SIAM J. MATH. ANAL.
Vol. 8, No. 1, February 1977

ON A MIXED BOUNDARY VALUE PROBLEM FOR THE REDUCED WAVE EQUATION ON A SPHERE*

H. L. JOHNSON†

Abstract. The paper considers the mixed boundary value problem $\nabla^2 w + \lambda^2 w = 0$, $0 \leqq \rho < 1$, $0 \leqq \varphi \leqq \pi$, $w(1, \varphi) = H_1(\varphi)$, $0 \leqq \varphi < \alpha$; $(\partial w/\partial \rho)(1, \varphi) = H_2(\varphi)$, $\alpha < \varphi \leqq \pi$, for an axial symmetric potential $w = w(\rho, \varphi)$ in a unit sphere. The problem is transformed into a linear integral equation of the second kind with a weakly singular kernel. An orthogonality condition is imposed on the given boundary functions H_1 and H_2 in order that the solution have a desired degree of regularity. The paper extends some previous work of the author and is related to some earlier work of W. D. Collins on dual series equations.

1. Introduction. This paper extends the integral equation formulation of [5] to the mixed boundary value problem

$$\text{I} \begin{cases} \nabla^2 w + \lambda^2 w = \dfrac{1}{\rho^2}\left[\dfrac{\partial w}{\partial \rho}\left(\rho^2 \dfrac{\partial w}{\partial \rho}\right) + \dfrac{1}{\sin \varphi}\dfrac{\partial}{\partial \varphi}\left(\sin \varphi \dfrac{\partial w}{\partial \varphi}\right)\right] + \lambda^2 w = 0, & (1.1) \\ \qquad (\rho, \varphi) \in G^+ = \{(\rho, \varphi) | 0 \leqq \rho < 1, 0 \leqq \varphi < \pi\}, & \\ w(1, \varphi) = H_1(\varphi), \qquad \varphi \in S_1 = \{\varphi | 0 \leqq \varphi < \alpha\}, & (1.2) \\ \dfrac{\partial w}{\partial \rho}(1, \varphi) = H_2(\varphi), \qquad \varphi \in S_2 = \{\varphi | \alpha < \varphi \leqq \pi\}. & (1.3) \end{cases}$$

The continuity properties of the prescribed functions H_1 and H_2 are described in § 6.

In a series of papers, [2], [3], and [4], W. D. Collins has shown how certain diffraction and electrostatic problems with mixed boundary conditions on a spherical boundary can be reduced to the solution of a system of dual series, and how these series can be transformed into an integral equation of the second kind. For axial symmetric problems, the series studied by Collins, in his notation, take the form

$$\sum_{n=0}^{\infty} (1 + H_n) C_n P_n(\cos \varphi) = f(\varphi), \qquad 0 \leqq \varphi < \alpha, \tag{1.4}$$

$$\sum_{n=0}^{\infty} (2n + 1) C_n P_n(\cos \varphi) = g(\varphi), \qquad \alpha < \varphi \leqq \pi, \tag{1.4}'$$

where $f(\varphi)$ and $g(\varphi)$ are prescribed functions.

In his examples, $H_n = H_n(\lambda)$, $n \geqq 0$, are described entire functions of the wave number λ with the property that $H_n = O(1/n)$.

It would be natural to seek a solution w of the boundary value problem I in the series form

$$w(\rho, \varphi) = \sum_{n=0}^{\infty} B_n j_n(\lambda \rho) P_n(\cos \varphi),$$

* Received by the editors February 20, 1975, and in revised form August 14, 1975.

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

where $j_n(\lambda\rho)$, $n \geqq 0$, are spherical Bessel functions, and to rewrite the boundary conditions (1.2), (1.3) in the form (1.4) and (1.4)′. We do this and find that

$$H_n = H_n(\lambda) = \frac{(2n+1)j_n(\lambda)}{j_n'(\lambda)} - 1.$$

Since these H_n have poles at the zeros of j_n', the kernel of the integral equation generated by Collins' method would, as a function of λ, necessarily have singularities at all of these poles.

The method presented in this paper develops a pair of Fredholm integral equations of the second kind for which the kernel is an entire function of λ. This formulation is based on two principal ideas. The first idea is that a solution of (1.1) can be written in the form of an integral that depends on an unprescribed harmonic function of two variables. The second idea can be viewed as an extension of Mehler's formula for Legendre polynomials, and is expressed as Theorem 1 in the paper. Once these facts have been applied to the boundary value problem I, the course of the ensuing analysis is, for the most part, determined.

To set the stage for the analysis, we note that the change of variables, $x = \rho \cos \varphi$, $y = \rho \sin \varphi$, $W = W(x, y) = w(\rho, \varphi)$, transforms (1.1) into

$$\frac{\partial^2 W}{\partial x^2} + \frac{\partial^2 W}{\partial y^2} + \frac{1}{y}\frac{\partial W}{\partial y} + \lambda^2 W = 0. \tag{1.5}$$

It is known [7] that every solution W of (1.5) can be written as

$$W = W(x, y) = \frac{2}{\pi}\int_0^y \frac{\cos(\lambda\sqrt{y^2 - s^2})u(x, s)\, ds}{\sqrt{y^2 - s^2}} \tag{1.6}$$

for some two-dimensional harmonic function $u = u(x, s)$, $s > 0$. The function u can be continued harmonically into $x^2 + s^2 < 1$ as an even function of s. By setting $s = y \cos \theta$ and using $u(x, -s) = u(x, s)$, (1.6) can be rewritten as

$$w(\rho, \varphi) = \frac{1}{\pi}\int_0^\pi \cos(\lambda\rho \sin \varphi \sin \theta) u(\rho \cos \varphi, \rho \sin \varphi \cos \theta)\, d\theta, \tag{1.7}$$

and

$$\begin{aligned} \frac{\partial w}{\partial \rho}(\rho, \varphi) = &-\frac{1}{\pi}\int_0^\pi \lambda \sin \varphi \sin \theta \sin(\lambda\rho \sin \varphi \sin \theta) \\ &\cdot u(\rho \cos \varphi, \rho \sin \varphi \cos \theta)\, d\theta \\ &+\frac{1}{\pi}\int_0^\pi \cos(\lambda\rho \sin \varphi \sin \theta)\frac{1}{\rho}\operatorname{Re}\left(\frac{df}{dz}z\right) d\theta, \end{aligned} \tag{1.8}$$

where

$$f(z) = u(x, s) + iv(x, s) \tag{1.9}$$

is analytic in $|z| < 1$ and

$$z = \rho \cos \varphi + i\rho \sin \varphi \cos \theta. \tag{1.10}$$

2. A basic theorem and corollary. Before taking limits of $w(\rho, \varphi)$ and $(\partial w/\partial\rho)(\rho, \varphi)$ as $\rho \to 1^-$, it is convenient to put each term on the right-hand side of (1.7) and (1.8) in the form

$$\text{Re}\left(\frac{1}{\pi}\int_0^{\pi} g(z)\, d\theta\right), \tag{2.1}$$

where g is analytic, $\text{Re}\,(g(\bar{z})) = \text{Re}\,(g(z))$ in $|z|<1$, and z is related to θ by (1.10).

Let $S = \{\varphi | 0 < \varphi < \alpha\} \cup \{\varphi | \alpha < \varphi < \pi\}$; let $H^{\beta}(S)$ be the class of Hölder-continuous functions of order β; and let $L_2(S)$ be the class of Lebesgue-integrable functions on S. The $\lim_{\rho\to 1^-}$ of the integral (2.1) is described by the following theorem.

THEOREM 1. *If*

(i) $g = g(z)$ *is analytic in* $|z|<1$,

(ii) $g(e^{is}) \in H^{\beta}(S) \cap L_2(S)$,

(iii) $\text{Re}\,(g(\bar{z})) = \text{Re}\,(g(z))$, *then*

$$\lim_{\rho\to 1^-} \text{Re}\left(\frac{1}{\pi}\int_0^{\pi} g(\rho\cos\varphi + i\rho\sin\varphi\cos\theta)\, d\theta\right) = A(g)(\varphi), \tag{2.2}$$

where

$$A(g)(\varphi) = \frac{F(g)(\varphi) + G(g)(\varphi)}{2}, \tag{2.3}$$

$$F(g)(\varphi) = \text{Re}\left(\frac{1}{\pi}\int_0^{\varphi} \frac{g(e^{is})\, e^{is/2}\, ds}{K(\varphi, s)}\right), \tag{2.4}$$

$$G(g)(\varphi) = \text{Im}\left(\frac{1}{\pi}\int_{\varphi}^{\pi} \frac{g(e^{is})\, e^{is/2}}{K(\varphi, s)}\, ds\right), \tag{2.5}$$

$$K(\varphi, s) = \sqrt{|\sin^2(\varphi/2) - \sin^2(s/2)|}. \tag{2.6}$$

To prove Theorem 1, we may insert

$$g(z) = \frac{1}{2\pi i}\oint_{|\delta|=1} \frac{g(\delta)d\delta}{(\delta - z)}$$

into $\int_0^{\pi} g(z)d\theta$, interchange orders of integration, carry out the inner integration, and use the dominating convergence theorem to pass the limit under the integral sign.

COROLLARY 1. *If* $f(e^{is}) = u(s) + iv(s) = \overline{f(e^{-is})}$, $s \in S$, *are the boundary values of an analytic function in* $|z|<1$ *with* $f(e^{is}) \in H^{\beta}(S) \cap L_2(S)$, *then*

$$A(f)(\varphi) = F(f)(\varphi) = G(f)(\varphi), \qquad \varphi \in S. \tag{2.7}$$

Proof. If $f = f_k = e^{ik\varphi}$, then

$$F(f_k)(\varphi) = G(f_k)(\varphi) \tag{2.8}$$

is a consequence of Mehler's formula [8, p. 57]. In the more general case, the partial sums of the Fourier series of f exist and can be written as $s_n(\varphi) =$

$\sum_{k=0}^{n} a_k e^{ik\varphi}$, where the a_k are real. Let $\sigma_n = (1/(n+1)) \sum_{k=0}^{n} s_k(\varphi)$. An immediate consequence of (2.8) is

$$F(\sigma_n)(\varphi) = G(\sigma_n)(\varphi). \tag{2.9}$$

It is known that $\|f-\sigma_n\|_{L_2(S)} \to 0$ as $n \to \infty$, and that $\lim_{n\to\infty} \sigma_n(\varphi) = f(e^{i\varphi})$ uniformly on closed subintervals of S, where f is continuous [9, p. 45]. Let $\varphi \in S$, $M = M(\varphi) = \max(\csc(\varphi/2), \sec(\varphi/2), 1/(K(\varphi, \alpha)))$. There exists a set $E \subset \bar{S}$ containing the points 0, α, and π such that $\bar{S} - E$ is closed and $1/(K(\varphi, s)) \leqq 2M$ for $s \in E$. Equation (2.9) and the triangle inequality imply that

$$|F(f)(\varphi) - G(f)(\varphi)| \leqq R_n = |F(f-\sigma_n)(\varphi)| + |G(f-\sigma_n)(\varphi)|. \tag{2.10}$$

Moreover, from the definitions of M and E, and the form of the operators F and G, it follows that

$$R_n \leqq 4M \int_E (|\mathrm{Re}(f-\sigma n)| + |\mathrm{Im}(f-\sigma_n)|)\, ds + 2 \int_{S-E} \frac{(|\mathrm{Re}(f-\sigma_n)| + |\mathrm{Im}(f-\sigma_n)|)\, ds}{K(\varphi, s)}.$$

The L_2 convergence on S and the uniform convergence on $\bar{S} - E$ of σ_n to f imply that $\lim_{n\to\infty} R_n = 0$. Q.E.D.

The significance of Corollary 1 is that it allows one to use the operator A in its F form, as given by (2.4), when working with the first boundary condition (1.2) and to use A in its G form, as given by (2.5), when working with the second boundary condition (1.3).

3. Transformations of the first boundary condition. To transform (1.7) into the form (2.1), we start with the well-known identity

$$\cos(x \sin\theta) = J_0(x) + 2 \sum_{n=1}^{\infty} J_{2n}(x) \cos(2n\theta), \tag{3.1}$$

where $J_{2n}(x)$ are Bessel functions of the first kind. It is known that the right-hand side of (3.1) converges uniformly for $|\gamma x| \leqq 1$.

Again, let $f = u + iv$ be analytic in $|z| < 1$, and set $z = \rho \cos\varphi + i\rho \sin\varphi \cos\theta$. Let

$$Z_1(z, \rho) = \frac{z - \rho\cos\varphi}{i\rho \sin\varphi} = \cos\theta. \tag{3.2}$$

The binomial identity

$$\cos(2n\theta) = \mathrm{Re}(\cos\theta + i\sin\theta)^{2n} = \sum_{j=0}^{n} \binom{2n}{2j} \cos^{2n-2j}(\theta)(-1)^j (1-\cos^2\theta)^j$$

permits (1.7) to be written as $w(\rho, \varphi) = \mathrm{Re}\,(1/\pi) \int_0^{\pi} g_1(z)\, d\theta)$ with

$$g_1(z) = \left[J_0(\lambda\rho\sin\varphi) + 2 \sum_{n=1}^{\infty} J_{2n}(\lambda\rho\sin\varphi) \left(\sum_{j=0}^{n} \binom{2n}{2j} (-1)^j Z_1^{2n-2j}(z, \rho) \cdot (1 - Z_1^2(z, \rho))^j \right) \right] f(z). \tag{3.3}$$

Since the series (3.3) converges uniformly for $\frac{1}{2} \leqq \rho \leqq 1$, Theorem 1 can be applied term by term, and the identity

$$\frac{(Z_1+iZ_2)^{2n}+(Z_1-iZ_2)^{2n}}{2} = \sum_{j=0}^{n} \binom{2n}{2j} (-1)^j Z_1^{2n-2j} Z_2^{2j} \tag{3.4}$$

can be used to obtain

$$w(1,\varphi) = A\left[\left(J_0(\lambda \sin\varphi) + \sum_{n=1}^{\infty} J_{2n}(\lambda \sin\varphi)((Z_1+iZ_2)^{2n}+(Z_1-iZ_2)^{2n})\right) f(e^{is})\right](\varphi) \tag{3.5}$$

with

$$Z_1 = \frac{e^{is}-\cos\varphi}{i \sin\varphi}, \tag{3.6}$$

$$Z_2 = \sqrt{1-Z_1^2} = \frac{\sqrt{2(\cos s - \cos\varphi)}\, e^{is/2}}{\sin\varphi}. \tag{3.7}$$

Setting $t = Z_1 + iZ_2$, it follows that $t^{-1} = Z_1 - iZ_2$. The identities

$$e^{(x/2)(t-t^{-1})} = \sum_{n=-\infty}^{\infty} J_n(x) t^n, \tag{3.8}$$

$$(-1)^n J_n(x) = J_{-n}(x) = J_n(-x) \tag{3.9}$$

imply that

$$\cosh\left(\frac{x}{2}(t-t^{-1})\right) = J_0(x) + \sum_{n=1}^{\infty} J_{2n}(x)(t^{2n}+t^{-2n}). \tag{3.10}$$

These facts permit (3.5) to be written as

$$w(1,\varphi) = A\left[\cosh\left(\frac{\lambda \sin\varphi}{2}(t-t^{-1})\right) f(e^{is})\right](\varphi); \tag{3.11}$$

furthermore,

$$t - t^{-1} = 2iZ_2 = \frac{4i\sqrt{K_1(\varphi,s)}\, e^{is/2}}{\sin\varphi}, \tag{3.12}$$

where

$$K_1(\varphi,s) = \sin^2(\varphi/2) - \sin^2(s/2). \tag{3.13}$$

Since

$$\cosh(iz) = \cos(z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!},$$

equation (3.11) becomes

$$\begin{aligned} w(1,\varphi) &= A[\cos(2\lambda K(\varphi,s)\, e^{is/2}) f(e^{is})](\varphi) \\ &= \sum_{n=0}^{\infty} \frac{A[(2\lambda)^{2n}(-1)^n K^{2n}(\varphi,s)\, e^{ins} f(e^{is})]}{(2n)!}(\varphi) = H_1(\varphi), \qquad \varphi \in S_1. \end{aligned} \tag{3.14}$$

Equation (3.14) can be partially inverted by setting $\varphi = \eta$, multiplying it by $(\sin \eta/2)(1/K(\varphi, \eta))$ and integrating on η from 0 to φ. The integrals generated by performing these operations on (3.14) take the form

$$\int_0^{\varphi} \frac{\sin \eta}{2} \frac{1}{K(\varphi, \eta)} \operatorname{Re}\left(\frac{1}{\pi} \int_0^{\eta} (K(\eta, s))^{n-1/2} e^{ins} f(e^{is/2})\, ds\, d\eta\right)$$

$$= \operatorname{Re}\left(\int_0^{\varphi} f(e^{is})\, e^{is/2}\, e^{ins}\left(\frac{1}{\pi} \int_s^{\varphi} \frac{\sin \eta}{2} \frac{(K(\eta, s))^{n-1/2}}{K(\varphi, \eta)}\, d\eta\right) ds\right)$$

$$= \operatorname{Re}\left(\int_0^{\varphi} f_1(s)\, e^{ins} (K_1(\varphi, s))^n \left(\frac{1}{\pi} B(n+1/2, 1/2)\right) ds\right),$$

where

$$f_1(s) = u_1(s) + iv_1(s) = f(e^{is})\, e^{is/2} \tag{3.15}$$

and

$$\frac{B(n+1/2, 1/2)}{\pi} = \frac{\Gamma(1/2)\Gamma(n+1/2)}{\pi \Gamma(n+1)} = \frac{2}{4^n} \frac{(2n-1)!}{(n-1)!n!}.$$

The integrated form of (3.14) thus becomes

$$\operatorname{Re}\left(\int_0^{\varphi} f_1(s)\left(\sum_{n=0}^{\infty} (-1)^n \frac{\lambda^{2n} (K(\varphi, s))^{2n} e^{ins}}{(n!)^2}\right) ds\right)$$

$$= \operatorname{Re}\left(\int_0^{\varphi} f_1(s) J_0(z(\varphi, s))\, ds\right) = h_1(\varphi), \qquad \varphi \in S_1, \tag{3.16}$$

where

$$z(\varphi, s) = 2\lambda K(\varphi, s)\, e^{is/2}, \tag{3.17}$$

$$h_1(\varphi) = \int_0^{\varphi} \frac{\sin \eta}{2} \frac{H_1(\eta)}{K(\varphi, \eta)}\, d\eta. \tag{3.18}$$

Upon differentiating (3.16) and using some well-known properties of Bessel functions, one can show that

$$u_1(\varphi) = h_1'(\varphi) + \lambda^2 \sin \varphi \operatorname{Re}\left(\int_0^{\varphi} f_1(s)\, e^{is}\left(\frac{J_1(z(\varphi, s))}{z(\varphi, s)}\right) ds\right), \qquad \varphi \in S_1. \tag{3.19}$$

4. Transformation of the second boundary condition. The second term of (1.8) is equal to (1.7) with $u = \operatorname{Re}(f)$ replaced by $\operatorname{Re}(z(df/dz))$. The limit of this term is

$$A\left[\cosh\left(\lambda \frac{\sin \varphi}{2}(t - t^{-1})\right) e^{is} f'(e^{is})\right](\varphi). \tag{4.1}$$

It is now appropriate to take the operator A in its G form. In this case, the expression

$$\frac{\lambda \sin(\varphi)}{2}(t - t^{-1}) = 2\lambda \sqrt{K_1(s, \varphi)}\, e^{is/2} = 2\lambda K(\varphi, s)\, e^{is/2} = z(\varphi, s), \tag{4.2}$$

and

$$e^{is}f'(e^{is}) = -\frac{id}{ds}f(e^{is}).$$

To modify the first term in (1.8), we start with the identity

$$\sin(x\sin\theta) = 2\sum_{n=1}^{\infty} J_{2n-1}(x)\sin((2n-1)\theta). \tag{4.3}$$

It follows that

$$\sin\theta\sin(\lambda\rho\sin\varphi\sin\theta) = J_1(\lambda\rho\sin\varphi) - 2\sum_{n=1}^{\infty} J'_{2n}(\lambda\rho\sin\varphi)\cos(2n\theta).$$

Proceeding now as in § 3, we write the term in (1.8) as

$$\operatorname{Re}\left(\frac{1}{\pi}\int_0^{\pi} g_2(z)\,d\theta\right), \tag{4.4}$$

where

$$\begin{aligned} g_2(z) = -\lambda\sin\varphi\Big[J_1(\lambda\rho\sin\varphi) - 2\sum_{n=1}^{\infty} J'_{2n}(\lambda\rho\sin\varphi)\Big(\sum_{j=0}^{n}\binom{2n}{2j} \\ \cdot(-1)^j Z_1^{2n-2j}(z,\rho)(1-Z_1^2(z,\rho))^j\Big)\Big], \end{aligned} \tag{4.5}$$

where $Z_1(z,\rho)$ is defined by (3.2). Theorem 1 applied to (4.4) yields

$$A\Big[(-\lambda\sin\varphi)\Big(J_1(\lambda\sin\varphi) - \sum_{n=1}^{\infty} J'_{2n}(\lambda\sin\varphi)(t^{2n}+t^{-2n})\Big)(e^{is})\Big](\varphi), \tag{4.6}$$

where $t = Z_1 + iZ_2$ is defined as in § 3.

The above series of Bessel functions may be written more completely by using the identity

$$J_1(x) - \sum_{n=1}^{\infty} J'_{2n}(x)(t^{2n}+t^{-2n}) = -\frac{(t-t^{-1})}{2}\sinh\left(\frac{x}{2}(t-t^{-1})\right). \tag{4.7}$$

To verify this identity, observe that (3.8) leads to

$$\begin{aligned} \sinh\left(\frac{x}{2}(t-t^{-1})\right) &= \sum_{n=-\infty}^{\infty} J_{2n+1}(x)t^{2n+1} \\ &= \sum_{n=1}^{\infty} J_{2n+1}(x)t^{2n+1} + J_1(x)t + \sum_{n=1}^{\infty} J_{-(2n-1)}(x)t^{-(2n-1)}. \end{aligned} \tag{4.8}$$

Equation (4.7) is obtained by regrouping the terms in this series and using the identity $J_{p+1} - J_{p-1} = 2J'_p$.

In our application of (4.7), $x = \lambda\sin\varphi$ and $\lambda((\sin\varphi)/2)(t-t^{-1}) = z(\varphi,s)$. Hence the limit of the first term in the second boundary condition becomes

$$A[z(\varphi,s)\sinh(z(\varphi,s))f(e^{is})](\varphi). \tag{4.9}$$

The second boundary condition is the sum of the expressions (4.1) and (4.9); this condition can now be written as

$$G[z(\varphi, s)\sinh(z(\varphi, s))f(e^{is})](\varphi)+G[\cosh(z(\varphi, s))\left(-i\frac{df}{ds}\right)](\varphi)=H_2(\varphi). \tag{4.10}$$

Equation (4.10) may be partially inverted if we set $\varphi=\eta$, multiply by $((\sin\eta)/2)(1/K(\eta, \varphi))$ integrate on η from φ to π, and use the power series form of $\sinh(z(\eta, s))$ and $\cosh(z(\eta, s))$ to obtain

$$\begin{aligned} h_2(\varphi) \doteq \int_\varphi^\pi \frac{\sin\eta}{2K(\eta, \varphi)}H_2(\eta)\,d\eta &= \operatorname{Im}\left[\int_\varphi^\pi f_1(s)z(\varphi, s)I_1(z(\varphi, s))\,ds\right] \\ &\quad -\operatorname{Re}\left[\int_\varphi^\pi \frac{df}{ds}1(s)I_0(z(\varphi, s))\,ds\right]-\operatorname{Im}\left[\int_\varphi^\pi \frac{f_1(s)}{2}I_0(z(\varphi, s))\,ds\right], \end{aligned} \tag{4.11}$$

where $f_1(s)$ is defined by (3.15) and the terms $I_k(x)$ are modified Bessel functions. The second term on the right-hand side of (4.11) is integrated by parts and $\operatorname{Re}(f_1(\pi))=0$ is used to obtain

$$\begin{aligned} u_1(\varphi) = h_2(\varphi) &- \operatorname{Re}\left(\lambda^2\int_\varphi^\pi f_1(s)e^{is}\sin(s)\frac{I_1(z(\varphi, s))\,ds}{z(\varphi, s)}\right) \\ &-\frac{\operatorname{Im}}{2}\left(\int_\varphi^\pi f_1(s)(z(\varphi, s)I_1(z(\varphi, s))-I_0(z(\varphi, s)))\,ds\right), \qquad \varphi\in S_2. \end{aligned} \tag{4.12}$$

5. The integral equation for v_1. Equations (3.19) and (4.12) give the function u_1 in terms of $f_1(s)=u_1(s)+iv_1(s)$. Hence, we still need to develop an integral equation for v_1 on $S_1\cup S_2$. To do this, we assume that $u(\varphi)=u(1, \varphi)$ and $v(\varphi)=v(1, \varphi)$ are representable in the Fourier series $u(\varphi)=\sum_{n=0}^{\infty}a_n\cos(n\varphi)$, $v(\varphi)=\sum_{n=1}^{\infty}a_n\sin(n\varphi)$.

It follows that

$$u_1(\varphi)=\sum_{n=0}^{\infty}a_n\cos((n+\tfrac{1}{2})\varphi), \tag{5.1}$$

$$v_1(\varphi)=\sum_{n=0}^{\infty}a_n\sin((n+\tfrac{1}{2})\varphi) \tag{5.2}$$

with

$$a_n=\frac{2}{\pi}\int_0^\pi u_1(t)\cos((n+\tfrac{1}{2})t)\,dt.$$

Thus

$$\begin{aligned} v_1(\varphi)=\sum_{n=0}^{\infty}\Bigg[\frac{2}{\pi}\Bigg(&\int_0^\alpha u_1(t)\cos((n+\tfrac{1}{2})t)\,dt+\int_\alpha^\pi u_1(t)\cos((n+\tfrac{1}{2})t)\,dt\Bigg) \\ &\cdot\sin((n+\tfrac{1}{2})\varphi)\Bigg]. \end{aligned} \tag{5.3}$$

A formal integration by parts of the integrals contained in (5.3) yields

$$v_1(\varphi)=-2[u_1](\alpha)k(\varphi, \alpha)-2\int_0^\pi u'1(t)k(\varphi, t)\,dt, \tag{5.4}$$

where

$$k(\varphi, t) = -\frac{1}{2\pi} \ln \left| \frac{\sin(\varphi/2) - \sin(t/2)}{\sin(\varphi/2) + \sin(t/2)} \right| = \frac{2}{\pi} \sum_{n=0}^{\infty} \frac{\sin((n+\frac{1}{2})\varphi) \sin((n+\frac{1}{2})t)}{(2n+1)} \tag{5.5}$$

and

$$[u_1](\alpha) = u_1(\alpha + 0) - u_1(\alpha - 0). \tag{5.6}$$

Since we have formally applied Theorem 1 and Corollary 1 to the function with boundary values $v'(\varphi) - iu'(\varphi)$, it is necessary, under our proof, that $u(\varphi)$ and $v(\varphi)$, and hence $u_1(\varphi)$ and $v_1(\varphi)$, not have a logarithmic singularity at $\varphi = \alpha$. Thus we require that

$$[u_1](\alpha) = 0. \tag{5.7}$$

The implications of this condition are discussed in § 8.

To obtain an integral equation from (5.4), we need expressions for u_1'. By formally differentiating (3.19) and (4.12), one obtains

$$\begin{aligned} u_1'(\varphi) = h_1''(\varphi) + \operatorname{Re}\Big[& \frac{\lambda^2 \sin(\varphi) f_1(\varphi) e^{i\varphi}}{2} \\ & + \int_0^{\varphi} \lambda^2 \cos(\varphi) f_1(s) e^{is} \frac{J_1(z(\varphi, s))}{z(\varphi, s)}\, ds \\ & - \int_0^{\varphi} \lambda^4 \sin^2(\varphi) f_1(s) e^{2is} \frac{J_2(z(\varphi, s))}{(z(\varphi, s))^2}\, ds \Big], \qquad \varphi \in S_1, \end{aligned} \tag{5.8}$$

$$\begin{aligned} u_1'(\varphi) = h_2'(\varphi) - \frac{v_1(\varphi)}{2} + \operatorname{Re} \frac{\lambda^2 \sin(\varphi) f_1(\varphi) e^{i\varphi}}{2}, \\ + \operatorname{Im}\left(\frac{\lambda^2 \sin(\varphi)}{4} \int_{\varphi}^{\pi} f_1(s) e^{is} (I_0(z(\varphi, s)) + I_2(z(\varphi, s)))\, ds \right) \\ + \operatorname{Re}\left(\int_{\varphi}^{\pi} \lambda^4 \sin(\varphi) \sin(s) f_1(s) e^{2is} \frac{I_2(z(\varphi, s))}{(z(\varphi, s))^2}\, ds \right), \qquad \varphi \in S_2. \end{aligned} \tag{5.9}$$

Inserting (5.8) and (5.9) into (5.4) with $[u_1](\alpha) = 0$, one obtains

$$\begin{aligned} v_1(\varphi) = F_2(\varphi) + \int_{\alpha}^{\pi} k(\varphi, s) v_1(s)\, ds + \operatorname{Re}\left(\int_0^{\pi} f_1(s) \left(\sum_{j=1}^{3} K_{2,j}(\varphi, s) \right) ds \right), \\ \varphi \in S_1 \cup S_2, \end{aligned} \tag{5.10}$$

where

$$F_2(\varphi) = -\int_0^{\alpha} k(\varphi, s) 2h_1''(s)\, ds - \int_{\alpha}^{\pi} k(\varphi, s) 2h_2'(s)\, ds, \tag{5.11}$$

$$K_{2,1}(\varphi, s) = -\lambda^2 \sin(s) e^{is} k(\varphi, s), \tag{5.12}$$

$$K_{2,2}(\varphi, s) = \begin{cases} -\lambda^2 e^{is} \int_s^{\alpha} k(\varphi, t) \cos(t) \\ \qquad (J_2(z(t,s)) + J_0(z(t,s))) \, dt, & 0 < s < \alpha, \\ i\lambda^2 e^{is} \int_{\alpha}^{s} k(\varphi, t) \dfrac{\sin(t)}{2} (I_2(z(t,s)) + I_0(z(t,s))) \, dt, & \alpha < s < \pi, \end{cases} \tag{5.13}$$

$$K_{2,3}(\varphi, s) = \begin{cases} 2(\lambda^2 e^{is})^2 \int_s^{\alpha} k(\varphi, t) \sin^2(t) \dfrac{J_2(z(t,s))}{(z(t,s))^2} \, dt, & 0 < s < \alpha, \\ -2(\lambda^2 e^{is})^2 \int_{\alpha}^{s} k(\varphi, t) \sin(s) \sin(t) \dfrac{I_2(z(t,s))}{(z(t,s))^2} \, dt, & \alpha < s < \pi. \end{cases} \tag{5.14}$$

Hereafter we shall refer to equations (3.19), (4.12), and (5.10) as the *derived integral equation.*

6. Continuity properties of solutions of the derived integral equations. At this point we have formally derived a system of two linear integral equations of the second kind. To write the second boundary condition as (4.10), we have assumed that $v'(\varphi) - iu'(\varphi) \in H^{\beta}(S) \cap L_2(S)$ for some $\beta > 0$. Our next goal is to show that v' and u' will be in this class if H_1 and H_2 are suitably smooth. Toward this end, we note the following smoothness properties of h_1, h_2, and F_2.

LEMMA 1. *If* $H_1 \in C^2(\bar{S}_1)$, *then* $h_1'' \in H^{1/2}(\bar{S}_1)$.
If $H_2 \in C^1(\bar{S}_2)$, *then* $h_2' \in H^{1/2}(\bar{S}_2)$.

This lemma is proved by integrating by parts the integrals (3.18) and (4.11), differentiating the results, and estimating the difference of the final integrals to prove that they are in the stated Hölder class.

COROLLARY 2. *Under the hypotheses of Lemma* 1, $F_2 \in C(\bar{S}) \cap H^{1/2}(S)$ *and* $F_2' \in H^{1/2}(S)$.

The property $F_2' \in H^{1/2}(S)$ is a consequence of Lemma 1 and the well-known fact that Cauchy singular integrals map $H^{\beta}(S)$ into $H^{\beta}(S)$ [6, p. 46].

The expected smoothness properties of u_1 and v_1 are described in the following theorem.

THEOREM 2. *If* $H_1 \in C^2(\bar{S}_1)$ *and* $H_2 \in C^1(\bar{S}_2)$, *then every solution* (u_1, v_1) *of the derived integral equations which is an element of* $L_2(S) \times L_2(S)$ *has the property that each of the functions* u_1, v_1, u_1', *and* v_1' *are elements of* $H^{1/2}(S) \cap L_2(S)$.

The proof of Theorem 2 consists of a sequence of bootstrap arguments. First notice that u_1 is continuous except possibly at $\varphi = \alpha$. The terms making up v_1 are all integrals of the form $\int_0^{\pi} k(\varphi, s) f(s) \, ds$ with $f \in L_2(S)$. The absolute continuity of the Lebesgue integral and the continuity property of the logarithm function imply that these integrals are continuous on $\bar{S}$. Next, it is easy to prove that an integral of the form $F(x) = \int_a^b \log|x - s| f(s) \, ds$ with $f(s)$ piecewise continuous on $[a, b]$ is in $H^{\beta}((a, b))$ for any $\beta < 1$. It follows from Corollary 2 and the above remarks that both u_1 and v_1 are elements of $H^{1/2}(S)$. The formal differentiation that was carried

out to derive (5.8) from (3.19) and (5.9) from (4.12) is now justified. Furthermore, it follows from Lemma 1, the fact that $f_1 = u_1 + iv_1 \in H^{1/2}(S)$, and the differentiability properties of $J_k(x)$ and $I_k(x)$ that $u_1' \in H^{1/2}(S) \cap L_2(S)$. Again since $f_1 \in H^{1/2}(S)$, it follows that v_1 exists as a Cauchy singular integral. The property that Cauchy singular integrals map H^β into H^β, together with Corollary 2, implies that $v_1' \in H^{1/2}(S)$. The logarithmic character of Cauchy singular integrals at the endpoints of the interval of integration implies that $v_1' \in L_2(S)$. Q.E.D.

7. The case of $\lambda = 0$. For $\lambda = 0$, the derived integral equations become

$$v_1(\varphi) = F_2(\varphi) + \int_\alpha^\pi k(\varphi, s) v_1(s)\, ds, \qquad \varphi \in S_1 \cup S_2; \tag{7.1}$$

$$u_1(\varphi) = \begin{cases} h_1'(\varphi), & \varphi \in s_1 \\ h_2(\varphi) + \displaystyle\int_\varphi^\pi \frac{v_1(s)\, ds}{2}, & \varphi \in S_2, \end{cases} \tag{7.2}$$

where $k(\varphi, s)$ is the symmetric positive kernel defined by (5.5). The domain of $k(\varphi, s)$ as used in the integral equation (7.1) is dependent on α. To emphasize this fact, we write

$$k_\alpha = k_\alpha(\varphi, s) = k(\varphi, s), \qquad \alpha \leqq \varphi \leqq \pi, \quad \alpha \leqq s \leqq \pi. \tag{7.3}$$

The solution of (7.1) depends on the characteristic values, c.v., of the kernel k_α. Let μ_α denote the lowest c.v. of k_α. Our next goal is to show that

$$\mu_\alpha > 1, \quad \alpha > 0. \tag{7.4}$$

To prove (7.4), we first introduce the extended kernel

$$\hat{k}_\alpha = \hat{k}_\alpha(s, t) = \begin{cases} k(s, t), & \alpha \leqq s \leqq \pi, \quad \alpha \leqq t \leqq \pi, \\ 0, & 0 \leqq s < \alpha, \quad 0 \leqq t \leqq \pi, \\ 0, & 0 \leqq s \leqq \pi, \quad 0 \leqq t < \alpha, \end{cases} \tag{7.5}$$

and study some of its properties. $\hat{k}_\alpha$ is a symmetric nonnegative kernel. It is known [1, p. 285] that the lowest c.v. of a nonnegative kernel has at most one independent characteristic function, c.f., and that the c.f. corresponding to the lowest c.v. can be taken to be nonnegative. Let $\hat{\mu}_\alpha$ denote the lowest c.v. of $\hat{k}_\alpha$ and $\hat{\Phi}_\alpha = \hat{\Phi}_\alpha(t)$ its corresponding normalized nonnegative c.f.. $\hat{\Phi}_\alpha$ satisfies

$$\hat{\Phi}_\alpha(t) = \hat{\mu}_\alpha \int_0^\pi \hat{k}_\alpha(t, s) \hat{\Phi}_\alpha(s)\, ds, \qquad 0 \leqq t \leqq \pi. \tag{7.6}$$

From the definition of $\hat{k}_\alpha$, it follows that

$$\hat{\Phi}_\alpha(t) = 0, \quad 0 \leqq t < \alpha. \tag{7.7}$$

When $\alpha = 0$, the c.v.'s and the c.f.'s of $\hat{k}_\alpha$ are known, and in the above notation,

$$\hat{\mu}_0 = 1, \quad \hat{\Phi}_0(t) = \sqrt{(2/\pi)} \sin(t/2). \tag{7.8}$$

The Rayleigh quotient formulation of the lowest c.v. for the kernel $\hat{k}_0$ states that

$$1/1 = \sup_{(\Phi,\Phi)=1} (\hat{k}_0\Phi, \Phi) > (\hat{k}_0\hat{\Phi}_\alpha, \hat{\Phi}_\alpha). \tag{7.9}$$

This last inequality is strict, for if $(\hat{k}_0\hat{\Phi}_\alpha, \hat{\Phi}_\alpha) = 1$, then $\hat{\Phi}_\alpha$ would be a c.f. of $\hat{k}_0$ (see [1, p. 108]) and, since $\hat{\Phi}_\alpha$ is normalized and nonnegative, it would necessarily follow that $\hat{\Phi}_\alpha = \hat{\Phi}_0(t)$. This result is contradicted by (7.7) and (7.8).

Next, observe that

$$\frac{1}{\hat{\mu}_\alpha} = (\hat{k}_\alpha\hat{\Phi}_\alpha, \hat{\Phi}_\alpha) = \int_0^\pi \int_0^\pi \hat{k}_\alpha(s, t)\hat{\Phi}_\alpha(t)\hat{\Phi}_\alpha(s)\, dt\, ds = (\hat{k}_0\hat{\Phi}_\alpha, \hat{\Phi}_\alpha). \tag{7.10}$$

The last equality is a consequence of (7.7). Taken together, equations (7.9) and (7.10) state that $1/\hat{\mu}_\alpha < 1$.

Another consequence of (7.6) is that

$$\hat{\Phi}_\alpha(t) = \hat{\mu}_\alpha \int_\alpha^\pi k_\alpha(t, s)\hat{\Phi}_\alpha(s)\, ds, \qquad \alpha < t < \pi. \tag{7.11}$$

Equation (7.11) asserts that $\hat{\mu}_\alpha$ is a c.v. of k_α and that $\hat{\Phi}_\alpha$ is the corresponding c.f. Since $\hat{\Phi}_\alpha(t) \geqq 0$, it is known [1, p. 288] that $\hat{\Phi}_\alpha$ must be associated with the smallest c.v. of k_α, i.e., $\hat{\mu}_\alpha = \mu_\alpha$. Q.E.D.

Since the lowest c.v. of k_α is greater than one, its resolvent kernel is

$$R_\alpha(t, s) = \sum_{\nu=1}^{\infty} k_\alpha^\nu(t, s), \tag{7.12}$$

where $k_\alpha^\nu(t, s)$ are the iterate kernels of k_α, and therefore the solution of (7.1) is

$$v_1(\varphi) = F_2(\varphi) + \int_\alpha^\pi R_\alpha(\varphi, s)F_2(s)\, ds, \qquad \varphi \in S_2. \tag{7.13}$$

8. The condition $[u_1](\alpha) = 0$. We have found that u_1 must be continuous at α if the integrals and limits that we have been working with are to exist. The continuity at α is not required if $\alpha = \pi$, the Dirichlet problem, nor if $\alpha = 0$, the Neumann problem. For all other values of α, the condition $[u_1](\alpha) = 0$ is effectively an orthogonality condition on the input functions H_1 and H_2 and some of their derivatives. To see this, we consider again the case of $\lambda = 0$.

For $\lambda = 0$, the function u_1 is given by equations (7.2), and continuity of u_1 at $\varphi = \alpha$ asserts that

$$u_1(\alpha - 0) = h_1'(\alpha) = u_1(\alpha + 0) = h_2(\alpha) + \int_\alpha^\pi \frac{v_1(t)\, dt}{2}, \tag{8.1}$$

which by (7.13) can be rewritten as

$$h_2(\alpha) - h_1'(\alpha) = -\frac{1}{2}\left(\int_\alpha^\pi \left(F_2(t) + \int_\alpha^\pi R_\alpha(t, \sigma)F_2(\sigma)\, d\sigma\right) dt\right). \tag{8.2}$$

If we insert the definition (5.11) for F_2 and use a Fredholm identity between a kernel and its resolvent, equation (8.2) becomes

$$h_2(\alpha)-h_1'(\alpha)=\int_0^{\alpha} h_1''(s)R_\alpha(s)\,ds+\int_\alpha^{\pi} h_2'(s)R_\alpha(s)\,ds, \tag{8.3}$$

where

$$R_\alpha(s)=\int_\alpha^{\pi} R_\alpha(s,t)\,dt. \tag{8.4}$$

Under the hypotheses $H_1\in C^2(\bar{S}_1)$ and $H_2\in C^1(\bar{S}_2)$, the integrals (3.18) and (4.11) which define h_1 and h_2 respectively, may be integrated by parts. The results are

$$h_2(\alpha)-h_1'(\alpha)=\int_\alpha^{\pi}\frac{\sin(s)H_2(s)\,ds}{2K(s,\alpha)}-H_1(0)\cos(\alpha/2)-\int_0^{\alpha}\frac{\sin(\alpha)H_1'(s)\,ds}{2K(\alpha,s)}, \tag{8.5}$$

$$\begin{aligned} h_1''(s)=&-\frac{1}{2}\sin\left(\frac{s}{2}\right)H_1(0)+\frac{\pi}{2}\cos\left(\frac{s}{2}\right)H_1'(0)\\ &-\int_0^{s}\left(\cos(s)\arcsin\left(\frac{\sin(t/2)}{\sin(s/2)}\right)\right.\\ &\left.+\frac{\cos^2(s/2)\sin(t/2)}{K(s,t)}\right)\frac{d}{dt}\left(\frac{H_1'(t)}{\cos(t/2)}\right)dt, \end{aligned} \tag{8.6}$$

$$h_2'(s)=-H_2(\pi)\sin\left(\frac{s}{2}\right)+\int_s^{\pi}\frac{\sin(s)H_2'(t)\,dt}{2K(t,s)}. \tag{8.7}$$

The substitution of (8.5), (8.6) and (8.7) into (8.3) leads to a restriction on H_1 and H_2.

In the general case of $\lambda\geqq 0$, we start with (3.19) written in the form

$$u_1(\varphi)=h_1(\varphi)-\int_0^{\varphi}M_1(\varphi,s)v_1(s)\,ds+\int_0^{\varphi}L_1(\varphi,s)u_1(s)\,ds,\qquad \varphi\in S_1, \tag{8.8}$$

where $L_1(\varphi,s)$ and $M_1(\varphi,s)$ are real and

$$L_1(\varphi,s)+iM_1(\varphi,s)=\lambda^2 e^{is}\sin\varphi\frac{J_1(z(\varphi,s))}{z(\varphi,s)}. \tag{8.9}$$

Equation (8.8) is a Volterra integral equation for $u_1(\varphi)$, $\varphi\in S_1$. The resolvent kernel for this integral equation exists in the usual Neumann series form for all values of λ and α. Designating the resolvent by $R_1(\varphi,s)$, we have

$$u_1(\varphi)=g_1(\varphi)+\int_0^{\varphi}N_1(\varphi,s)v_1(s)\,ds,\qquad \varphi\in S_1, \tag{8.10}$$

where

$$g_1(\varphi)=h_1'(\varphi)+\int_0^{\varphi}R_1(\varphi,t)h_1'(t)\,dt, \tag{8.11}$$

$$N_1(\varphi, s) = -\left(M_1(\varphi, s) + \int_s^{\varphi} R_1(\varphi, t) M_1(t, s)\, dt\right). \tag{8.12}$$

Equation (4.12) can be treated in a similar way. It can be written as

$$u_1(\varphi) = h_2(\varphi) - \int_{\varphi}^{\pi} M_2(\varphi, s) v_1(s)\, ds + \int_{\varphi}^{\pi} L_2(\varphi, s) u_1(s)\, ds, \qquad \varphi \in S_2, \tag{8.13}$$

where $L_2(\varphi, s)$ and $M_2(\varphi, s)$ are real and

$$L_2(\varphi, s) + iM_2(\varphi, s) = -\lambda^2 e^{is} \sin(s) \frac{I_1(z(\varphi, s))}{z(\varphi, s)} + \frac{i}{2}(z(\varphi, s) I_1(z(\varphi, s)) - I_0(z(\varphi, s))).$$

We designate the resolvent kernel of the Volterra equation (8.13) by $R_2(\varphi, s)$ and use it to write the solution of (8.13) as

$$u_1(\varphi) = g_2(\varphi) + \int_{\varphi}^{\pi} N_2(\varphi, s) v_1(s)\, ds, \qquad \varphi \in S_2, \tag{8.14}$$

where

$$g_2(\varphi) = h_2(\varphi) + \int_{\varphi}^{\pi} R_2(\varphi, t) h_2(t)\, dt, \tag{8.15}$$

$$N_2(\varphi, s) = -\left(M_2(\varphi, s) + \int_{\varphi}^{s} R_2(\varphi, t) M_2(t \cdot s)\, dt\right). \tag{8.16}$$

A single Fredholm integral equation can now be obtained for the function v_1 by setting

$$L_3(\varphi, s) + iM_3(\varphi, s) = \sum_{j=1}^{3} K_{2,j}(\varphi, s),$$

where the right-hand side is defined through equations (5.12), (5.13), and (5.14), and $L_3(\varphi, s)$ and $M_3(\varphi, s)$ are real kernels. Using this notation, equation (5.10) becomes

$$v_1(\varphi) = F_2(\varphi) + \int_{\alpha}^{\pi} k(\varphi, s) v_1(s)\, ds + \int_0^{\pi} L_3(\varphi, s) u_1(s) - M_3(\varphi, s) v_1(s)\, ds. \tag{8.17}$$

Substituting (8.10) and (8.14) into (8.17), one obtains

$$v_1(\varphi) = F_3(\varphi) + \int_0^{\pi} K_3(\varphi, s) v_1(s)\, ds, \qquad \varphi \in S_1 \cup S_2, \tag{8.18}$$

where

$$F_3(\varphi) = F_2(\varphi) + \int_0^{\alpha} L_3(\varphi, t) g_1(t)\, dt + \int_{\alpha}^{\pi} L_3(\varphi, t) g_2(t)\, dt, \tag{8.19}$$

$$K_3(\varphi, s) = -M_3(\varphi, s) + N_3(\varphi, s), \tag{8.20}$$

$$N_3(\varphi, s) = \begin{cases} \displaystyle\int_s^{\alpha} L_3(\varphi, t) N_1(t, s)\, dt, & 0 < s < \alpha, \\ \displaystyle k(\varphi, s) + \int_{\alpha}^{s} L_3(\varphi, t) N_2(t, s)\, dt, & \alpha < s < \pi. \end{cases} \tag{8.21}$$

Let $R_3(\varphi, s)$ denote the resolvent kernel of $K_3(\varphi, s)$. On the basis of our analysis in § 7, R_3 exists for λ sufficiently small. When this resolvent exists and the homogeneous form of (8.18) has only the trivial solution, the solution of (8.18) is

$$v_1(\varphi) = F_3(\varphi) + \int_0^{\pi} R_3(\varphi, t) F_3(t)\, dt. \tag{8.22}$$

As a result of equations (8.10) and (8.14), the continuity condition on u_1 at α becomes

$$g_1(\alpha) + \int_0^{\alpha} N_1(\alpha, s) v_1(s)\, ds = g_2(\alpha) + \int_{\alpha}^{\pi} N_2(\alpha, s) v_1(s)\, ds, \tag{8.23}$$

where v_1 is given by (8.22).

The condition (8.23), although complicated, is explicit once the resolvent kernel R_3 has been found. As in the case $\lambda = 0$, the condition $[u_1](\alpha) = 0$, when written in the form (8.23), can be viewed as an orthogonality condition on the input function H_1 and H_2.

REFERENCES

[1] J. A. COCHRAN, *Analysis of Linear Integral Equations*, McGraw-Hill, New York, 1972.

[2] W. D. COLLINS, *On some dual series equations and their applications to electrostatic problems for spheroidal caps*, Proc. Cambridge Philos. Soc., 57 (1961), pp. 367–384.

[3] ———, *Some scalar diffraction problems for spherical caps*, Arch. Rational Mech. Anal., 10 (1962), pp. 249–266.

[4] ———, *On some triple series equations and their applications*, Ibid., 11 (1962), pp. 122–137.

[5] H. L. JOHNSON, *An integral equation formulation of a mixed boundary value problem on a sphere*, this Journal, pp. 417–426.

[6] N. I. MUSKELISHVILI, *Singular Integral Equations*, P. Noordhoff, Groningen, the Netherlands, 1953.

[7] K. B. RANGER, *A note on some mixed boundary value problems for the heat equation*, SIAM J. Applied Math., 24 (1973), pp. 556–561.

[8] I. N. SNEDDON, *Mixed Boundary Value Problems in Potential Theory*, North-Holland, Amsterdam, 1966.

[9] A. ZYGMUND, *Trigonometrical Series*, Dover, New York, 1955.

SIAM J. MATH. ANAL.
Vol. 8, No. 1, February 1977

ON THE COMPONENTS OF EXTREMAL SOLUTIONS OF SECOND ORDER SYSTEMS*

SHAIR AHMAD† AND ALAN C. LAZER‡

Abstract. Sign properties of components of extremal solutions of the linear system

$$x''+A(t)x=0,$$

where $A(t)$ is a continuous $n\times n$ matrix, are investigated. For most of the considerations, $A(t)$ is assumed to be symmetric. The proofs are based on variational arguments.

1. Introduction. In this paper we study the second order linear system

$$y''+A(t)y=0, \tag{1}$$

where $A(t)=(a_{ij}(t))$ is a continuous $n\times n$ matrix function. In most of our considerations, $A(t)$ will also be assumed to be symmetric. We shall be mainly concerned with the sign properties of components of *extremal solutions* of (1). For pertinent bibliography, one might consult [1], [2] and [5]; particularly, Chapter VII of [5]. A solution $y(t)$ of (1) is an *extremal solution* if $y(t)\not\equiv 0$, $y(a)=y(b)=0$ for some a and b with $a<b$, and such that there exists no nontrivial solution $x(t)$ of (1) such that $x(a)=x(c)=0$ with $a<c<b$. To motivate one of the main results of this paper, consider the case where A is a *constant* symmetric matrix. Let $\lambda_1\geqq\lambda_2\geqq\cdots\geqq\lambda_n$ denote the eigenvalues of A. Every constant symmetric real matrix A is similar to a diagonal matrix via an orthogonal change of variables. Let T be an orthogonal matrix such that $T^{-1}AT=\operatorname{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$. Then the substitution $y(t)=Tx(t)$ reduces (1) to the uncoupled system

$$x_k''+\lambda_k x_k=0, \qquad k=1,2,\cdots,n.$$

If $x(a)=0$, then $x_k(t)=c_k\sin\sqrt{\lambda_k}(t-a)$ if $\lambda_k>0$, $x_k(t)=c_k(t-a)$ if $\lambda_k=0$, and $x_k(t)=c_k\sinh\sqrt{-\lambda_k}(t-a)$ if $\lambda_k<0$. Consequently if $\lambda_1>0$, then the smallest number $b>a$ such that there exists a nontrivial solution $y(t)$ of (1) with $y(a)=y(b)=0$ is $b=a+\pi/\sqrt{\lambda_1}$; and if v is a nonzero constant vector such that $Av=\lambda_1 v$, then $y(t)=\sin\sqrt{\lambda_1}(t-a)v$ is a nontrivial solution of the boundary value problem $y''+Ay=0$, $y(a)=y(b)=0$ such that *no component of $y(t)$ changes sign on (a,b)* (although some component may be identically zero).

This leads to the questions of when does the same type of behavior hold for nonconstant $A(t)$, and when can extremal solutions be characterized by the sign properties of their components?

We shall show that if $A(t)$ is symmetric and the off-diagonal elements are nonnegative, then the same type of behavior again holds. We also obtain more general results along this line. We shall also be concerned with the question of uniqueness of extremal solutions (up to multiplication by constants) for given a and b.

* Received by the editors May 12, 1975, and in revised form September 22, 1975.
† Department of Mathematics, Oklahoma State University, Stillwater, Oklahoma 74074.
‡ Department of Mathematics, University of Cincinnati, Cincinnati, Ohio 45221.

2. Lemmas and theorems. We first prove an elementary useful lemma.

LEMMA 1. *Let $A(t)=(a_{ij}(t))$ be an $n\times n$ continuous matrix on $[a,b]$ with $a_{ij}(t)\geqq 0$. Let $y=\operatorname{col}(y_1,\cdots,y_n)$ be a solution of* (1) *with $y(a)=y(b)=0$ and $y_i(t)\geqq 0$ for all $t\in((a,b)$ and all $i, i=1,2,\cdots,n$. If for some $k, k=1,\cdots,n$, either* (i) $y_k'(a)=0$, (ii) $y_k'(b)=0$ *or* (iii) $y_k(c)=0$ *for some $c, a<c<b$, then $y_k(t)\equiv 0$ on (a,b).*

Although an analytic proof of the above lemma is not difficult to see, it follows rather quickly from the following geometric observations, suggested to us by one of the referees of this paper. Since $y_k''(t)\leqq 0$ implies that the graph of y_k cannot rise above any tangent line, cases (i) and (ii) bound y_k above by the t axis, and hence $y_k\equiv 0$. The third case follows similarly upon observing that $y_k'(c)$ must be 0, as otherwise would imply negative values for $y_k'(t)$.

We recall that an $n\times n$ matrix $A=(a_{ij})$ is called irreducible if it is impossible to have $\{1,2,\cdots,n\}=I\cup J, I\cap J=\varnothing, I\neq\varnothing\neq J$, and $a_{ij}=0$ for all $i\in I, j\in J$.

THEOREM 1. *Let $A(t)$ satisfy the same conditions as in Lemma* 1, *and suppose that $A(t_0)$ is irreducible for some $t_0\in(a,b)$. If $y(t)$ is a nontrivial solution of* (1) *with $y(a)=y(b)=0$ and $y_i(t)\geqq 0$ on $(a,b), i=1,\cdots,n$, then for each $k, k=1,\cdots,n$, we have* (i) $y_k'(a)>0$, (ii) $y_k'(b)<0$ *and* (iii) $y_k(t)>0$ *on (a,b). Moreover, if w is any solution of* (1) *such that $w(a)=w(b)=0$, then $w(t)=\alpha y(t)$ for some constant α.*

Proof. Assume that for some $k, k=1,\cdots,n$, either (i), (ii) or (iii) fails. Then it follows from Lemma 1 that $y_k(t)\equiv 0$. Let $I=\{i\in\{1,2,\cdots,n\}|y_i(t)\equiv 0\}$, and let $J=\{1,2,\cdots,n\}-I$. By Lemma 1, for each $j\in J$ we have $y_j'(a)>0$, $y_j'(b)<0$, and $y_j(t)>0$ on (a,b). We note for each $i\in I, s\in(a,b)$, we have

$$0=y_i''(s)+\sum_{k=1}^{n}a_{ik}(s)y_k(s)=\sum_{k=1}^{n}a_{ik}(s)y_k(s)=\sum_{j\in J}a_{ij}(s)y_j(s).$$

However, in the last sum, since $y_j(s)>0$ and $a_{ij}(s)\geqq 0$, we must have $a_{ij}(s)=0$. Thus we have $a_{ij}(s)=0$ on (a,b) for $i\in I, j\in J$, contradicting the assumption that $A(t_0)$ is irreducible. This establishes the proof of the first part of the theorem.

To prove the second part, let $w(t)=\operatorname{col}(w_1,\cdots,w_n)$ be a solution of (1) such that $w(a)=w(b)=0$. We may assume, without loss of generality, that $w_{\bar{k}}(t_0)>0$ for some $\bar{k}, \bar{k}=1,\cdots,n$, and for some $t_0\in(a,b)$. Since for each k, $y_k'(a)>0$ and $y_k'(b)<0$, we can choose $\alpha>0$ such that $y_k'(a)-\alpha w_k'(a)>0$ and $y_k'(b)-\alpha w_k'(b)<0$, $k=1,\cdots,n$. Let $z_k(t)=y_k(t)-\alpha w_k(t)$. We note that $z_k(a)=z_k(b)=0$ by assumption. Since we have chosen α such that $z_k'(a)>0$ and $z_k'(b)<0$ for each k, there exists $\delta>0$ such that $z_k(t)>0$ on $(a,a+\delta), (b-\delta,b), k=1,\cdots,n$. Clearly, by choosing $\alpha>0$ sufficiently small (smaller if necessary) we have $z_k(t)=y_k(t)-\alpha w_k(t)>0$ on $[a+\delta,b-\delta], k=1,\cdots,n$. Let $A=\{\alpha>0|y_k(t)-\alpha w_k(t)>0$ on $(a,b), k=1,\cdots,n\}$. We have shown that $A\neq\varnothing$. A is bounded, since $\alpha\in A$ implies $\alpha<u_{\bar{k}}(t_0)/w_{\bar{k}}(t_0)$. Let $\alpha^*=$ l.u.b. A. We assume that w is independent of y. Let $z^*(t)=y(t)-\alpha^*w(t)$. It follows from the definition of α^* that $z_k^*(t)\geqq 0$ on $(a,b), k=1,\cdots,n$. Thus by the first part of the theorem, $z^*(t)\not\equiv 0$ implies $z_k^{*\prime}(a)>0$, $z_k^{*\prime}(b)<0$, and $z_k^*(t)>0$ on $(a,b), k=1,\cdots,n$. By the argument just given (y replaced by z^*), for $\beta>0$ sufficiently small, $z_k^*(t)-\beta w_k(t)>0$ on $(a,b), k=1,\cdots,n$. However, $z^*(t)-\beta w(t)=y(t)-(\alpha^*+\beta)w(t)$, contradicting the definition of α^*. Therefore, some component of z^* vanishes in (a,b) and hence $z^*(t)\equiv 0$ or $w=(\alpha^*)^{-1}y$. This completes the proof.

We recall that two distinct points a and b are said to be *conjugates relative* to (1) if there exists a nontrivial solution $y(t)$ of (1) such that $y(a)=y(b)=0$. Equation (1) is said to be *disconjugate* on an interval I if no nontrivial solution of (1) has more than one zero in I. It is well-known that if $A(t)$ in (1) is symmetric, then conjugate points are isolated. Let $\mathscr{A}[a, b]$ denote the set of absolutely continuous R^n-valued functions $h(t)$ on $[a, b]$ such that $|h'| \in L^2[a, b]$ and $h(a)=h(b)=0$. Let $J[h]$ define the functional

$$J[h]=\int_a^b (\langle h', h'\rangle - \langle A(t)h, h\rangle)\, dt$$

over the set $\mathscr{A}[a, b]$ of admissible functions. It follows (see [2, p. 123]) that if $[a, b]$ contains no point conjugate to a, then for all $h \in \mathscr{A}[a, b]$, $h(t) \not\equiv 0$, $J[h]>0$ (assuming that $A(t)$ is symmetric). Our next lemma is a slight modification of this result.

LEMMA 2. *If $A(t)$ in (1) is symmetric, then $J[h] \geqq 0$ for all $h \in \mathscr{A}[a, b]$ if and only if $[a, b]$ contains no point conjugate to a in its interior.*

Proof. For a proof of the sufficiency of this lemma, one is referred to [5, prob. 4.2, p. 332]. To see the necessity, we recall (see, e.g., [3]) that if $J[h] \geqq 0$ for all $h \in \mathscr{A}[a, b]$, then for any $z \in \mathscr{A}[a, b]$, $J[z]=0$ implies that z is a solution of (1). Thus if c is a point in the interior of $[a, b]$ which is conjugate to a, then there is a nontrivial solution $u(t)$, $u(t) \not\equiv 0$, of (1) with $u(a)=u(c)=0$. Let $z(t)$ be the function defined by $z(t)=u(t)$ if $a \leqq t \leqq c$ and $z(t)=0$ if $c \leqq t \leqq b$. Then $z \in \mathscr{A}[a, b]$ and $J[z]=0$. Hence $z(t)$ is a nontrivial solution of (1) defined on $[a, b]$, which is impossible since $z(t) \equiv 0$ on $[c, b]$.

THEOREM 2. *Let $A(t)=(a_{ik}(t))$ in (1) be an $n \times n$ symmetric matrix such that $a_{ik}(t) \geqq 0$ whenever $i \neq k$. If $b>a$ is the first conjugate point to a (relative to (1)), then there exists a nontrivial solution $u(t)=\operatorname{col}(u_1, \cdots, u_n)$ of (1) with $u(a)=u(b)=0$, and $u_k(t) \geqq 0$ on $[a, b]$, $k=1, \cdots, n$.*

Remark. One of the referees of this paper has pointed out to us that the statement of Theorem 2 would not be any weaker if one assumed $a_{ii}(t) \geqq 0$, $i=1, \cdots, n$. For by making a substitution $y(t)=e^{kt}u(t)$ and a classical change of variable from t to s, it follows that (1) is equivalent to an equation of the form

$$\frac{d^2u}{ds^2}+[k^2I+A(t(s))]\, e^{4kt(s)}u(s)=0.$$

Proof of Theorem 2. By definition of conjugate point, there exists a nontrivial solution $y(t)=\operatorname{col}(y_1, \cdots, y_n)$ of (1) with $y(a)=y(b)=0$. From $y''+A(t)y=0$ we obtain

$$-\langle y, y''\rangle - \langle y, Ay\rangle = 0,$$

and integrating the latter from a to b we have

$$\begin{aligned} 0 &= \langle y, y'\rangle\Big|_a^b + \int_a^b (\langle y', y'\rangle - \langle y, Ay\rangle)\, dt \\ &= \int_a^b (\langle y', y'\rangle - \langle y, Ay\rangle)\, dt = J[y], \end{aligned} \tag{2}$$

since $y(a)=y(b)=0$ by hypothesis. Let $u(t)=\operatorname{col}(u_1,\cdots,u_n)$, where $u_k(t)=|y_k(t)|$, $k=1,2,\cdots,n$. Clearly, $u(t)\not\equiv 0$ and $u(t)\in\mathscr{A}[a,b]$. We have

$$
\begin{aligned}
J[u]&=\int_a^b (\langle u',u'\rangle-\langle u,Au\rangle)\,dt\\
&=\int_a^b \Big(\langle u',u'\rangle-\sum_{i=1}^n\sum_{k=1}^n a_{ik}(t)u_i(t)u_k(t)\Big)\,dt \qquad (3)\\
&=\int_a^b \Big(\langle u',u'\rangle-\sum_{i=k} a_{ii}u_i^2-\sum\sum_{i\neq k} a_{ik}u_iu_k\Big)\,dt.
\end{aligned}
$$

Now $\langle u',u'\rangle=\langle y',y'\rangle$, and

$$\sum_{i=1}^n a_{ii}u_i^2=\sum_{i=1}^n a_{ii}|y_i|^2=\sum_{i=1}^n a_{ii}y_i^2.$$

Furthermore, since $a_{ik}\geqq 0$ for $i\neq k$, we have

$$
\begin{aligned}
\sum\sum_{i\neq k} a_{ik}u_iu_k&=\sum\sum_{i\neq k} a_{ik}|y_i|\,|y_k|\\
&\geqq\sum\sum_{i\neq k} a_{ik}y_iy_k.
\end{aligned}
$$

Hence from (3) and (2) we have

$$
\begin{aligned}
J[u]&\leqq\int_a^b \Big(\langle y',y'\rangle-\sum_{i=1}^n a_{ii}y_i^2-\sum\sum_{i\neq k} a_{ik}y_iy_k\Big)\,dt\\
&=\int_a^b (\langle y',y'\rangle-\langle y,Ay\rangle)\,dt\\
&=J[y]=0.
\end{aligned}
$$

But since b is the first conjugate point to a, by Lemma 2, we also have $J[u]\geqq 0$; thus $J[u]=0$. Hence by a standard result in the calculus of variations (see, e.g., [3]), $u\in C^2[a,b]$ with $u''+A(t)u=0$ and $u(a)=u(b)=0$, because u affords a minimum to J for the class $\mathscr{A}[a,b]$ of admissible functions. This completes the proof.

Remark 1. It follows that the components of the solution $u(t)$ of Theorem 2 cannot have simple zeros on (a,b).

Making use of a certain transformation, we now obtain a more general theorem.

THEOREM 2′. *Assume that $A(t)=(a_{ij}(t))$ in* (1) *is symmetric, and let b be the first conjugate point to a. Suppose that $\{1,2,\cdots,n\}=P\cup N$ with $P\cap N=\varnothing$ and*

(i) $a_{ij}(t)\geqq 0$ *on* $[a,b]$ *if* $i\neq j$ *and either* $i,j\in P$ *or* $i,j\in N$,

(ii) $a_{ij}(t)\leqq 0$ *on* $[a,b]$ *if* $i\neq j$ *and either* $i\in P$, $j\in N$ *or* $i\in N$, $j\in P$.

Then there exists a nontrivial solution $y(t)=\operatorname{col}(y_1,\cdots,y_n)$ of (1) *such that $y(a)=y(b)=0$, $y_k(t)\geqq 0$ on $[a,b]$ for $k\in P$, and $y_k(t)\leqq 0$ on $[a,b]$ for $k\in N$.*

Proof. Define T to be the matrix $T = \operatorname{diag}(t_1, t_2, \cdots, t_n)$, where $t_i = 1$ if $i \in P$ and $t_i = -1$ if $i \in N$. Obviously, $T^{-1} = T$ and T is symmetric. By setting $y = Tx$, (1) can be transformed into the system

$$x'' + B(t)x = 0, \tag{1'}$$

where $B(t) = TA(t)T$. It is easy to verify that $B(t)$ satisfies the hypothesis of Theorem 2, i.e., it is symmetric and its elements b_{ij} are nonnegative on $[a, b]$ for $i \neq j$. Hence, by Theorem 2, there exists a nontrivial solution $x(t) = \operatorname{col}(x_1, \cdots, x_n)$ of (1') satisfying $x(a) = x(b) = 0$, and $x_k(t) \geqq 0$ on $[a, b]$. But then $y = Tx$ is a solution of (1) satisfying the conclusion of Theorem 2'.

Remark 2. We note that Theorem 2 follows from Theorem 2' when $N = \varnothing$.

Remark 3. It can be verified that Theorem 1, as well as our next two theorems, have similar generalizations.

Example. If $n = 3$ and $N \neq \varnothing \neq P$, then one of the sets N and P contains one element and the other two elements. Suppose that $P = \{1\}$ and $N = \{2, 3\}$. This means that $a_{12}(t) = a_{21}(t) \leqq 0$, $a_{13}(t) = a_{31}(t) \leqq 0$ and $a_{23}(t) = a_{32}(t) \geqq 0$ for $t \in [a, b]$. In this case, $y(t) = \operatorname{col}(y_1, y_2, y_3)$ with $y_1(t) \geqq 0$, $y_2(t) \leqq 0$, and $y_3(t) \leqq 0$ on $[a, b]$.

THEOREM 3. *Assume that $A(t) = (a_{ij}(t))$ in (1) is symmetric, and positive definite in (a, b) except at isolated points. If $a_{ij}(t) > 0$ on (a, b), $i, j = 1, \cdots, n$, and if there exists a nontrivial solution $v(t) = \operatorname{col}(v_1, \cdots, v_n)$ of (1) with $v(a) = v(b) = 0$ and $v_k(t) \geqq 0$, $k = 1, \cdots, n$, then b is the first conjugate point to a relative to (1).*

Proof. Assume that b is not conjugate to a. Let $s^* = \text{g.l.b.}\{s \mid$ there exist points conjugate to a on $[a, b]$ relative to $y'' + sA(t)y = 0\}$. Clearly, if $s < 0$, then

$$\int_a^b (\langle y', y' \rangle - s\langle Ay, y \rangle)\, dt > 0$$

for $y \in \mathcal{A}[a, b]$, $y \not\equiv 0$. Therefore, s^* exists and $s^* \geqq 0$. Since b is not the first conjugate point of a, by Lemma 2 there exists $\bar{y} \in \mathcal{A}[a, b]$ such that

$$J[\bar{y}] = \int_a^b (\langle \bar{y}', \bar{y}' \rangle - \langle A\bar{y}, \bar{y} \rangle)\, dt < 0.$$

Obviously, if $s < 1$ and s is sufficiently close to 1, then

$$\int_a^b (\langle \bar{y}', \bar{y}' \rangle - s\langle A\bar{y}, \bar{y} \rangle)\, dt < 0.$$

Therefore, by Lemma 2, s belongs to the above set. This shows that $s^* < 1$.

Consider the differential equation

$$y'' + s^* A(t)y = 0. \tag{4}$$

It follows that $[a, b)$ contains no points conjugate to a relative to (4). For otherwise, by Lemma 2, there would exist $\bar{z} \in \mathcal{A}[a, b]$ satisfying the inequality

$$\int_a^b (\langle \bar{z}', \bar{z}' \rangle - s^* \langle A\bar{z}, \bar{z} \rangle)\, dt < 0.$$

But then one could choose a number s sufficiently close to s^*, $s<s^*$, satisfying

$$\int_a^b (\langle \bar{z}', \bar{z}' \rangle - s\langle A\bar{z}, \bar{z} \rangle)\, dt < 0.$$

This would imply, by Lemma 2, that there exists a point conjugate to a in (a, b) relative to $y''+sAy=0$, contradicting the definition of s^*.

Next, we wish to show that b is conjugate to a relative to (4), and hence the first point conjugate to a relative to (4). Let $Y(t, s)$ denote the $n \times n$ matrix satisfying

$$Y''+sA(t)Y=0,$$

and initial conditions $Y(a, s)=0$, $Y'(a, s)=I$, where I is the identity matrix. We note that the first conjugate point of a relative to

$$y''+sA(t)y=0$$

is the first zero of $\det Y(t, s)$ on (a, ∞). By continuity with respect to parameters, $Y(t, s)$ is continuous in both variables. Now, let M be a number such that $M>0$ and $A(t)<M \cdot I$ on $[a, b]$. Choose a positive number σ such that $\sigma<b-a$ and $\sigma<\pi/\sqrt{M(s^*+1)}$. The first point conjugate to a relative to $y''+(s^*+1)MIy=0$ is $a+(\pi/\sqrt{M(s^*+1)})$. If $s^* \leqq s \leqq s^*+1$, then $sA(t) \leqq (s^*+1)MI$; and hence by the Sturmian comparison theorem for systems (see [4]), the first point conjugate to a relative to $y''+sAy=0$ is greater than the first point conjugate to a relative to $y''+(s^*+1)MIy=0$. Since $a+(\pi/\sqrt{M(s^*+1)})>a+\sigma$, it follows that for $s^* \leqq s \leqq s^*+1$, $y''+sAy=0$ is disconjugate on $[a, a+\sigma]$. If we assume that b is not conjugate to a relative to (4), then for all t in $[a+\sigma, b]$, $\det Y(t, s^*) \neq 0$. By continuity, there exists δ, $0<\delta<1$, such that $\det Y(t, s) \neq 0$ for $t \in [a+\sigma, b]$ and $s^* \leqq s \leqq s^*+\delta$. Therefore, $\det Y(t, s) \neq 0$ for $t \in [a, b]$ and $s^* \leqq s \leqq s^*+\delta$, since we have shown that for such s, $y''+sAy=0$ is disconjugate on $[a, a+\delta]$. This shows that for $s^* \leqq s \leqq s^*+\delta$, $y''+sAy=0$ is disconjugate on $[a, b]$, which is a contradiction to the definition of s^*. Thus we have shown that b is the first conjugate point of a relative to (4). Thus, by Theorem 2, there exists a nontrivial solution $u(t)=\operatorname{col}(u_1, \cdots, u_n)$ of (4) such that $u(a)=u(b)=0$, and $u_k(t) \geq 0$, $k=1, \cdots, n$. We also have, by hypothesis,

$$v''+A(t)v=0$$

with $v(a)=v(b)=0$, and $v_k(t) \geqq 0$, $k=1, \cdots, n$; $v(t) \not\equiv 0$. Noting that $A(t)$ is symmetric, we have

$$\begin{aligned}\langle u, v'' \rangle - \langle v, u'' \rangle &= s^*\langle v, Au \rangle - \langle u, Av \rangle \\ &= (s^*-1)\langle u, Av \rangle.\end{aligned}$$

Therefore

$$(\langle u, v' \rangle - \langle v, u' \rangle)' = (s^*-1)\langle u, A(t)v \rangle.$$

Integrating both sides of the above equation from a to b, and noting that u and v both vanish at a and b, one obtains

$$\int_a^b \langle u, A(t)v \rangle\, dt = 0,$$

which is a contradiction, since u and v are nontrivial with nonnegative components and $A(t)$ has positive entries. This completes the proof.

THEOREM 4. *If* (1) *is disconjugate on* (a, ∞), $a_{ij}(t) \geqq 0$, $A(t)$ *is symmetric and* $A(t_0)$ *is irreducible for some* $t_0 > a$, *then there is a solution* $y(t) = \operatorname{col}(y_1, \cdots, y_n)$ *of* (1) *with* $y(a) = 0$ *and* $y_k(t) > 0$ *on* (a, ∞), $k = 1, \cdots, n$.

Proof. For each integer $m > 0$, consider the equations

$$y'' + (A(t) + (1/m^2)I)y = 0 \tag{5}$$

and

$$z'' + ((1/m^2)I)z = 0. \tag{6}$$

Let $z(t)$ be the solution of (6) defined as $z(t) = \sin(1/m)(t-a) \cdot c$, where c is a constant vector in R^n. Since $z(a) = z(a + m\pi) = 0$, z must have a first conjugate point in $(a, a + m\pi]$. From (6) we have

$$\int_a^{a+m\pi} \left(\langle z', z' \rangle - \frac{1}{m^2} \langle z, z \rangle \right) dt = 0.$$

Therefore,

$$\int_a^{a+m\pi} \left(\langle z', z' \rangle - \left\langle \left(A(t) + \frac{1}{m^2} I \right) z, z \right\rangle \right) dt \leqq 0. \tag{7}$$

Now, it is well known that (5) is disconjugate on an interval $[c, d]$, if and only if

$$J[u] = \int_c^d \left(\langle u', u' \rangle - \left\langle \left(A(t) + \frac{1}{m^2} I \right) u, u \right\rangle \right) dt > 0$$

for all $u \in \mathcal{A}[c, d]$, $u \not\equiv 0$. Thus (7) implies that the first conjugate point to a relative to (5) must lie in the interval $(a, a + m\pi)$. Let $b(m)$ denote the first conjugate point to a relative to (5); $b(m)$ exists by the Sturmian comparison theorem. By Theorem 2, there exists a solution $y_m(t) = \operatorname{col}(y_{m1}, \cdots, y_{mn})$ of (5) with $y_m(a) = y_m(b) = 0$, $y_{mk}(t) \geqq 0$, $k = 1, \cdots, n$, and $y_m(t) \not\equiv 0$. Using the Sturmian comparison theorem, it follows, as in the proof of Theorem 3, that $m_1 < m_2$ implies $b(m_1) < b(m_2)$ and $b(m) \to \infty$ as $m \to \infty$, since

$$y'' + A(t) = 0 \tag{1}$$

is disconjugate on (a, ∞) ($b(m)$ denotes the first conjugate point of a relative to (5)). Without loss of generality, we can assume that $\|y_m'(a)\| = 1$, and hence can further assume that $y_m'(a) \to c$, where c is a constant vector with $\|c\| = 1$. Let $u(t)$ be a solution of (1) with $u(a) = 0$ and $u'(a) = c$. Then $y_m(t) \to u(t)$ on (a, ∞). Now, $y_{mk}(t) \geqq 0$ on $(a, b(m))$ and $b(m) \to \infty$ as $m \to \infty$ implies that $u_k(t) \geqq 0$ on (a, ∞), $k = 1, \cdots, n$. By Lemma 1, if $u_p(s) = 0$ for some $s > a$ and some p, $p = 1, \cdots, n$, then $u_p(t) \equiv 0$ for all $t > a$. Thus the proof of the theorem follows from an argument similar to that given to the proof of Theorem 1.

REFERENCES

[1] W. A. COPPEL, *Disconjugacy*, Lecture Notes in Mathematics, vol. 220, Springer-Verlag, Berlin, 1971.

[2] I. M. GELFAND AND S. V. FOMIN, *Calculus of Variations*, Prentice-Hall, Englewood Cliffs, N.J., 1963.

[3] M. R. HESTENES, *Calculus of Variations and Optimal Control Theory*, John Wiley, New York, 1966.

[4] M. MORSE, *A generalization of the Sturm separation and comparison theorems in n-space*, Math. Ann., 103 (1930), pp. 52–69.

[5] W. T. REID, *Ordinary Differential Equations*, John Wiley, New York, 1971.

SIAM J. MATH. ANAL.
Vol. 8, No. 1, February 1977

AXIAL RADIATION RECEPTION. PART II*

LIM CHEE-SENG†

Abstract. This paper again concerns the axial reception of radiation from a suddenly triggered, constantly maintained point source within an n-dimensional, nondispersive, axially symmetric medium. Part I (Chee-Seng (1973)) deals solely with the situation where n is odd and $\geqq 3$. The present part II completes the picture by focusing on the case for even n ($\geqq 2$), and extends a hyperbolicity-cum-ellipticity to include nonstrictness. In most aspects, the odd and even n-problems are, expectantly, quite distinct regarding their respective analyses, results and subsequent interpretations, as with multi-dimensional Cauchy problems in general. Additionally, nonstrictness leads to some rather surprising developments. It first crops up in both applications to magnetogasdynamic flow aligned receptions along, as well as crossing, the magnetic field. The mainstream is largely responsible.

1. Introduction. Suppose the position vector $\mathbf{x} = (x_1, x_2, \cdots, x_n) \in R_n$, the unbounded n-dimensional Cartesian space. The preceding paper (part I, Chee-Seng (1973)) deals with the following unsteady radiation problem:

$$Q\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x_1}, \nabla_1^2\right)\phi = P\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x_1}, \nabla_1^2\right)H(t)\delta(\mathbf{x}), \tag{1.1}$$

such that

$$\phi = 0 \text{ during (time) } t < 0, \tag{1.2}$$

solved solely in the case of *odd* n ($\geqq 3$) and for *axial reception* along the (x_1-) axis of symmetry through the radiating point source $\delta(\mathbf{x})H(t)$ which is abruptly activated at instant $t = 0$, $H(t)$ being the Heaviside unit function, while $\delta(\mathbf{x})$ denotes the Dirac delta function in R_n. Also, $\nabla_1^2 \equiv \partial^2/\partial x_2^2 + \cdots + \partial^2/\partial x_n^2$, the $(n-1)$-dimensional Laplacian. The (polynomial) Q-operator is of order m, an even integer, whilst the P-operator is of (integral) order $m - l$: $1 \leqq l \leqq m$. Both operators correspond to a real, homogeneous, nondispersive, transmitting medium. The highest order attained by the ∇_1^2 term in Q need not be $\frac{1}{2}m$ but, generally, $\leqq \frac{1}{2}m$; similarly for P. So there exist even integers p and q, called *Laplacian indices* of the respective P- and Q-operators, satisfying $l \leqq p \leqq m$ and $0 \leqq q < m$, whereby

$$\underset{(\text{in } \nabla_1^2)}{\deg \text{ poly } P}\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x_1}, \nabla_1^2\right) = \tfrac{1}{2}(m - p), \tag{1.3}$$

meaning that the degree of the polynomial P, when referred to ∇_1^2, is $\frac{1}{2}(m - p)$,

$$\underset{(\text{in } \nabla_1^2)}{\deg \text{ poly } Q}\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x_1}, \nabla_1^2\right) = \tfrac{1}{2}(m - q). \tag{1.4}$$

* Received by the editors November 19, 1974, and in revised form July 7, 1975.

† Institute of Geophysics and Planetary Physics, University of California at Los Angeles, Los Angeles, California 90024. On leave from Department of Mathematics, University of Malaya, Kuala Lumpur, Malaysia. This work was supported by the Institute of Geophysics and Planetary Physics, University of California at Los Angeles, and appears as Publication No. 1546.

Generally, then,

$$(1.5)\qquad P\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x_1}, \nabla_1^2\right) \equiv \sum_{\mu=0}^{(1/2)(m-p)} \nabla_1^{2\mu} \sum_{\nu=0}^{m-l-2\mu} A_\mu^\nu \left(\frac{\partial}{\partial x_1}\right)^\nu \left(\frac{\partial}{\partial t}\right)^{m-l-2\mu-\nu},$$

$$(1.6)\qquad Q\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x_1}, \nabla_1^2\right) \equiv \sum_{\mu=0}^{(1/2)(m-q)} \nabla_1^{2\mu} \sum_{\nu=0}^{m-2\mu} B_\mu^\nu \left(\frac{\partial}{\partial x_1}\right)^\nu \left(\frac{\partial}{\partial t}\right)^{m-2\mu-\nu},$$

where A_μ^ν and B_μ^ν are real constants, the $A_{(1/2)(m-p)}^\nu$'s are not all zeros for $\nu = 0, 1, \cdots, p - l$, the $B_{(1/2)(m-q)}^\nu$'s are not all zeros for $\nu = 0, 1, \cdots, q$. Observe a homogeneity consistent with no-dispersion:

$$(1.7)\qquad P(\beta\lambda, \beta\xi, \beta^2\chi^2) = \beta^{m-l}P(\lambda, \xi, \chi^2), \qquad Q(\beta\lambda, \beta\xi, \beta^2\chi^2) = \beta^m Q(\lambda, \xi, \chi^2)$$

for all parameters β, λ, ξ, χ.

The solution to (1.1) and (1.2) may be formally expressed, following Lighthill (1960, Appendix B), as a Fourier integral, viz., (part I)

$$(1.8)\qquad \phi = \frac{i}{(2\pi)^{n+1}} \int_{R_n} \exp(i\boldsymbol{\alpha}\cdot\mathbf{x})\, d\boldsymbol{\alpha} \int_{-\infty+i\varepsilon}^{\infty+i\varepsilon} \frac{P(-i\omega, i\alpha_1, \alpha_1^2 - \boldsymbol{\alpha}^2)}{Q(-i\omega, i\alpha_1, \alpha_1^2 - \boldsymbol{\alpha}^2)} \frac{e^{-i\omega t}}{\omega}\, d\omega,$$

valid for both odd and even n, with $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ ranging over R_n, $d\boldsymbol{\alpha} = d\alpha_1\, d\alpha_2 \cdots d\alpha_n$, $\boldsymbol{\alpha}\cdot\mathbf{x} = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$ (a scalar product). The ω-integral path $(-\infty + i\varepsilon, \infty + i\varepsilon)$ must be compatible with the (partial) zero condition (1.2).

The *purpose of the present paper* is to tackle axial reception in the complementary case where n is *even* and $\geqq 2$, the basic assumption throughout. (Note that when $n = 2, \nabla_1^2 \equiv \partial^2/\partial x_2^2$). For odd and even n, respective solutions to the classical Cauchy problem embodying any of the standard time-dependent equations are known to differ substantially in their derivations, forms and interpretations (see, e.g., Courant and Hilbert (1962)). This is also true with regard to axial reception of the Cauchy-type problem governed by (1.1) and (1.2). So the present case study, essential for completeness, is also vital from a comparison standpoint, besides its various applicabilities.

2. Hyperbolicity cum ellipticity. From now on, we suppose that the reception is axially aligned and avoids the source point, i.e., $\mathbf{x} = (x_1, 0, 0, \cdots, 0)$ with $x_1 \neq 0$, so that $\alpha_1 = x_1^{-1}\boldsymbol{\alpha}\cdot\mathbf{x}$. Then it can be shown, by a law of the spherical mean (John (1955)), that

$$\int_{R_n} f(\boldsymbol{\alpha}\cdot\mathbf{x}) g(\alpha_1, \boldsymbol{\alpha}^2)\, d\boldsymbol{\alpha}$$

$$= \tfrac{1}{2}\Omega_{n-1} \int_{-1}^{1} (1 - \xi^2)^{(1/2)(n-3)}\, d\xi \int_{-\infty}^{\infty} (\operatorname{sgn}\alpha) f(\alpha\xi x_1) g(\alpha\xi, \alpha^2)\alpha^{n-1}\, d\alpha,$$

where $\Omega_n = 2\pi^{(1/2)n}/\Gamma(\frac{1}{2}n)$, and $\operatorname{sgn}\alpha = 1$ $(\alpha > 0)$, -1 $(\alpha < 0)$. For odd n (part I), the factor $\operatorname{sgn}\alpha$ must be replaced by 1. This difference, though apparently minor to begin with, is to lead to a substantial deviation of the present analysis. The

representation (1.8) is now expressible in the reduced form:

$$\phi = \frac{i\Omega_{n-1}}{2(2\pi)^{n+1}} \left(\int_{-1}^{0} + \int_{0}^{1} \right) (1 - \xi^2)^{(1/2)(n-3)} \, d\xi \int_{-\infty}^{\infty} (\operatorname{sgn} \alpha) \exp(i\alpha\xi x_1) \alpha^{n-1} \, d\alpha \cdot \int_{-\infty + i\varepsilon}^{\infty + i\varepsilon} \frac{P(-i\omega, i\alpha\xi, \alpha^2\xi^2 - \alpha^2)}{Q(-i\omega, i\alpha\xi, \alpha^2\xi^2 - \alpha^2)} \frac{e^{-i\omega t}}{\omega} \, d\omega. \tag{2.1}$$

Unless otherwise specified, it is normally understood that $|x_1| \neq \infty$.

For a compatible, stable configuration, let us postulate, following Chee-Seng (op. cit.), that, under axial alignment, the Q-operator is *hyperbolic*, but evolves ultimately into a time-independent form which is *elliptic* (cf. Gårding (1951), Courant and Hilbert (1962), Bers et al. (1964)), viz. for any real $\xi \in [-1, 1]$:

$$\text{deg poly } Q(-\lambda, \xi, 1 - \xi^2) = m, \quad \text{or equivalently} \quad Q(1, 0, 0) \neq 0; \tag{2.2}$$
(in λ)

furthermore all (subsequent) m λ-roots, $\lambda \cdot = \lambda_j(\xi)$ $(j = 1, \cdots, m)$ to the characteristic relation

$$Q(-\lambda, \xi, 1 - \xi^2) = 0 \tag{2.3}$$

are real and, moreover, if $\xi \neq 0$, are nonvanishing as well as distinct. Hence, the *characteristic polynomial* is factorizable on $-1 \leqq \xi \leqq 1$:

$$Q(-\lambda, \xi, 1 - \xi^2) \equiv Q(1, 0, 0) \prod_{j=1}^{m} \lambda - \lambda_j(\xi), \tag{2.4}$$

where $\lambda_j(\xi) \neq \lambda_k(\xi)$ $(j \neq k)$ whenever $0 < |\xi| \leqq 1$, throughout which

$$Q(0, \xi, 1 - \xi^2) \equiv Q(1, 0, 0) \prod_{j=1}^{m} \lambda_j(\xi) \neq 0. \tag{2.5}$$

Both the hyperbolicity and ellipticity are *strict* with regard to $-1 \leqq \xi < 0$, $0 < \xi \leqq 1$. Under these circumstances, Chee-Seng (op. cit.) observed that, with reference to (2.1), the ω-integrand factor accompanying $e^{-i\omega t}$ is a meromorphic function which satisfies a uniform convergence requirement of Jordan's lemma in the entire complex ω-plane, and possesses $m + 1$ real, simple poles at $\omega = 0$, $\alpha\lambda_j(\xi)$ $(j = 1, \cdots, m)$. The condition (1.2) is then fulfilled if one chooses ε to be real and positive. Residue theory then yields

$$\int_{-\infty + i\varepsilon}^{\infty + i\varepsilon} (\cdot) \, d\omega = -2\pi i H(t) \left[\operatorname*{residue}_{\omega = 0} (\cdot) + \sum_{j=1}^{m} \operatorname*{residue}_{\omega = \alpha\lambda_j(\xi)} (\cdot) \right] \tag{2.6}$$

for the innermost ω-integral in (2.1). Hereafter, we assume that $t > 0$, so that the Heaviside factor $H(t) \equiv 1$. We now depart from the arguments of the preceding paper. Consider

$$\int_{-\infty}^{\infty} (\cdot) \, d\omega \tag{2.7}$$

involving the same ω-integrand, but evaluated along the pole-studded Re ω-axis. As these poles are all simple, (2.7) is meaningful in the sense of a Cauchy principal value, by infinitesimally indenting the real path about every pole. This indented path may now be closed by an infinite, clockwise-directed semicircle, necessarily described into the lower half-plane $\operatorname{Im}\omega < 0$, by virtue of Jordan's lemma. Contour integration then produces a result for (2.7) equaling half the value on the right side of (2.6) (with $t > 0$). Thus by a substitution of the integration variable for (2.7) and an appeal to the homogeneity rules of (1.7), we obtain

$$(2.8)\qquad \int_{-\infty+i\varepsilon}^{\infty+i\varepsilon}(\cdot)\,d\omega = 2(i\alpha)^{-l}(\operatorname{sgn}\alpha)\int_{-\infty}^{\infty}\frac{P(-\lambda,\xi,1-\xi^2)}{\lambda Q(-\lambda,\xi,1-\xi^2)}\exp(-i\alpha\lambda t)\,d\lambda.$$

Whereupon, (2.1) can be manipulated into the following:

$$(2.9)\qquad \phi = \phi_1(x_1,t) = \frac{\Omega_{n-1}}{(2\pi)^{n-1}}\left(\frac{\partial}{\partial x_1}\right)^{n-l-1}\frac{K(X)}{x_1} \quad \text{if } n \geqq l+1,$$

$$(2.10)\qquad \left(\frac{\partial}{\partial x_1}\right)^{l+1-n}\phi = \phi_2(x_1,t) = \frac{\Omega_{n-1}}{(2\pi)^{n-1}}\frac{K(X)}{x_1} \quad \text{if } n \leqq l+1,$$

in terms of a K-element, dependent on just $X = x_1/t$ and given by

$$K(X) = \frac{(-1)^{(1/2)n-1}}{(2\pi)^2}\left(\int_{-1}^{0}+\int_{0}^{1}\right)(1-\xi^2)^{(1/2)(n-3)}\xi^{l+1-n}\,d\xi$$
$$\cdot\int_{-\infty}^{\infty}\exp(i\alpha\xi)\,d\alpha\int_{-\infty}^{\infty}\frac{P(-\lambda X,\xi,1-\xi^2)}{\lambda Q(-\lambda X,\xi,1-\xi^2)}\exp(-i\alpha\lambda)\,d\lambda.$$

But, by Fourier transformation and inversion:

$$2\pi f(\xi) = \int_{-\infty}^{\infty}\exp(i\alpha\xi)\,d\alpha\int_{-\infty}^{\infty}f(\lambda)\exp(-i\alpha\lambda)\,d\lambda.$$

Whence,

$$(2.11)\quad K(X) = \frac{(-1)^{(1/2)n-1}}{2\pi}\left(\int_{-1}^{0}+\int_{0}^{1}\right)\frac{(1-\xi^2)^{(1/2)(n-3)}P(-\xi X,\xi,1-\xi^2)}{\xi^{n-l}Q(-\xi X,\xi,1-\xi^2)}\,d\xi.$$

The corresponding expression for odd n (part I) differs drastically in its form, the subsequent approach accorded to it, and the eventual solution it leads to.

The case of nonstrict hyperbolicity and ellipticity is a highly complicated one and deferred to §§ 7 and 8.

Some of the concepts we use, in dealing with the n-dimensional system, are, roughly speaking, adapted and modified from those of Weitzner (1961), Burridge (1967) and Payton (1971) for two-dimensional anisotropic waves emitted by the instantaneous point impulse $\delta(t)\delta(\mathbf{x})$, the corresponding solution being the Green's function. This is, within the present context, the quantity $\partial\phi/\partial t$. Should the process ultimately attain a steady state, the Green's function vanishes identically throughout this steady state, thereby becoming trivial. In fact, as far as the Green's

function is concerned, the notion of a steady state seems entirely detached. No elliptic-type condition is necessary. In contrast, owing to an indefinitely radiating source, our problem requires an ellipticity for the establishment of $K(X)$, and to ensure a bounded transition into a valid steady state. The governing differential equations considered in the three papers cited are strictly hyperbolic.

3. The *K*-element. To tackle the representation (2.11) for $K(X)$, we first apply the transformation

$$\zeta = \xi^{-1}(1 - \xi^2)^{1/2} \quad (\operatorname{Im} \xi = 0). \tag{3.1}$$

This maps the two real, right-directed ξ-paths $(-1, 0)$ and $(0, 1)$ onto two real, left-directed, semi-infinite ζ-paths, viz. $(0, -\infty)$ and $(\infty, 0)$, respectively. Furthermore, via (1.7):

$$P(-\xi X, \xi, 1 - \xi^2) \equiv \xi^{m-1} P(-X, 1, \zeta^2), \tag{3.2}$$

$$Q(-\xi X, \xi, 1 - \xi^2) \equiv \xi^m Q(-X, 1, \zeta^2). \tag{3.3}$$

Consequently, (2.11) converts to

$$K(X) = \frac{(-1)^{(1/2)n-1}}{2\pi} \int_{-\infty}^{\infty} \zeta^{n-2} \frac{P(-X, 1, \zeta^2)}{Q(-X, 1, \zeta^2)} d\zeta, \tag{3.4}$$

which can be determined by means of contour integration. To accomplish this, however, a preliminary study of the integrand is vital. In this section, we concentrate essentially on the case $|X| > 0$ (i.e., $t < \infty$).

Now, by (1.6), the polynomial factor (involved in (3.4))

$$Q(-X, 1, \zeta^2)$$

$$\equiv \zeta^{m-q} \sum_{\nu=0}^{q} B^{\nu}_{(1/2)(m-q)}(-X)^{q-\nu} + \zeta^{m-q-2} \sum_{\nu=0}^{q+2} B^{\nu}_{(1/2)(m-q)-1}(-X)^{q+2-\nu} \tag{3.5}$$

$$+ \cdots + \zeta^2 \sum_{\nu=0}^{m-2} B^{\nu}_1(-X)^{m-2-\nu} + \sum_{\nu=0}^{m} B^{\nu}_0(-X)^{m-\nu}.$$

The term independent of ζ is a polynomial in X, viz.,

$$\sum_{\nu=0}^{m} B^{\nu}_0(-X)^{m-\nu} \equiv Q(-X, 1, 0) \equiv Q(1, \dot{0}, 0) \prod_{j=1}^{m} X - \lambda_j(1), \tag{3.6}$$

taking into account (2.4). Now, the characteristic relation (2.3) defines a family of m phase curves $\lambda = \lambda_1(\xi), \cdots, \lambda_m(\xi)$ within the infinite rectangular strip $-1 \leqq \xi \leqq 1$ of the $\xi - \lambda$ plane. According to Chee-Seng (op. cit.), corresponding to the Laplacian index q of the Q-operator, exactly q phase curves pass through the origin to form a restricted subfamily $\lambda = \lambda_1(\xi), \cdots, \lambda_q(\xi)$, say, with $\lambda_j(0) = 0$ $(j = 1, \cdots, q)$ but $\lambda_j(0) \neq 0$ $(j = q + 1, \cdots, m)$. If $q = 0$, no phase curve goes

through the origin. Combining (2.4) with (3.1) and (3.3), we have, for $q > 0$,

$$\zeta^{q-m}Q(-X,1,\zeta^2) \equiv (1-\xi^2)^{(1/2)(q-m)}Q(1,0,0)\prod_{j=1}^{q}[X-\lambda_j(\xi)/\xi] \cdot \prod_{j=q+1}^{m}\xi X-\lambda_j(\xi) \tag{3.7}$$

inside $-1 < \xi < 1$. Let $\xi \to 0$, in which case $|\zeta| \to \infty$, and the limit of (3.7) then indicates, with reference to (3.5), that the coefficient of ζ^{m-q} in $Q(-X,1,\zeta^2)$ is

$$Q(1,0,0)\prod_{j=1}^{q}[X-\lambda_j'(0)]\prod_{j=q+1}^{m}\lambda_j(0) \tag{3.8}$$

if $q > 0$; but if $q = 0$, the particular coefficient simply equals

$$Q(1,0,0)\prod_{j=1}^{m}\lambda_j(0) \equiv Q(0,0,1) \neq 0. \tag{3.9}$$

Evidently, (3.1) maps those real ξ-zeros of $Q(-\xi X,\xi,1-\xi^2)$ that do not coincide with $\xi = 0$ onto corresponding ζ-zeros of $Q(-X,1,\zeta^2)$ (see (3.3)). In particular, along the Re ζ-axis alone, the latter vanishes at the ζ-images of those real ξ-roots to the set of m equations $\xi X = \lambda_j(\xi)$ $(j = 1,\cdots,m)$ within $0 < |\xi| \leqq 1$. Geometrically, these ξ-roots are the abscissas of those intersections off the λ-axis, but inside the strip $-1 \leqq \xi \leqq 1$, between the straight line $\lambda = \xi X$ and the entire phase curve family $\lambda = \lambda_j(\xi)$ $(j = 1,\cdots,m)$.

In a space of odd n-dimensions, it has been demonstrated (part I) that the axial reception is mostly singular whenever the observer's velocity X takes any of the following discrete values:

$\lambda_1'(0),\cdots,\lambda_q'(0)$, the q radiated singularities of type 1 (r.s.1's), which exist only if $q > 0$;

$\lambda_1(1),\cdots,\lambda_m(1)$, the m radiated singularities of type 2 (r.s.2's);

all those $\lambda_j'(\xi)$'s satisfying, within $0 < \xi < 1$, $\lambda_j'(\xi) = \lambda_j(\xi)/\xi$ $(j = 1,\cdots,m)$, the radiated singularities of type 3 (r.s.3's).

In the present structure, the r.s.1's and r.s.2's coincide with the X-zeros (all real) of the respective polynomial expressions (3.8) and (3.6). To interpret the r.s.3's, we first derive from (3.1), (3.3) and (2.4),

$$m(1+\zeta^2)^{-1}Q(-X,1,\zeta^2) - \zeta^{-1}\,\partial Q(-X,1,\zeta^2)/\partial\zeta \equiv \xi^2 Q(1,0,0)\sum_{j=1}^{m}[X-\lambda_j'(\xi)]\prod_{k=1:k\neq j}^{m}X-\lambda_k(\xi)/\xi. \tag{3.10}$$

The right side holds if $0 < \xi < 1$, corresponding to $0 < \zeta < \infty$ for the left side. We then conclude that the r.s.3's are, in fact, the real X-roots to the two simultaneous equations

$$Q(-X,1,\zeta^2) = 0 \quad \text{and} \quad \partial Q(-X,1,\zeta^2)/\partial\zeta = 0 \quad \text{in } 0 < \zeta < \infty. \tag{3.11}$$

This is because the left equation is satisfied along $0 < \zeta < \infty$ if, and only if, $X = \lambda_j(\xi)/\xi$ within $0 < \xi < 1$ and for some $j \in \{1,\cdots,m\}$, in which event,

$\sum_{j=1}^{m}$ of (3.10) reduces to just

$$[\lambda_j(\xi)/\xi - \lambda_j'(\xi)] \prod_{k=1:k\neq j}^{m} [\lambda_j(\xi) - \lambda_k(\xi)]/\xi, \tag{3.12}$$

so that $\lambda_j'(\xi) = \lambda_j(\xi)/\xi = X$ is, by (3.10), the necessary and sufficient condition for (3.11) to be completely satisfied. The various radiated singularities can also be related to (i) the phase-curve family, (ii) the group velocities of energy transmission for the associated multidimensional waves, (part I). They correspond to the intersections of wavefronts with the x_1-axis.

To evaluate $K(X)$, we suppose that the variable X avoids every r.s.1, r.s.2, and r.s.3. Under such a restriction, the axial radiation reception for odd n (part I) is, mathematically and physically, nonsingular. In view of (3.5) and (3.8) or (3.9), the ζ-polynomial $Q(-X, 1, \zeta^2)$ does attain its highest possible degree of $m - q$, so that

$$\underset{(\text{in } \zeta)}{\text{deg poly}}\,[Q(-X, 1, \zeta^2)] - \underset{(\text{in } \zeta)}{\text{deg poly}}\,[\zeta^{n-2}P(-X, 1, \zeta^2)] \geqq 2 + p - n - q. \tag{3.13}$$

Whence, assuming

$$p \geqq n + q, \tag{3.14}$$

the infinite integral of (3.4) satisfies a convergence requirement. The integrand, extended into the complex ζ-plane, is meromorphic, its only singularities being poles located at the $m - q$ ζ-roots to

$$Q(-X, 1, \zeta^2) = 0. \tag{3.15}$$

These collect in pairs, each symmetric about the origin $\zeta = 0$. Complex roots may be simple or multiple. Existing real roots turn out to be nonvanishing and simple, in which case, the integral of (3.4) can be envisaged in the sense of a Cauchy principal value. (Note that if X coincides with an r.s.2, then according to (3.5) and (3.6), equation (3.15) possesses at $\zeta = 0$ a repeated root having an even order of at least two. In the physical $n = 2$ situation, this root constitutes a pole of the same order, assuming $P(-X, 1, 0) \neq 0$. On the other hand, if X coincides with an r.s.3, then in view of (3.11), one generally encounters two opposing, repeated real poles, each with an order $\geqq 2$. Due to the blockage of its path by repeated poles in both these cases, the particular integral cannot normally be validly interpreted, but actually diverges. The reception is therefore singular.)

Let the given path $(-\infty, \infty)$ for (3.4) be infinitesimally indented about existing real poles, and closed by an infinite semicircle described anticlockwise into the upper half-plane $\operatorname{Im} \zeta > 0$. The diverted contour integration along this semicircle vanishes under a uniform convergence law, on account of (3.13) and (3.14). Thus, in accordance with Cauchy theory, the desired principal value is

$$K(X) = (-1)^{(1/2)n-1}\left\{\sum_{\operatorname{Im}\zeta_\nu = 0} \tfrac{1}{2}i \operatorname{Res}\,[\zeta_\nu(X)] + \sum_{\operatorname{Im}\zeta_\nu > 0} i \operatorname{Res}\,[\zeta_\nu(X)]\right\}, \tag{3.16}$$

where $\sum_{\mathrm{Im}\,\zeta_\nu=0}$ and $\sum_{\mathrm{Im}\,\zeta_\nu>0}$ run over, respectively, all real ζ-roots and all complex ζ-roots within $\mathrm{Im}\,\zeta > 0$ to equation (3.15), while

$$\text{(3.17)}\quad \mathrm{Res}\,[\zeta_\nu(X)] = \frac{1}{(m_\nu - 1)!}\lim_{\zeta\to\zeta_\nu(X)}\left(\frac{\partial}{\partial\zeta}\right)^{m_\nu-1}\left[(\zeta-\zeta_\nu(X))^{m_\nu}\zeta^{n-2}\frac{P(-X,1,\zeta^2)}{Q(-X,1,\zeta^2)}\right],$$

the pertinent residue at the root $\zeta = \zeta_\nu(X)$ of arbitrary order m_ν. Alternatively, a closed contour may be completed by an infinite semicircle described clockwise into $\mathrm{Im}\,\zeta < 0$ instead. Nonetheless this ultimately produces the same end result.

Since real ζ-roots to (3.15) are all of order one and form nonvanishing symmetric pairs of the type $\zeta_\nu(X)$, $-\zeta_\nu(X)$, the net contribution linked with each such pair is

$$\mathrm{Res}\,[\zeta_\nu(X)] + \mathrm{Res}\,[-\zeta_\nu(X)] \equiv 0;$$

the vanishment here follows from the fact that, in this case, (3.17) gives

$$\text{(3.18)}\qquad \mathrm{Res}\,[\zeta_\nu(X)] = \{\zeta^{n-2}P(-X,1,\zeta^2)/\partial Q(-X,1,\zeta^2)/\partial\zeta\}_{\zeta=\zeta_\nu(X)},$$

the curly-bracketed quantity being an odd function of ζ. Regarding (3.16), then, $\sum_{\mathrm{Im}\,\zeta_\nu=0} \equiv 0$. Evidently, the representation (3.16) remains virtually unaltered, by virtue of residue theory, whenever (3.15) possesses no real root; in this event, however, one dispenses with the principal value approach.

Purely imaginary ζ-roots to (3.15) appear in symmetric conjugate pairs of the type $i|\zeta_\nu(X)|$, $-i|\zeta_\nu(X)|$, both being of the same order; between them, only the upper root $i|\zeta_\nu(X)|$ contributes. Since the coefficients of the ζ-polynomial $Q(-X,1,\zeta^2)$ are real, its complex zeros lying off the $\mathrm{Im}\,\zeta$-axis form symmetric quadruple groups of the type

$$\zeta_\nu(X),\quad -\zeta_\nu(X),\quad \overline{\zeta_\nu(X)},\quad -\overline{\zeta_\nu(X)},$$

all of the same order. Here, the bar denotes a complex conjugate. Only two (contributing) elements in each such group are encountered inside $\mathrm{Im}\,\zeta > 0$, these being $\zeta_\nu(X)$ and $-\overline{\zeta_\nu(X)}$, provided $\mathrm{Im}\,\zeta_\nu(X) > 0$. Accordingly, by applying a differentiation rule in conjugate complex variables to (3.17), it can be proved that

$$i\,\mathrm{Res}\,[\zeta_\nu(X)] + i\,\mathrm{Res}\,[-\overline{\zeta_\nu(X)}] = -2\,\mathrm{Im}\,\{\mathrm{Res}\,[\zeta_\nu(X)]\}.$$

Whereupon, (3.16) and (3.17) eventually lead to

$$\begin{aligned} K(X) = {} & 2(-1)^{(1/2)n}\sum_{0<\arg\zeta_\nu<(1/2)\pi}\frac{1}{(m_\nu-1)!} \\ & \cdot \mathrm{Im}\left\{\lim_{\zeta\to\zeta_\nu(X)}\left(\frac{\partial}{\partial\zeta}\right)^{m_\nu-1}\left[(\zeta-\zeta_\nu(X))^{m_\nu}\zeta^{n-2}\frac{P(-X,1,\zeta^2)}{Q(-X,1,\zeta^2)}\right]\right\} \\ & - \sum_{\arg\zeta_\nu=(1/2)\pi}\frac{1}{(m_\nu-1)!}\lim_{\zeta\to|\zeta_\nu(X)|}\left(\frac{\partial}{\partial\zeta}\right)^{m_\nu-1} \\ & \cdot\left[(\zeta-|\zeta_\nu(X)|)^{m_\nu}\zeta^{n-2}\frac{P(-X,1,-\zeta^2)}{Q(-X,1,-\zeta^2)}\right], \end{aligned} \tag{3.19}$$

where the leading sum ranges over every complex root $\zeta = \zeta_\nu$ (of order m_ν) to (3.15) well inside the first quadrant: $0 < \arg \zeta < \frac{1}{2}\pi$, while the succeeding sum ranges over every purely imaginary root $\zeta = i|\zeta_\nu|$ (of order m_ν) along the positive Im ζ-axis: $\arg \zeta = \frac{1}{2}\pi$. For computational convenience, any term, corresponding to $m_\nu \geqq 2$, of the leading sum is expressible as an $m_\nu \times m_\nu$ determinant (cf. Hille (1963)), viz.,

$$\text{Im}\left\{ q_0^{-m_\nu} \begin{vmatrix} q_0 & 0 & 0 & \cdots & p_0 \\ q_1 & q_0 & 0 & \cdots & p_1 \\ q_2 & q_1 & q_0 & \cdots & p_2 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ q_{m_\nu - 1} & q_{m_\nu - 2} & q_{m_\nu - 3} & \cdots & p_{m_\nu - 1} \end{vmatrix} \right\}, \tag{3.20}$$

wherein, for $k = 0, 1, \cdots, m_\nu - 1$,

$$p_k = \frac{1}{k!}\left[\left(\frac{\partial}{\partial \zeta}\right)^k \zeta^{n-2} P(-X, 1, \zeta^2)\right]_{\zeta = \zeta_\nu(X)}, \tag{3.21}$$

$$q_k = \frac{1}{(m_\nu + k)!}\left[\left(\frac{\partial}{\partial \zeta}\right)^{m_\nu + k} Q(-X, 1, \zeta^2)\right]_{\zeta = \zeta_\nu(X)}. \tag{3.22}$$

A similar formula applies to each ($m_\nu \geqq 2$) term in the succeeding sum. For $m_\nu = 1$, the desired form is directly computable. The radiation function (axial solution) for nonsingular reception is determined from (2.9) or (2.10), together with (3.19). Bearing in mind the assumption (3.14), the relevant radiation function is ϕ_1 if $l + q + 1 \leqq n + q \leqq p$; it is ϕ_2 for either $n + q \leqq l + q + 1 \leqq p$, or $n + q \leqq p \leqq l + q + 1$.

Let us account for a way whereby $K(X)$ acquires its various contributions. First, consider the equation

$$Q(-\xi X, \xi, 1 - \xi^2) = 0 \tag{3.23}$$

within the complex ξ-plane, and for nonsingular reception. Only those of its ξ-roots along the real slit $-1 \leqq \xi \leqq 1$ are determinable from the equations $\xi X = \lambda_j(\xi)$ ($j = 1, \cdots, m$). Now, (3.1) maps those real ξ-roots over $1 < |\xi| < \infty$ upon the purely imaginary ζ-roots to (3.15) that are confined to the two adjacent vertical segments $0 < |\text{Im } \zeta| < 1$. Whenever $|X|$ *decreases* (or increases) slightly across $|\lambda_j(1)|$, i.e., an |r.s.2|, an intersection between the line $\lambda = \xi X$ and phase curve $\lambda = \lambda_j(\xi)$ is *broken* (or made) in the region $0 < \xi < 1$, within which, (3.23) therefore *loses* (or gains) a corresponding ξ-root; there are now two possibilities: (i) if $\lambda = \lambda_j(\xi)$ possesses a real continuation slightly beyond $\xi = 1$ and into $\xi > 1$, the particular ξ-root invariably *reappears* within (or departs from) $\xi > 1$, in which case, $K(X)$ *secures* (or forfeits) a contribution, associated with a specific purely imaginary root $\zeta = \zeta_\nu(X)$: $0 < \text{Im } \zeta_\nu(X) < 1$, for (or from) the second component series $\sum_{\arg \zeta_\nu = (1/2)\pi}$ of (3.19); (ii) otherwise, the ξ-root transferred in exchange is either purely imaginary or generally complex, in which case, $K(X)$ again *secures* (or forfeits) an appropriate contribution.

In a parallel situation, the K-element for *odd* n (part I) *forfeits* (or secures) a corresponding contribution instead; actually this K-element is a finite superposition of terms, each jointly contributed by two opposing real ξ-roots (i.e., eigenvalues) to (3.23) along $-1 < \xi < 1$; ξ-roots beyond this interval do not count. *Conversely, for the present even n-problem, it is always the latter set of (exterior) roots, never the former set, that matters* (with regard to the eigenvalue-spectrum). This illustrates one of the more conspicuous discrepancies, in this instance, a mathematical one, between odd and even n-systems. In the following §§ 4 and 5, other significant differences, including physical ones, are also encountered.

4. Physical phenomena. Since, in view of (2.2),

$$\underset{(\text{in } X)}{\deg \text{ poly } Q(-X, 1, \zeta^2)} - \underset{(\text{in } X)}{\deg \text{ poly } P(-X, 1, \zeta^2)} \geqq l \geqq 1,$$

(3.4) discloses that $K(X) = 0$ when $|X| = \infty$ (i.e., when $|x_1| = \infty$). Such a vanishment, in fact, extends inwards from the two ends $X = \pm\infty$. To prove this, we first appeal to an observation in part I, viz., for each of the two partially infinite exteriors

$$(4.1) \qquad -\infty < X < \min_{j=1,\cdots,m} \lambda_j(1), \qquad \max_{j=1,\cdots,m} \lambda_j(1) < X < \infty,$$

the straight line $\lambda = \xi X$ always intersects once, nontangentially and inside the strip $0 < \xi < 1$, every member of a specific subfamily of $\frac{1}{2}(m - q)$ phase curves. By our earlier reasoning, each such intersection associates with a real positive root $\zeta = |\zeta_\nu(X)|$ to (3.15), in which event, $\zeta = -|\zeta_\nu(X)|$ is also a root. Thus all ($m - q$, the maximal number of) ζ-roots are real, nonvanishing and, evidently, simple. Note that no radiated singularity occurs within either X-interval of (4.1). In particular, (3.15) issues no complex (or imaginary) ζ-root to participate in (3.19). Consequently, $K(X) \not\equiv 0$ throughout both X-intervals of (4.1), as well as at their outer limits $X = -\infty, \infty$. These are therefore external ranges of absolute silence with null reception. Their two inner limits are occupied by the r.s.2's $\min_{j=1,\cdots,m} \lambda_j(1)$ and $\max_{j=1,\cdots,m} \lambda_j(1)$, which are incidentally, the respective *fastest* radiated singularities to the left and right of the source point (cf. part I). The time-dependent, axial radiation flux is totally confined to the finite interval·

$$(4.2) \qquad \min_{j=1,\cdots,m} \lambda_j(1) < X < \max_{j=1,\cdots,m} \lambda_j(1)$$

expanding outwards from the source, *a consistency with the radiation principle.* The corresponding phenomenon of silence is also demonstrated for odd n (part I) by a relatively devious technique, however, due to the fact that the K-element is not directly trivial, but actually comprises a maximal set of contributions whose net effect cancels out through fulfillment of auxiliary conditions that play no part in our present problem.

If, for any subinterval of (4.2), there are also exactly $m - q$ real, simple ζ-roots to (3.15), it again followed that $K(X) \equiv 0$ within such a subinterval. This interior

of silence is, effectively, *a linear intercept, with the symmetry axis, of a Petrowsky's lacuna* (Petrowsky (1945), Gårding (1951 and 1970)), commonly associated with the Green's (impulse) function. Our "all real roots" criterion is comparable with a neat sufficiency condition of Burridge (1967) for the formation of a lacuna in two-dimensions (see also Weitzner (1961), Bazer and Yen (1969), Payton (1971)). In the case of odd n, such a criterion is not enough (part I) to establish a lacunary intercept; various additional requirements are necessary, depending on the situation.

In the odd n configuration, the K-element usually vanishes identically (or it may assume a time-independent composition if the Laplacian index $q = 0$) whenever $X \neq \lambda_j(\xi)/\xi$ on $0 < \xi \leqq 1$ for every $j = 1, \cdots, m$, and $X \neq \lambda_j'(0)$ $(j = 1, \cdots, q)$ if $q > 0$. Under these circumstances, X avoids all radiated singularities, and (3.15) admits no real root. However, the result (3.19) neither vanishes nor becomes time-independent. In this instance, the corresponding odd and even n phenomena disagree completely.

In a physically compatible, constantly sustained radiative system, one normally expects a steady (i.e., time-independent) state to develop eventually, if not at finite time, at least after an infinite period, measured from the activation instant. Does our present system conform to this accepted pattern of behavior? It does along the symmetry axis and during time $t = \infty$ if $K(0)$ exists. Now, the strictly elliptic hypothesis (2.5) implies, via (3.3), that $Q(0, 1, \zeta^2) \neq 0$ along the entire Re ζ path. Suppose each existing r.s.1 never vanishes, or equivalently, with reference to the Q-operator of (1.1), (1.6):

$$B^q_{(1/2)(m-q)} = \text{coefficient of } (\partial/\partial x_1)^q(\nabla_1^2)^{(1/2)(m-q)} \neq 0, \tag{4.3}$$

by virtue of (3.5) and (3.8). Note that this automatically holds for $q = 0$ on account of (3.9). In view of (3.8) or (3.9), and (3.13), then, the inequality (3.14), once imposed, continues to satisfy a convergence requirement when $X = 0$, which again overlaps no radiated singularity. At this limit, therefore, the K-element has a valid, *non-singular integral* representation, viz., from (3.4),

$$K(0) = \frac{(-1)^{(1/2)n-1}}{2\pi} \int_{-\infty}^{\infty} \zeta^{n-2} \frac{P(0, 1, \zeta^2)}{Q(0, 1, \zeta^2)} d\zeta. \tag{4.4}$$

It accumulates finite, X-independent residues precipitated by exactly $\frac{1}{2}(m - q)$ zeros (counting multiplicities) of $Q(0, 1, \zeta^2)$ inside Im $\zeta > 0$, in a manner resembling (3.19). Should the latter be practicably explicit in terms of X, then by letting $X \to 0$, one also arrives at $K(0)$. The associated radiation functions ϕ_1 and ϕ_2, defined by (2.9) and (2.10), likewise achieve time-independence. A steady state is thus (axially) attained at instant $t = \infty$. For odd n satisfying $p + 1 > n + q$ (cf. (3.14)), $K(0) \equiv 0$ (part I).

If one is merely concerned with exploring the steady state of, say, pulsatory radiation, it seems expedient to ignore initial conditions and substitute them with a radiation condition incorporated in the manner of Lighthill (1960, 1965, 1967); see also Chee-Seng (1971). Such an application would be redundant for the present

unsteady analysis. Instead, *the agreement with the radiation principle, as well as the steady state development, is an indirect outcome of the zero initial condition* (1.2) *enforced upon* (1.1). A related discussion on this matter can be found in a current paper (Chee-Seng (1974)).

5. Magnetogasdynamic flow aligned reception. Weitzner's (1961) paper essentially dealt with the two-dimensional magnetogasdynamic Green's function for a stationary fluid. Burridge (1967) applied his own construction to confirm the lacunas. The three-dimensional problem has been successfully attempted much earlier by Friedlander (1959). Owing to the complicated anisotropy, results are unavoidably presented in rather general forms, principally in terms of geometrical characteristics. Unfortunately, explicit reduced versions along the axes of symmetry have been apparently overlooked.

The magnetogasdynamics of a perfect fluid are basically governed by first order equations of continuity, state, momentum and induction, the last of which being compatible with a no-divergence constraint on the solenoidal magnetic field (see, e.g., Kulikovskiy and Lyubimov (1965)). Consider an unbounded uniform flow with velocity $\mathbf{V}$ past a weak body force $\mathbf{f}$ (per unit mass). The equations of motion can then be linearized and combined into a second order form, governing the velocity perturbation $\mathbf{v}$:

$$[D^2/Dt^2 - (\mathbf{a}\cdot\nabla)^2]\mathbf{v} + [\mathbf{a}(\mathbf{a}\cdot\nabla) - (a^2 + c^2)\nabla](\nabla\cdot\mathbf{v}) + \nabla(\mathbf{a}\cdot\nabla)(\mathbf{a}\cdot\mathbf{v}) = D\mathbf{f}/Dt, \tag{5.1}$$

where $\mathbf{a}$ is the Alfvén velocity parallel to the equilibrium magnetic field, $a = |\mathbf{a}|$, c is the sound speed in the compressible medium, ∇ is the gradient operator, while $D/Dt \equiv \partial/\partial t + \mathbf{V}\cdot\nabla$ denotes differentiation following the motion. Note that $c \neq 0$; also, ignoring ordinary (i.e., nonmagnetic) gas flow, $a \neq 0$.

Let us propose that the body force exerts constantly, starting from the instant $t = 0$, is point-localized at the origin of a two-dimensional $\mathbf{r} = (x, y)$ coordinate frame, and acts transversely across the equilibrium field with which the positive x-direction is aligned. In particular then,

$$\mathbf{f} = H(t)\delta(\mathbf{r})(0, 1), \tag{5.2}$$

while $\mathbf{a} = (a, 0)$. Naturally, any perturbation depends exclusively on $\mathbf{r}$ and t, and commences only at instant $t = 0$. It can then be shown from (5.1) that $v_2 = v_2(\mathbf{r}, t)$, viz., the component of $\mathbf{v}$ perpendicular to $\mathbf{a}$, satisfies

$$G\left(\frac{D^2}{Dt^2}, \frac{\partial^2}{\partial x^2}, \frac{\partial^2}{\partial y^2}\right)v_2 = F\left(\frac{D}{Dt}, \frac{\partial^2}{\partial x^2}\right)H(t)\delta(\mathbf{r}), \tag{5.3}$$

in which the operators

$$F\left(\frac{D}{Dt}, \frac{\partial^2}{\partial x^2}\right) \equiv \left(\frac{D^2}{Dt^2} - c^2\frac{\partial^2}{\partial x^2}\right)\frac{D}{Dt}, \tag{5.4}$$

$$G\left(\frac{D^2}{Dt^2}, \frac{\partial^2}{\partial x^2}, \frac{\partial^2}{\partial y^2}\right) \equiv \frac{D^4}{Dt^4} - (a^2 + c^2)\nabla^2\frac{D^2}{Dt^2} + a^2c^2\nabla^2\frac{\partial^2}{\partial x^2}, \tag{5.5}$$

are homogeneous in D/Dt, $\partial/\partial x$ and $\partial/\partial y$. But for the convective part $\mathbf{V}\cdot\nabla$ of D/Dt, the radiative system here would have been simultaneously symmetric about both x and y axes.

In the following § 6, a cross-flow configuration is considered. Meanwhile, throughout the rest of this section, we assume that the flow and the permeating field are originally in alignment (includes possible opposition), viz., $\mathbf{V} = (V, 0)$. Let $x = x_1$, $y = x_2$. So $D/Dt \equiv \partial/\partial t + V\partial/\partial x_1$. A uniaxial symmetry (about the x_1-axis) is thus imparted to the motion governed by (5.3); this now takes the form (1.1) with

$$P\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x_1}, \nabla_1^2\right) \equiv F\left(\frac{D}{Dt}, \frac{\partial^2}{\partial x_1^2}\right), \tag{5.6}$$

which, incidentally, is independent of ∇_1^2, while

$$Q\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x_1}, \nabla_1^2\right) \equiv G\left(\frac{D^2}{Dt^2}, \frac{\partial^2}{\partial x_1^2}, \nabla_1^2\right), \tag{5.7}$$

the F- and G-operators being given by (5.4) and (5.5) wherein $\nabla^2 \equiv \partial^2/\partial x_1^2 + \nabla_1^2$. Evidently, these operator forms are consistent with (1.5) and (1.6); here, $m = 4$, $l = 1$, while the respective Laplacian indices $p = 4$ and $q = 2$.

Now, from (5.5) and (5.7): $Q(1, 0, 0) \equiv G(1, 0, 0) \equiv 1$, so that (2.2) is satisfied. Also, the characteristic polynomial

$$Q(-\lambda, \xi, 1 - \xi^2) \equiv G\{(\lambda - V\xi)^2, \xi^2, 1 - \xi^2\} \equiv \prod_{j=1}^{4} \lambda - \lambda_j(\xi) \tag{5.8}$$

for $-1 \leqq \xi \leqq 1$, whereon, by introducing

$$\chi_+(\xi) = \{\tfrac{1}{2}(a^2 + c^2 + \sqrt{(a^2 + c^2)^2 - 4a^2c^2|\xi|^2})\}^{1/2}, \tag{5.9}$$

$$\chi_-(\xi) = \begin{cases} -\{\tfrac{1}{2}(a^2 + c^2 - \sqrt{(a^2 + c^2)^2 - 4a^2c^2|\xi|^2})\}^{1/2}, & -1 \leqq \xi < 0, \quad (5.10) \\ 0, & \xi = 0, \quad (5.11) \\ \{\tfrac{1}{2}(a^2 + c^2 - \sqrt{(a^2 + c^2)^2 - 4a^2c^2|\xi|^2})\}^{1/2} & 0 < \xi \leqq 1, \quad (5.12) \end{cases}$$

we have

$$\lambda_1(\xi) \equiv V\xi + \chi_-(\xi), \qquad \lambda_2(\xi) \equiv V\xi - \chi_-(\xi), \tag{5.13}$$

$$\lambda_3(\xi) \equiv V\xi + \chi_+(\xi), \qquad \lambda_4(\xi) \equiv V\xi - \chi_+(\xi). \tag{5.14}$$

Throughout $-1 \leqq \xi \leqq 1$, $\chi_+(\xi)$ and $\chi_-(\xi)$ are real and continuous, and $\chi_+(\xi) \neq 0$, whilst $\chi_-(\xi)$ vanishes only at $\xi = 0$. Furthermore, for $-1 < \xi < 1$, $\chi_+(\xi) \neq \pm\chi_-(\xi)$. Note that, generally, $\chi_+(\pm 1) = \max(a, c)$, $\chi_-(1) = \min(a, c) = -\chi_-(-1)$. If $a = c$, then $\chi_+(\pm 1) = a = \chi_-(1) = -\chi_-(-1)$. Thus the functions $\lambda_j(\xi)$ $(j = 1, 2, 3, 4)$, as defined by (5.13) and (5.14), satisfy a strictly hyperbolic condition of being real and distinct for every $\xi \in (-1, 0)$ and $(0, 1)$ as well as at both endpoints

$\xi = +1, -1$ provided $a \neq c$. If $a = c$, the hyperbolicity is nonstrict at $\xi = +1, -1$. The strictly elliptic condition depicted by (2.5) is satisfied if and only if within the vertical strips $0 < |\xi| \leqq 1$, the nonvertical line $\lambda = V\xi$ ($|V| < \infty$) never intersects any one of the four phase curves for a stationary fluid, viz., $\lambda = \chi_+(\xi)$, $-\chi_+(\xi)$, $\chi_-(\xi)$, $-\chi_-(\xi)$. However, we shall allow for possible intersections, corresponding to nonstrict ellipticity.

There are two r.s.1's at

$$\begin{aligned} \lambda_1'(0) &= V + \lim_{\xi \to 0} \chi_-(\xi)/\xi = V + ac(a^2 + c^2)^{-1/2}, \\ \lambda_2'(0) &= V - \lim_{\xi \to 0} \chi_-(\xi)/\xi = V - ac(a^2 + c^2)^{-1/2}; \end{aligned} \tag{5.15}$$

and four r.s.2's at

$$\begin{aligned} \lambda_1(1) &= V + \min(a, c), \qquad \lambda_2(1) = V - \min(a, c), \\ \lambda_3(1) &= V + \max(a, c), \qquad \lambda_4(1) = V - \max(a, c). \end{aligned} \tag{5.16}$$

From (5.5) and (5.7), the function

$$\begin{aligned} Q(-X, 1, \zeta^2) &\equiv G\{(X - V)^2, 1, \zeta^2\} \\ &\equiv [a^2c^2 - (a^2 + c^2)(X - V)^2]\zeta^2 \\ &\quad + [(X - V)^2 - a^2][(X - V)^2 - c^2]. \end{aligned} \tag{5.17}$$

Note that the term independent of ζ agrees with (3.6) via (5.16), while the coefficient of ζ^2 agrees with (3.8) via (5.9), (5.14) and (5.15). Also from (5.4) and (5.6),

$$P(-X, 1, \zeta^2) \equiv F(V - X, 1) \equiv (V - X)[(V - X)^2 - c^2]. \tag{5.18}$$

Suppose X avoids all six r.s.1's and r.s.2's. Then the ζ-polynomial expression (5.17) maintains its maximal degree of $m - q = 2$, and always vanishes off $\zeta = 0$ at $\zeta = \zeta_1(X)$, $-\zeta_1(X)$:

$$\zeta_1(X) \equiv \{[(X - V)^2 - a^2][(X - V)^2 - c^2]/[(a^2 + c^2)(X - V)^2 - a^2c^2]\}^{1/2}. \tag{5.19}$$

Both these zeros being obviously distinct, the second equation in (3.11) is never satisfied. Consequently, there are no r.s.3's. Since $p - n - q = 0$, the condition (3.14) holds. Although we are relaxing the strictness hypothesis on hyperbolicity (by allowing for the possibility $a = c$), as well as on ellipticity, nonetheless the general result (3.19) for $K(X)$ remains applicable.[1] Whereupon, for a magnetically aligned flow and reception: $\mathbf{V} = (V, 0)$ and $\mathbf{r} = (x, 0) \neq \mathbf{0}$ during $t > 0$, the solution to (5.3) is, via (2.9) or (2.10),

$$v_2 = (\pi x)^{-1} K(X), \tag{5.20}$$

with $X = x/t$.

Both $\zeta_1(X)$, $-\zeta_1(X)$ are real if and only if either

$$|X - V| > \max(a, c), \tag{5.21}$$

[1] This is guaranteed by a general rule stipulated in the last paragraph of § 8.

or

$$ac(a^2 + c^2)^{-1/2} < |X - V| < \min(a, c). \tag{5.22}$$

Within the corresponding intervals, absolute quiet therefore prevails:

$$K(X) \equiv 0, \tag{5.23}$$

by virtue of arguments in § 3. The radiation principle is clearly verified with reference to the two outermost intervals described by (5.21). These are precisely of the types of (4.1). They contract as the two fastest radiated singularities $\lambda_3(1)$ and $\lambda_4(1)$, coincident with their inner limits, retreat right and left, respectively, from a flow-transported center at $\mathbf{r} = (Vt, 0)$. The other two intervals of silence defined by (5.22) are interior to both $\lambda_3(1)$ and $\lambda_4(1)$. Each is, therefore, the linear intercept of a Petrowsky's lacuna with the equilibrium line of force through the source point. For a stationary ($V \equiv 0$) fluid, this feature supports both Weitzner's (1961) and Burridge's (1967) conclusions (see also Bazer and Yen (1969)). The three-dimensional stationary fluid does not apparently propagate such a lacunary intercept (part I).

For the remaining possibilities, viz.,

$$|X - V| < ac(a^2 + c^2)^{-1/2}, \tag{5.24}$$

$$\min(a, c) < |X - V| < \max(a, c), \tag{5.25}$$

the simple ζ-zero $\zeta_1(X) = i|\zeta_1(X)|$, a purely imaginary value. In accordance with (3.19), it is the sole contributing element, in this case, to the sum $\sum_{\arg \zeta_\nu = (1/2)\pi}$; the accompanying zero $-\zeta_1(X)$ plays no part. Hence, via (5.17)–(5.19), one eventually arrives at

$$K(X) = \frac{\frac{1}{2}(X - V)|(X - V)^2 - c^2|^{1/2}}{[|(X - V)^2 - a^2|\,|(a^2 + c^2)(X - V)^2 - a^2c^2|]^{1/2}} \tag{5.26}$$

throughout the innermost interval expressed by (5.24). Along the two complementary intervals corresponding to (5.25), the result is practically the same except for an extra sign factor of $\operatorname{sgn}(|X - V| - c)$. This close relationship seems to be missing for the three-dimensional stationary fluid (part I).

6. Aligned reception in cross flow. Suppose the flow originally traverses the magnetic field orthogonally instead, i.e., $\mathbf{V} = (0, V)$. Now choose $x = -x_2$, $y = x_1$. Then, again, $D/Dt \equiv \partial/\partial t + V\partial/\partial x_1$, while (5.3), (5.4) and (5.5) fall within the class comprising (1.1), (1.5), (1.6), but now with

$$P\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x_1}, \nabla_1^2\right) \equiv F\left(\frac{D}{Dt}, \nabla_1^2\right), \tag{6.1}$$

$$Q\left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x_1}, \nabla_1^2\right) \equiv G\left(\frac{D^2}{Dt^2}, \nabla_1^2, \frac{\partial^2}{\partial x_1^2}\right). \tag{6.2}$$

Thus, a (transverse) symmetry about the x_1-axis is maintained. Again $m = 4$, $l = 1$; but now, the Laplacian indices $p = 2$ and $q = 0$. Also, $Q(1, 0, 0) \equiv 1$

(see (2.2)). Defining

$$\Omega_{\pm}(\xi) = \{\tfrac{1}{2}(a^2 + c^2 \pm \sqrt{(a^2 + c^2)^2 - 4a^2c^2(1 - \xi^2)})\}^{1/2}, \tag{6.3}$$

$$\lambda_1(\xi) \equiv V\xi + \Omega_-(\xi), \qquad \lambda_2(\xi) \equiv V\xi + \Omega_+(\xi), \tag{6.4}$$

$$\lambda_3(\xi) \equiv V\xi - \Omega_-(\xi), \qquad \lambda_4(\xi) \equiv V\xi - \Omega_+(\xi), \tag{6.5}$$

we have, from (6.2) and (5.5),

$$Q(-\lambda, \xi, 1 - \xi^2) \equiv G\{(\lambda - V\xi)^2, 1 - \xi^2, \xi^2\} \equiv \prod_{j=1}^{4} \lambda - \lambda_j(\xi). \tag{6.6}$$

Within $-1 \leqq \xi \leqq 1$, all four $\lambda_j(\xi)$'s are real and continuous; they are also distinct for real ξ satisfying $0 < |\xi| < 1$, equivalent to a strict hyperbolicity. Since $q = 0$, no r.s.1 exists. Regarding the r.s.2's, two overlap: $\lambda_1(1) = V = \lambda_3(1)$, whilst the remaining two, viz., $\lambda_2(1) = V + \sqrt{a^2 + c^2}$ and $\lambda_4(1) = V - \sqrt{a^2 + c^2}$, are distinct. Evidently, the hyperbolicity is nonstrict at $\xi = 1$. This is also true at $\xi = -1$, with $\lambda_1(-1) = -V = \lambda_3(-1)$, but $\lambda_2(-1) \neq \lambda_4(-1)$. At $\xi = 0$, the hyperbolicity is nonstrict if, and only if, $a \neq c$. Otherwise, when $a = c$, $\lambda_1(0) = \lambda_2(0) = a$ and $\lambda_3(0) = \lambda_4(0) = -a$, a violation of strictness. As in § 5, we do accommodate nonstrict ellipticity, corresponding to vanishments of $\lambda_j(\xi)$ $(j = 1, 2, 3, 4)$ within $-1 \leqq \xi \leqq 1$. Note a compounded nonstrictness in both hyperbolicity and ellipticity whenever $V = 0 : \lambda_1(\pm 1) = \lambda_3(\pm 1) = 0$.

From (6.1) and (5.4), we have the P-polynomial

$$P(-X, 1, \zeta^2) \equiv F(V - X, \zeta^2) \equiv (V - X)[(V - X)^2 - c^2\zeta^2] \tag{6.7}$$

which, unlike the form (5.18), is dependent on ζ. Furthermore, from (6.2) and (5.5), the Q-polynomial

$$Q(-X, 1, \zeta^2) \equiv G\{(X - V)^2, \zeta^2, 1\} \equiv a^2c^2\zeta^4 - Q_2(X)\zeta^2 + Q_0(X), \tag{6.8}$$

where

$$Q_0(X) \equiv Q(-X, 1, 0) \equiv (X - V)^2[(X - V)^2 - (a^2 + c^2)] \tag{6.9}$$

(cf. (3.6)), and

$$Q_2(X) \equiv (a^2 + c^2)(X - V)^2 - a^2c^2. \tag{6.10}$$

The ζ-expression (6.8) is permanently of degree $m - q = 4$. Its four zeros occur at $\zeta = \zeta_+(X), -\zeta_+(X), \zeta_-(X), -\zeta_-(X)$:

$$\zeta_{\pm}(X) = 2^{-1/2}(ac)^{-1}\{Q_2(X) \pm \sqrt{Q_2^2(X) - 4a^2c^2Q_0(X)}\}^{1/2}. \tag{6.11}$$

We now assume that X avoids all four r.s.2's, so that $Q_0(X) \neq 0$. Then all four zeros determined by (6.11) are nonvanishing. Moreover, they are all distinct since, from (6.9) and (6.10),

$$Q_2^2(X) - 4a^2c^2Q_0(X) \equiv (a^2 - c^2)^2(X - V)^4 + a^2c^2[2(a^2 + c^2)(X - V)^2 + a^2c^2] > 0.$$

In particular, then, there are no r.s.3's. The rule (3.14) is satisfied because $p - n - q = 0$. Once again, in spite of nonstrictness, one can still construct the K-element from (3.19). In terms of this K-element, the stream-aligned solution to (5.3) for cross-flow is, via (2.9) or (2.10),

$$v_2 = (\pi y)^{-1} K(X) \tag{6.12}$$

at point $\mathbf{r} = (0, y) \neq \mathbf{0}$ and instant $t > 0$; here $X = y/t$.

When

$$|X - V| > \sqrt{a^2 + c^2} \tag{6.13}$$

(the magnetosonic speed), we have $Q_0(X) > 0$ and $Q_2(X) > 0$. Both pairs of ζ-zeros $\pm\zeta_+(X)$, $\pm\zeta_-(X)$ are consequently real. Hence, again, as may be anticipated from a radiation principle, there is absolute quiet: $K(X) \equiv 0$ within two outermost intervals, this time associated with (6.13) (cf. (4.1)).

In the alternative situation,

$$0 < |X - V| < \sqrt{a^2 + c^2}, \tag{6.14}$$

one gets $Q_0(X) < 0$. So, $\zeta_+(X)$, $-\zeta_+(X)$ remain real and, therefore, noncontributing towards $K(X)$. However, $\zeta_-(X) = i|\zeta_-(X)|$, the only ζ-zero that counts towards formula (3.19), its complement $-\zeta_-(X)$ being a nonparticipant. Thus, employing (6.7) and (6.8), we subsequently arrive at

$$K(X) = \frac{(X - V)[(X - V)^2 + c^2|\zeta_-(X)|^2]}{2|\zeta_-(X)|[Q_2^2(X) - 4a^2c^2Q_0(X)]^{1/2}}, \tag{6.15}$$

valid in the inner interval (6.14) about, and avoiding, its flow-transported center. This is the repeated r.s.2 $\lambda_1(1) = V = \lambda_3(1)$. Evidently, our reception path never intercepts a Petrowsky's lacuna. The remaining r.s.2's, i.e., $\lambda_2(1)$ and $\lambda_4(1)$, form the outer limits of interval (6.14). One can show from (6.9)–(6.11) and (6.15) that as $X \to \lambda_j(1)$ ($j = 2$ or 4) from the inside, $K(X) \to \infty$ as $|X - \lambda_j(1)|^{-1/2}$. It then undergoes an infinite jump to zero once X crosses $\lambda_j(1)$.

7. Nonstrict conditions. Under strictly hyperbolic and elliptic conditions, the ω-integral in (2.1) is representable within the ξ-integration ranges $-1 \leqq \xi < 0$ and $0 < \xi \leqq 1$ by (2.6), wherein

$$\operatorname*{residue}_{\omega=0}(\cdot) = (i\alpha)^{-1} \frac{P(0, \xi, 1 - \xi^2)}{Q(0, \xi, 1 - \xi^2)}, \tag{7.1}$$

$$\operatorname*{residue}_{\omega=\alpha\lambda_j(\xi)}(\cdot) = (i\alpha)^{-1} \frac{P(-\lambda_j(\xi), \xi, 1 - \xi^2) \exp[-i\alpha\lambda_j(\xi)t]}{\lambda_j(\xi)[\partial Q(-\lambda, \xi, 1 - \xi^2)/\partial\lambda]_{\lambda=\lambda_j(\xi)}}, \tag{7.2}$$

which can be easily shown via (1.7).

Suppose the strictness on ellipticity is relaxed, say, at any ξ_0 within either ξ-range:

$$\lambda_\nu(\xi_0) = 0 \quad \text{for some } \nu \in \{1, 2, \cdots, m\} \qquad (0 < |\xi_0| \leqq 1), \tag{7.3}$$

i.e., the phase curve $\lambda = \lambda_\nu(\xi)$ meets the ξ-axis at the point $(\xi_0, 0)$; in this case, (2.4) discloses

$$Q(0, \xi_0, 1 - \xi_0^2) = 0. \tag{7.4}$$

Such a phenomenon is indeed encountered for magnetogasdynamic flow either aligned with or crossing the magnetic field (§§ 5, 6). Expression (7.1) invariably attains infinity at $\xi = \xi_0$; so does (7.2) when $j = \nu$. Now consider the joint $\xi = 0$, which, under the transformation (3.1) splits up into the two infinite limits of integration for (3.4). Unless the Laplacian index q of the Q-operator is zero, there is a permanent gap in the strictness at this joint, because here, q of the $\lambda_j(\xi)$'s vanish concurrently: $\lambda_j(0) = 0\,(j = 1, \cdots, q)$, and so $Q(0, 0, 1) = 0$. Again a failure of (7.1), as well as (7.2) for $j = 1, \cdots, q$, occurs, this time at $\xi = 0$. Although the implication has been initially overlooked in § 2, nevertheless, the difficulty posed is successfully confronted in the Appendix.

Alternatively, the hyperbolicity is nonstrict at $\xi = \xi_0 \in [-1, 0)$ or $(0, 1]$ provided the equation

$$Q(-\lambda, \xi_0, 1 - \xi_0^2) = 0 \tag{7.5}$$

has a real (repeated) λ-root, say, $\lambda = \lambda_\nu(\xi_0)$ of arbitrary order $n_\nu \geqq 2$ (but $\leqq m$). This happens whenever n_ν phase curves, one of them being $\lambda = \lambda_\nu(\xi)$, intersect at a point with abscissa ξ_0. Corresponding to each member $\lambda = \lambda_j(\xi)$ of the intersecting set, including $\lambda = \lambda_\nu(\xi)$:

$$[\partial^\kappa Q(-\lambda, \xi_0, 1 - \xi_0^2)/\partial\lambda^\kappa]_{\lambda = \lambda_j(\xi_0)} = 0 \qquad (\kappa = 0, 1, \cdots, n_\nu - 1), \tag{7.6}$$

but $\neq 0$ when $\kappa = n_\nu$. Yet again, (7.2) is singular at $\xi = \xi_0$, and for every such j. The applications in §§ 5 and 6 also cover possible nonstrict hyperbolicity. An example of a nonstrict hyperbolic equation of fourth order is briefly explored in Courant and Hilbert (1962), with special reference to the equation of crystal optics; the argument permits two simple roots to the dispersion (i.e., characteristic) relation to approach one another as a vector parameter approaches a critical value, thus producing a double root in the limiting coincidence.

The difficulty is compounded if both types of nonstrictness overlap at the same ξ-value, viz., $\lambda = 0$ is a repeated root to (7.5). Such a situation definitely exists at the joint $\xi = 0$ if the Laplacian index $q \geqq 2$, the vanishing root being of order q and corresponds to $\lambda = \lambda_j(0)\,(j = 1, \cdots, q)$. It also arises at $\xi = \pm 1$ during (i) a transverse reception in a stationary conducting gas (§ 6): $\lambda_1(\pm 1) = \lambda_3(\pm 1) = 0$; (ii) a magnetically aligned reception and sonic-flow when the sound and Alfvén speeds are equal (§ 5): $\lambda_2(1) = \lambda_4(1) = 0$, $\lambda_2(-1) = \lambda_3(-1) = 0$.

The reason (7.1) and (7.2) fail at $\xi = \xi_0$ is because each pertinent residue is structured on the basis that the precipitating pole at $\omega = 0$ or $\omega = \alpha\lambda_j(\xi)$ is uniformly simple on $-1 \leqq \xi \leqq 1$, a fact no longer true. Instead, multiplicities are now achieved at discrete ξ-parameters by one or more of these poles. Since these continue to block the real ω-path $(-\infty, \infty)$, the earlier Cauchy principal value interpretation of (2.7), via a multi-indented contour, is no longer feasible for any

such ξ-parameter, at which (2.8) therefore becomes suspect under classical complex theory. This in turn reflects controversially upon formula (2.11). To remedy the defect, one needs to return to (2.1). Adhering to definition (2.9) or (2.10), the K-element can be rewritten as

$$K = (-1)^{(1/2)n-1}(2\pi)^{-1}\left(\int_{-1}^{0} + \int_{0}^{1}\right)(1-\xi^2)^{(1/2)(n-3)}\xi^{l-n}L(x_1, t; \xi)\,d\xi, \tag{7.7}$$

where

$$\begin{aligned} L(x_1, t; \xi) &= (4\pi)^{-1}\xi x_1 \int_{-\infty}^{\infty} (\operatorname{sgn}\alpha)(i\alpha)^l \exp(i\alpha\xi x_1)\,d\alpha \\ &\cdot \int_{-\infty+i\varepsilon}^{\infty+i\varepsilon} \frac{P(-i\omega, i\alpha\xi, \alpha^2\xi^2 - \alpha^2)}{Q(-i\omega, i\alpha\xi, \alpha^2\xi^2 - \alpha^2)} \frac{e^{-i\omega t}}{\omega}\,d\omega \end{aligned} \tag{7.8}$$

which should, by comparison with (2.11), reduce to

$$L(x_1, t; \xi) \equiv P(-\xi X, \xi, 1-\xi^2)/Q(-\xi X, \xi, 1-\xi^2) \tag{7.9}$$

(with $X = x_1/t$) throughout $-1 \leqq \xi < 0$ and $0 < \xi \leqq 1$, for a strictly hyperbolic cum elliptic system. However, regarding the applications in §§ 5 and 6, we asserted that the general solution (3.19) holds in spite of nonstrictness. This solution stems from the integral (3.4) which is related to (2.11) via the mapping (3.1). For our assertion to qualify, the uniform validity of (7.9) must first be established under nonstrictness.

That $L(x_1, t; \xi)$ depends on X, rather than x_1 and t separately, can be immediately shown by suitably transforming both integration variables in (7.8) and exploiting (1.7). Thus

$$\begin{aligned} L(x_1, t; \xi) &= (4\pi)^{-1}\xi X \int_{-\infty}^{\infty} (\operatorname{sgn}\alpha)\exp(i\alpha\xi X)\,d\alpha \\ &\cdot \int_{\mathscr{L}} \frac{P(-\omega\alpha^{-1}, \xi, 1-\xi^2)}{Q(-\omega\alpha^{-1}, \xi, 1-\xi^2)} \frac{e^{-i\omega}}{\omega}\,d\omega, \end{aligned} \tag{7.10}$$

a function of X and ξ. Here, $\mathscr{L}$ denotes the horizontal path traced from $\omega = (-\infty + i\varepsilon)t$ to $(\infty + i\varepsilon)t$. As previously, $\varepsilon > 0$. Assuming, again, the primary hyperbolic postulate (2.2), Jordan's lemma is now fully satisfied only in the ω half-space below $\mathscr{L}$. Suppose an infinite, closed, semicircular contour $\mathscr{L}^*$ is completed from $\mathscr{L}$ within this half-space. The relevant ω-integration along the semicircular segment vanishes.

Consider the statement embodying (7.5), ignoring for the moment the compounded case where $\lambda_\nu(\xi_0) = 0$. The present implication is that $\omega = \alpha\lambda_\nu(\xi_0)$ is a real pole of order n_ν of the ω-integrand in (7.10) when $\xi = \xi_0 (\neq 0)$. The situation at $\xi = 0$ differs appreciably, and is handled separately (see Appendix). If $t < 0$, every such pole lies outside the contour $\mathscr{L}^*$ and, consistent with (1.2), therefore contributes nothing. However, as we normally assume $t > 0$, the particular pole

falls inside $\mathscr{L}^*$ and thus contributes, appropriately,

$$-2\pi i \operatorname*{residue}_{\omega=\alpha\lambda_\nu(\xi_0)} (\cdot)_{\xi_0} = \frac{-2\pi i}{(n_\nu - 1)!} \lim_{\omega\to\alpha\lambda_\nu(\xi_0)} \left(\frac{\partial}{\partial\omega}\right)^{n_\nu - 1} \{(\omega - \alpha\lambda_\nu(\xi_0))^{n_\nu}(\cdot)_{\xi_0}\}$$

$$= -2\pi i \lim_{\lambda\to\lambda_\nu(\xi_0)} \left(\frac{\partial}{\partial\lambda}\right)^{n_\nu - 1} \tag{7.11}$$

$$\cdot \left\{ \frac{(\lambda - \lambda_\nu(\xi_0))^{n_\nu} P(-\lambda, \xi_0, 1 - \xi_0^2) \exp(-i\alpha\lambda)}{(n_\nu - 1)!\lambda Q(-\lambda, \xi_0, 1 - \xi_0^2)} \right\},$$

a corresponding correction to (7.2). Note the present analyticity of the curly bracketed λ-function in the limit: $\lambda = \lambda_\nu(\xi_0)$. Incorporating (7.11) into (7.10) after a Leibnitz's differentiation, we deduce that $L(x_1, t; \xi_0)$ acquires from $\lambda_\nu(\xi_0)$ the contribution

$$\lim_{\lambda\to\lambda_\nu(\xi_0)} \sum_{k=0}^{n_\nu - 1} \binom{n_\nu - 1}{k} \left(\frac{\partial}{\partial\lambda}\right)^{n_\nu - 1 - k} \left\{ \frac{(\lambda - \lambda_\nu(\xi_0))^{n_\nu} P(-\lambda, \xi_0, 1 - \xi_0^2)}{(n_\nu - 1)!\, \lambda Q(-\lambda, \xi_0, 1 - \xi_0^2)} \right\}$$

$$\cdot (2i)^{-1} \xi_0 X \int_{-\infty}^{\infty} (\operatorname{sgn}\alpha)(-i\alpha)^k \exp[i\alpha(\xi_0 X - \lambda_\nu(\xi_0))]\, d\alpha. \tag{7.12}$$

Suppose, for each $\lambda_\nu(\xi_0)$,

$$\xi_0 X \neq \lambda_\nu(\xi_0), \tag{7.13}$$

which, as we shall see in § 8, is necessary for boundedness of the K-element. Then, within the context of generalized functions (Lighthill (1958)), the α-integral of (7.12) is a Fourier transform with the finite value

$$2ik!\,(\xi_0 X - \lambda_\nu(\xi_0))^{-k-1} \equiv \lim_{\lambda\to\lambda_\nu(\xi_0)} 2i(\partial/\partial\lambda)^k (\xi_0 X - \lambda)^{-1}. \tag{7.14}$$

Whereupon, by reversing Leibnitz's process, the expansion (7.12) is summable to yield the closed form

$$\xi_0 X \operatorname*{residue}_{\lambda=\lambda_\nu(\xi_0)} \left[\frac{P(-\lambda, \xi_0, 1 - \xi_0^2)}{\lambda(\xi_0 X - \lambda) Q(-\lambda, \xi_0, 1 - \xi_0^2)} \right], \tag{7.15}$$

after noting the fact that $\lambda = \lambda_\nu(\xi_0)$ is a pole of order n_ν of the function $[\cdot]$. Actually, the form (7.15) does not explicitly involve the order n_ν. It always determines the contribution to $L(x_1, t; \xi_0)$ from a typical root to (7.5), including any simple root. In considering the latter case, one takes $n_\nu = 1$.

Suppose the ellipticity is nonstrict in accordance with (7.4), but that

$$\partial Q(-\lambda, \xi_0, 1 - \xi_0^2)/\partial\lambda \neq 0 \quad \text{at } \lambda = 0, \qquad (\xi_0 \neq 0). \tag{7.16}$$

So $\lambda = 0$ constitutes a simple root to (7.5), corresponding to the vanishment of exactly one of the $\lambda_j(\xi)$'s at $\xi = \xi_0$. But owing to the linear factor ω of its denominator, the ω-integrand in (7.10) now possesses, when $\xi = \xi_0$, a double pole at $\omega = 0$, which is also surrounded by the contour $\mathscr{L}^*$. To rectify (7.1), one merely substitutes $n_\nu = 2$ and $\lambda_\nu(\xi_0) = 0$ everywhere in (7.11). Likewise, in

the compounded situation where several of the $\lambda_j(\xi)$'s vanish concurrently when $\xi = \xi_0$, one again substitutes $\lambda_\nu(\xi_0) = 0$ throughout (7.11). However, $n_\nu - 1$ must now be replaced by the multiplicity of the root at $\lambda = 0$ to (7.5). Whatever this multiplicity ($\geqq 1$), (7.13) must be accounted for; the following is conveniently inferred: if $X \neq 0$, $L(x_1, t; \xi_0)$ always acquires an odd contribution representable by (7.15), but with $\lambda_\nu(\xi_0) \equiv 0$. This is also generally true if the ellipticity is strict at $\xi = \xi_0$, and corresponds to having $n_\nu = 1$.

Ranging over all poles of the specific ω-integrand, one arrives at

$$L(x_1, t; \xi_0) = \xi_0 X \left\{ \operatorname*{residue}_{\lambda=0} [\,\cdot\,] + \sum_\nu \operatorname*{residue}_{\lambda=\lambda_\nu(\xi_0)} [\,\cdot\,] \right\}, \tag{7.17}$$

summed over all λ-zeros possessed by the denominator factor $\lambda Q(-\lambda, \xi_0, 1 - \xi_0^2)$ of the function $[\,\cdot\,]$ displayed in (7.15). Now, within the complex λ-plane, $\lambda[\,\cdot\,] \to 0$ uniformly over $0 \leqq \arg \lambda < 2\pi$ as $|\lambda| \to \infty$, because, in view of (2.2),

$$\underset{(\text{in } \lambda)}{\deg \text{poly}} [(\xi_0 X - \lambda) Q(-\lambda, \xi_0, 1 - \xi_0^2)] - \underset{(\text{in } \lambda)}{\deg \text{poly}} P(-\lambda, \xi_0, 1 - \xi_0^2)$$

$$\geqq l + 1 \geqq 2.$$

Hence

$$\oint [\,\cdot\,]\, d\lambda \equiv 0,$$

with the integration performed over an infinite closed contour, which obviously circumscribes all those λ-poles participating in (7.17). However (7.15) reveals that, under stipulations typified by (7.13) plus $X \neq 0$, $[\,\cdot\,]$ possesses an extra simple pole at $\lambda = \xi_0 X$. Consequently, by virtue of residue theory,

$$\begin{aligned} L(x_1, t; \xi_0) &= -\xi_0 X \operatorname*{residue}_{\lambda=\xi_0 X} [\,\cdot\,] \\ &\equiv P(-\xi_0 X, \xi_0, 1 - \xi_0^2)/Q(-\xi_0 X, \xi_0, 1 - \xi_0^2), \end{aligned} \tag{7.18}$$

which confirms the desired uniform validity.

Apart from (7.13), we also assumed $X \neq 0$. (In § 8, both these conditions are further discussed in relation to the K-element). Otherwise, with reference to (7.12), the α-integral corresponding to the pole $\lambda = \lambda_\nu(\xi_0)$, or $\lambda = 0$, diverges. With $\lambda = 0$, this divergent tendency is averted by the linear factor X (preceding the integral sign, see (7.12)), provided the ellipticity is strict. In this case, $\lambda = 0$ is a simple pole of $[\,\cdot\,]$. Thus by merely computing $\xi_0 X \operatorname{residue}_{\lambda=0} [\,\cdot\,]$ from (7.15), substituting into (7.17), and then letting $X \to 0$, we get in the limit,

$$L(x_1, \infty; \xi_0) = P(0, \xi_0, 1 - \xi_0^2)/Q(0, \xi_0, 1 - \xi_0^2), \tag{7.19}$$

which expectantly agrees with the limit of (7.18).

The above proof relies upon $\xi_0 \neq 0$, and cannot apply to $\xi_0 = 0$. However, compatibility of formula (7.9) at the joint $\xi = 0$ is also available, as demonstrated in the Appendix.

8. Effects on K-evaluation. The transformation (3.1) converts (2.11) into (3.4), whose evaluation is possible under certain conditions, one of which being that X avoids every radiated singularity. This must now be revised, among other things, within the context of nonstrictness. The representation (3.4) remains correct since that of (2.11) is proven to be so.

With nonstrict hyperbolicity, another class of radiation singularities arises, and must also be avoided during reception. Previously, on the basis of strict hyperbolicity, (3.11) was satisfied under a certain necessary and sufficient condition. This requires modification. Thus according to (3.10) and (3.12), (3.11) is now fully satisfied at a given $\zeta \in (0, \infty)$ if and only if, at the corresponding $\xi \in (0, 1)$, there exists $j \in \{1, 2, \cdots, m\}$: either (i) $X = \lambda_j(\xi)/\xi = \lambda_j'(\xi)$, or (ii) $X = \lambda_j(\xi)/\xi = \lambda_k(\xi)/\xi$ for at least one integer $k \in \{1, 2, \cdots, j-1, j+1, \cdots, m\}$, or (iii) both possibilities (i) and (ii) concur. The particular ζ-root corresponds to a multiple pole blocking the real path of integration for $K(X)$, thereby inducing divergence. Reception is consequently singular. (Note that both ζ- and X-roots to (3.11) essentially depend only on the B_μ^ν-coefficients in (1.6).) Possibility (i) states that X coincides with the r.s.3 $\lambda_j'(\xi)$; this is the case met in § 3. Regarding (ii), the hyperbolicity is nonstrict at the particular $\xi \in (0, 1)$, for which $\lambda_j(\xi)/\xi$ may be defined as a *pseudo-type* 3 radiated singularity, which is now coincident with X; however in the event of a concurrence with possibility (i), this singularity is identifiable as an r.s.3. Hence the X-roots to (3.11) locates not only the r.s.3's, but also a separate set of pseudo r.s.3's.

Corresponding to nonstrict hyperbolicity when $\xi = 1$, at least two of the r.s.2's overlap over one or more values (see, e.g., §§ 5 and 6). Otherwise, their influence is not significantly new.

Now, if parameter ζ satisfies (3.11) for a given X, then so does $-\zeta$, but along the left ζ-interval $(-\infty, 0)$. Both these parameters are, essentially, images under the mapping (3.1) of two symmetric ξ-points within, respectively, $(0, 1)$ and $(-1, 0)$. Our above reasoning leads to the interesting inference that, whenever the hyperbolicity is nonstrict for the real parametric ranges: $0 < |\xi| < 1$, it is invariably so at symmetrical, discrete ξ-pairs of the type $\{\xi, -\xi\}$. (Note that any pseudo r.s.3 always associates with one such pair, just as each r.s.3 associates with a specific ξ-pair too.) This symmetry principle also applies to nonstrict ellipticity since the left side of (7.4) is, from (1.6), an even function of ξ_0. In fact, the principle covers nonstrictness at both ends $\xi = 1, -1$ as well.

The condition (7.13) safeguards the finiteness of $L(x_1, t; \xi_0)$. It also happens to keep X away from the pseudo r.s.3 $\lambda_\nu(\xi_0)/\xi_0$. On the other hand, suppose $\lambda_\mu(\xi_0)$ is a nonrepeated λ-root to (7.5), noncoincident with $\xi_0\lambda_\mu'(\xi_0)$, and where $\xi_0 \in (-1, 0)$ or $(0, 1)$. Then $\lambda_\mu(\xi_0)/\xi_0$ is not a radiated singularity of any (so far as yet) known type. Hence a reception at

$$X = \lambda_\mu(\xi_0)/\xi_0 \tag{8.1}$$

is not normally expected to be singular. Nonetheless, it blatantly violates the boundedness of $L(x_1, t; \xi)$ at $\xi = \xi_0$. The question therefore arises of the latter's behavior about $\xi = \xi_0$. This is the center of a sufficiently narrow ξ-neighborhood

throughout which $\lambda_\mu(\xi)$ is a simple root to (2.3), and so (cf. (7.15)) imparts to $L(x_1, t; \xi)$ the near-singular term

$$\frac{\xi X P(-\lambda_\mu(\xi), \xi, 1-\xi^2)}{\lambda_\mu(\xi)[\xi X - \lambda_\mu(\xi)][\partial Q(-\lambda, \xi, 1-\xi^2)/\partial\lambda]_{\lambda=\lambda_\mu(\xi)}}; \tag{8.2}$$

near-singular because $\xi X - \lambda_\mu(\xi) \approx (\xi - \xi_0)[X - \lambda'_\mu(\xi_0)]$ under (8.1). The present method provides an alternative, more direct and elementary approach to (an approximation of) (7.18). Starting from (2.4), it is easily seen that the factor

$$[\partial Q(-\lambda, \xi, 1-\xi^2)/\partial\lambda]_{\lambda=\lambda_\mu(\xi)} \approx Q(1,0,0) \prod_{j:j\neq\mu} \xi_0 X - \lambda_j(\xi_0), \tag{8.3}$$

$$= [X - \lambda'_\mu(\xi_0)]^{-1} \cdot [\partial Q(-\xi X, \xi, 1-\xi^2)/\partial\xi]_{\xi=\xi_0}, \tag{8.4}$$

noting that $Q(-\xi_0 X, \xi_0, 1-\xi_0^2) = 0$. The residue-type contributions from the remaining λ-roots to (2.3), as well as from $\lambda = 0$, are bounded inside the particular ξ-neighborhood; consequently (cf. (7.17)),

$$L(x_1, t; \xi) \approx \frac{P(-\xi_0 X, \xi_0, 1-\xi_0^2)}{(\xi - \xi_0)[\partial Q(-\xi X, \xi, 1-\xi^2)/\partial\xi]_{\xi=\xi_0}}. \tag{8.5}$$

But the right side here is the complete principal part to a Laurent's expansion, viz., the term dominating the right most expression of (7.18). So the consistency is preserved right up to, but excluding the point $\xi = \xi_0$. By comparing (8.5) against (3.18) and accounting for (8.3), (8.4) and (3.10), it is readily deduced that the K-element acquires from $\xi = \xi_0$ (whose image under (3.1) constitutes a real simple pole at $\zeta = \zeta_0$, say), the temporary but nonsingular contribution proportional to

$$\text{Res}\,[\zeta_0] = -\operatorname*{residue}_{\xi=\xi_0} \{(1-\xi^2)^{(1/2)(n-3)} \xi^{l-n} L(x_1, t; \xi)\}, \tag{8.6}$$

and which is annulled by an opposing contribution obviously associated with $-\xi_0$ (see § 3). Note that $\{\cdot\}$ in fact encloses the ξ-integrand in (7.7), or (2.11).

As already indicated (§ 7), if the Laplacian index $q \neq 0$, nonstrictness occurs at the joint $\xi = 0$ which becomes infinite under transformation (3.1). The only bearing, with regard to ζ-integration for (3.4), is upon the behavior at infinity of the ζ-integrand. The latter continues to behave desirably if (3.14) remains true, and if X avoids every r.s.1 at $\lambda'_j(0)$ $(j = 1, \cdots, q)$. Moreover, if each such $\lambda'_j(0) \neq 0$, a steady state criterion stays satisfied (§ 4; see also Appendix).

It now remains to examine the effects of nonstrict ellipticity over $-1 \leqq \xi < 0$ and $0 < \xi \leqq 1$. Thus, suppose (7.3) holds. Then, owing to (7.4) and (3.3), $Q(0, 1, \zeta^2)$ has a real zero at $\zeta = \zeta_0 \equiv \xi_0^{-1}(1-\xi_0^2)^{1/2}$. (Note that the converse is also true). The steady state analysis for $X = 0$ is obviously effected. Unlike previously, only the principal value of the integral (4.4) is now conceivable. One sufficiency condition for its existence is that the zero ζ_0 be simple, i.e., via (3.10),

$$[\partial Q(0, 1, \zeta^2)/\partial\zeta]_{\zeta=\zeta_0} \equiv -\xi_0^{2-m}(1-\xi_0^2)^{1/2} Q(1,0,0)\lambda'_\nu(\xi_0) \prod_{j:j\neq\nu} \lambda_j(\xi_0) \neq 0;$$

actually valid whenever $0 < |\xi_0| \leqq 1$; so (i) $\xi_0 \neq \pm 1$, (ii) $\lambda_\nu'(\xi_0) \neq 0$, and (iii) $\lambda_j(\xi_0) \neq 0$ with $j = 1, 2, \cdots, \nu - 1, \nu + 1, \cdots, m$. Now, (i) implies that the ellipticity at $\xi = \pm 1$ is strict, or equivalently, no r.s.2 ever vanishes. If this is contradicted, a relevant multiple pole tends to form at $\zeta = 0$ and, hence, invalidate the integration. However, it is quite possible that this tendency may be arrested, e.g., when $n > 2$, or if $P(0, 1, \zeta^2)$ vanishes simultaneously at $\zeta = 0$. According to (ii) and (iii) together, the parameter ξ_0 does not associate with any vanishing r.s.3 or pseudo r.s.3. Thus, if no radiated singularity ever vanishes, the steady state solution is always determinable under (3.14) from (4.4). With the appropriate attainment of nonstrict ellipticity at various ξ-values, their corresponding contributions eliminate each other. In particular, if nonstrict ellipticity occurs at exactly $\frac{1}{2}(m - q)$ distinct ξ-values within $(0, 1)$, it follows that the ultimate steady state is invariably one of absolute quiet, in contrast (cf. § 4) to that in a strictly elliptic system. Thus, for the magnetically aligned flow (§ 5) which is either supersonic-superAlfvénic, or satisfies $ac(a^2 + c^2)^{-1/2} < |V| < \min(a, c)$, the axial reception during $t = \infty$ should be absolutely nil. This must also be the case with a supermagnetosonic cross-flow (§ 6).

Note. In arriving at (7.18), we specified that $X \neq 0$. This was subsequently recognized as being superfluous whenever the ellipticity is strict at $\xi = \xi_0$, say. Otherwise, the expression (7.19) would collapse under (7.3), even with the present conditions (i)–(iii) imposed. Nonetheless, this does not disqualify the admission of $X = 0$ for K-evaluation. The situation is clearly analogous to that posed by (8.1): in both parallels, the L-function has a singularity (precisely, a simple pole) at $\xi = \xi_0$, but which contributes no singular term at all to the K-element.

Overall, the only disadvantage is, at most, an increased number of convergence conditions required to absorb the additional effects of nonstrict hyperbolicity and nonstrict ellipticity in the consideration of (3.4) and (4.4) respectively. The *general rule*, covering nonstrictness, is as follows: provided, for the particular X, that the rational function

$$\zeta^{n-2} P(-X, 1, \zeta^2)/Q(-X, 1, \zeta^2) \tag{8.7}$$

has neither a pole (invariably multiple) at $\zeta = 0$ nor any (other) multiple real ζ-pole, and that the degree, with reference to ζ, of its denominator exceeds that of its numerator by at least two, the solution (3.19) holds. Under the same set of circumstances, but for $X = 0$, a steady state exists during $t = \infty$; its K-element is then derived from (3.19) by letting $X \to 0$, or, independently, through (4.4).

Note. The requirements on (8.7) are asserted in the most stringent sense. For their satisfaction, it is sufficient that every existing r.s.1, r.s.2, r.s.3 or pseudo r.s.3 never coincide with X (possibly zero), and that (3.14) apply.

Appendix. The objective of this Appendix is to seal a consistency of the integrand in (2.11), or equivalently of (7.9), at the gap $\xi = 0$ separating both integral ranges. Attention is confined as usual to $t > 0$. The starting point is (7.8) which, on transforming the α-integration variable to $\alpha\xi^{-1}$ and using (1.7),

becomes

$$L(x_1, t; \xi) = (4\pi)^{-1} x_1 \int_{-\infty}^{\infty} (\operatorname{sgn} \alpha) \exp(i\alpha x_1)\, d\alpha \int_{-\infty+i\varepsilon}^{\infty+i\varepsilon} \frac{P(-\xi\omega\alpha^{-1}, \xi, 1-\xi^2)}{Q(-\xi\omega\alpha^{-1}, \xi, 1-\xi^2)} \frac{e^{-i\omega t}}{\omega}\, d\omega, \tag{A.1}$$

where $\varepsilon > 0$. The situation at the gap must be explored by letting $\xi \to 0$ under the integral signs.

First, suppose the Laplacian index $q = 0$ (as encountered, e.g., in the application in § 6). Then according to (3.9), $Q(0, 0, 1) \neq 0$. Now,

$$\int_{-\infty}^{\infty} (\operatorname{sgn} \alpha) \exp(i\alpha x_1)\, d\alpha \int_{-\infty+i\varepsilon}^{\infty+i\varepsilon} \omega^{-1} \exp(-i\omega t)\, d\omega = 4\pi x_1^{-1},$$

which can be easily verified from an ω-contour integration coupled with the result (7.14) for the α-integral of (7.12); whence (A.1) implies

$$\lim_{\xi \to 0} L(x_1, t; \xi) = \frac{P(0, 0, 1)}{Q(0, 0, 1)} = \lim_{\xi \to 0} \frac{P(-\xi X, \xi, 1-\xi^2)}{Q(-\xi X, \xi, 1-\xi^2)}. \tag{A.2}$$

So (7.9) holds in the limit as $\xi \to 0$. Observe that it vanishes during this limit unless the Laplacian index of the P-operator is exactly minimal at $p = l$.

The case wherein $q \neq 0$ is less straightforward. Now, the ξ-integrand in (2.11) involves the factor

$$\xi^{l-n} P(-\xi X, \xi, 1-\xi^2)/Q(-\xi X, \xi, 1-\xi^2) = \xi^{p-q-n} R(X, \xi), \tag{A.3}$$

say, which must (cf. (7.7)) be compared against

$$\xi^{l-n} L(x_1, t; \xi) = \xi^{p-q-n} S(x_1, t; \xi), \tag{A.4}$$

where, via (A.1) and (A.3),

$$S(x_1, t; \xi) = (4\pi)^{-1} x_1 \int_{-\infty}^{\infty} (\operatorname{sgn} \alpha) \exp(i\alpha x_1)\, d\alpha \int_{-\infty+i\varepsilon}^{\infty+i\varepsilon} \omega^{-1} R(\omega\alpha^{-1}, \xi) \exp(-i\omega t)\, d\omega. \tag{A.5}$$

The quantity $R(X, \xi)$, introduced in (A.3), is known from (1.5) and (1.6), and found to be a rational function of X and ξ. In fact,

$$\lim_{\xi \to 0} \xi^{l-p} P(-\xi X, \xi, 1-\xi^2) = \sum_{\nu=0}^{p-l} A^{\nu}_{(1/2)(m-p)} (-X)^{p-l-\nu} = Y(X), \tag{A.6}$$

say, an X-polynomial whose degree $\leqq p - l$, while

$$\lim_{\xi \to 0} \xi^{-q} Q(-\xi X, \xi, 1-\xi^2) = \sum_{\nu=0}^{q} B^{\nu}_{(1/2)(m-q)} (-X)^{q-\nu} = Z(X), \tag{A.7}$$

say, another X-polynomial, which also serves as the coefficient of ζ^{m-q} in (3.5). From (3.8), then,

$$Z(X) = Q(1,0,0)\prod_{j=1}^{q}[X - \lambda_j'(0)]\prod_{j=q+1}^{m}\lambda_j(0), \tag{A.8}$$

whose degree is q, since $Q(1,0,0) \neq 0$ and $\lambda_j(0) \neq 0$ $(j = q+1, \cdots, m)$. Assuming nonsingular reception (§§ 3 and 4), so that X avoids every r.s.1,

$$X \neq \lambda_j'(0) \qquad (j = 1, \cdots, q). \tag{A.9}$$

Thus, from (A.3), (A.6) and (A.7),

$$\lim_{\xi \to 0} R(X, \xi) = Y(X)/Z(X), \tag{A.10}$$

a bounded rational function, in which event, the expression (A.3) and, subsequently, the ξ-integrand of (2.11) as well, approach zero (or $Y(X)/Z(X)$) under the inequality (3.14) as $\xi \to 0$. This should likewise be suspected of the expression (A.4). For a closer consistency, however, the two functions $R(X, \xi)$ and $S(x_1, t; \xi)$ must approach the same limiting value together. Note their permanent coincidence away from the limit (cf. § 7).

The limit of (A.5) is formally expressible, via (A.10), as

$$\begin{aligned}\lim_{\xi \to 0} S(x_1, t; \xi) = (4\pi)^{-1}x_1 \int_{-\infty}^{\infty} (\operatorname{sgn}\alpha) Y\left(i\alpha^{-1}\frac{\partial}{\partial t}\right) \exp(i\alpha x_1)\, d\alpha \\ \cdot \int_{-\infty+i\varepsilon}^{\infty+i\varepsilon} \frac{\exp(-i\omega t)}{\omega Z(\omega\alpha^{-1})}\, d\omega.\end{aligned} \tag{A.11}$$

The ω-integrand is meromorphic, possessing, in view of (A.8), $q+1$ real poles at

$$\omega = 0, \qquad \omega = \alpha\lambda_j'(0) \qquad (j = 1, \cdots, q). \tag{A.12}$$

Furthermore, it fully satisfies Jordan's lemma in the region $\operatorname{Im}\omega < \varepsilon$, which obviously contains all $q+1$ poles. To arrive at a steady state, it is stipulated in § 4 that every r.s.1 never vanishes (cf. (4.3)). Suppose, additionally, that the r.s.1's are all distinct. (This is exactly the situation for a flow and magnetically aligned magnetogasdynamic reception. The distinctness of both r.s.1's is explicit from (5.15), which also discloses their nonvanishment if the flow velocity $V \neq \pm ac(a^2 + c^2)^{-1/2}$). Consequently, the $q+1$ poles defined by (A.12) are all simple. Whereupon, residue theory yields for the ω-integral,

$$\int_{-\infty+i\varepsilon}^{\infty+i\varepsilon} = -2\pi i\left\{\frac{1}{Z(0)} + \sum_{j=1}^{q}\frac{\exp[-i\alpha\lambda_j'(0)t]}{\lambda_j'(0)Z'(\lambda_j'(0))}\right\}. \tag{A.13}$$

Following a modified reverse argument, as in dealing with (2.6) and (2.7), it is not difficult to demonstrate that the right side of (A.13) is actually the Cauchy principal value of

$$2(\operatorname{sgn}\alpha)\int_{-\infty}^{\infty}\frac{\exp(-i\alpha\lambda t)}{\lambda Z(\lambda)}\, d\lambda.$$

Thus, (A.11) subsequently becomes

$$\text{(A.14)} \qquad \lim_{\xi \to 0} S(x_1, t; \xi) = (2\pi)^{-1} X \int_{-\infty}^{\infty} \exp(i\alpha X)\, d\alpha \int_{-\infty}^{\infty} \frac{Y(\lambda)}{\lambda Z(\lambda)} \exp(-i\alpha\lambda)\, d\lambda$$

$$\text{(A.15)} \qquad \equiv Y(X)/Z(X) = \lim_{\xi \to 0} R(X, \xi),$$

on account of a Fourier transformation rule and (A.10). This completes the consistency test.

It seems paradoxical that only the compounded nonstrictness (when $q \neq 0$) effects the analysis at $\xi = 0$. Unlike the case away from this joint, any nonstrictness in pure hyperbolicity alone (i.e., corresponding to coincidences of $\lambda_{q+1}(0)$, $\lambda_{q+2}(0), \cdots, \lambda_m(0)$) plays no role. This is somewhat linked with the fact that the various nontrivial modes are attributed to the gradients $\lambda_j'(0)$ ($j = 1, \cdots, q$) in the restricted subfamily of q phase curves rather than, as might perhaps be originally anticipated, the m ordinates $\lambda_j(0)$ ($j = 1, \cdots, m$) of the complete family. As a consequence, the apparent nature of the local system is, virtually, strictly hyperbolic-cum-elliptic under the assumptions made.

REFERENCES

[1] J. BAZER AND D. H. Y. YEN (1969), *Lacunas of the Riemann matrix of symmetric hyperbolic systems in two space variables*, Comm. Pure Appl. Math., 22, pp. 279–333.

[2] L. BERS, F. JOHN AND M. SCHECHTER (1964), *Partial Differential Equations*, Interscience, New York.

[3] R. BURRIDGE (1967), *Lacunas in two-dimensional wave propagation*, Proc. Cambridge Philos. Soc., 63, pp. 819–825.

[4] L. CHEE-SENG (1971), *Isotropic radiation from a steadily pulsating multidimensional distribution*, Proc. Roy. Soc. London Ser. A, 323, pp. 555–580.

[5] ——— (1973), *Axial radiation reception* (referred to as part I), Proc. Cambridge Philos. Soc., 74, pp. 369–395.

[6] ——— (1974), *Unsteady current-induced perturbation of a magnetically contained magneto-hydrodynamic flow*, J. Fluid Mech., 63, pp. 273–299.

[7] R. COURANT AND D. HILBERT (1962), *Methods of Mathematical Physics*, vol. 2, Interscience, New York.

[8] F. G. FRIEDLANDER (1959), *Sound pulses in a conducting medium*, Proc. Cambridge Philos. Soc., 55, pp. 341–367.

[9] L. GÅRDING (1951), *Linear hyperbolic partial differential equations with constant coefficients*, Acta Math., 85, pp. 1–62.

[10] ——— (1970), *Theory of lacunas*, Hyperbolic Equations and Waves (M. Froissart, ed., Battelle Seattle 1968 Recontres), Springer-Verlag, Berlin.

[11] E. HILLE (1963), *Analytic Function Theory*, vol. 1, Blaisdell, New York.

[12] F. JOHN (1955), *Plane Waves and Spherical Means*, Interscience, New York.

[13] A. G. KULIKOVSKIY AND G. A. LYUBIMOV (1965), *Magnetohydrodynamics*, Addison-Wesley, Reading, Mass.

[14] M. J. LIGHTHILL (1958), *Fourier Analysis and Generalized Functions*, Cambridge University Press, London.

[15] ——— (1960), *Studies on magnetohydrodynamic waves and other anisotropic wave motions*, Philos. Trans. Roy. Soc. London Ser. A, 252, pp. 397–430

[16] ——— (1965), *Group velocity*, J. Inst. Math. Appl., 1, pp. 1–28.

[17] ——— (1967), *Waves generated in dispersive systems by travelling forcing effects*, J. Fluid Mech., 27, pp. 725–752.

[18] R. G. PAYTON (1971), *Two-dimensional elastic waves emanating from a point source*, Proc. Cambridge Philos. Soc., 70, pp. 191–210.

[19] I. G. PETROWSKY (1945), *Diffusion of waves and the lacunas for hyperbolic equations*, Rec. Math. [Mat. Sb.], 17, pp. 289–368.

[20] H. WEITZNER (1961), *Green's function for two-dimensional magnetohydrodynamic waves*, Phys. Fluids, 4, pp. 1238–1245.

SIAM J. MATH. ANAL.
Vol. 8, No. 1, February 1977

SUBFUNCTIONS AND DISTRIBUTIONAL INEQUALITIES*

ROBERT CARMIGNANI AND KEITH SCHRADER†

Abstract. The subfunctions for linear second order differential equations and the locally bounded subfunctions for differential equations of the form $y'' = f(x, y)$ are shown to be locally Lipschitzian and are characterized by continuity and the satisfaction of the appropriate "distributional differential inequality"; moreover, the strict subfunctions are classified by "strict differential distributional inequality". Then the existence of solutions to differential inequalities in the distributional sense are used to establish the existence of classical solutions to boundary value problems.

Introduction. For a second order differential equation

$$y'' = f(x, y, y') \tag{1}$$

with $f:(a, b) \times R \times R \to R$ continuous and such that for arbitrary x_1, y_1, x_2, y_2 with $a < x_1 < x_2 < b$ there is a unique solution y of (1) satisfying $y(x_\alpha) = y_\alpha$ ($\alpha = 1, 2$), a function $u:(a, b) \to R$ is called a subfunction (or strict subfunction) on (a, b) for (1) if for any interval $[x_1, x_2] \subset (a, b)$, we have that $u(x) \leqq$ (or $<$) $y(x)$ for all x in (x_1, x_2) where y is the solution to (1) passing through the points $(x_1, u(x_1))$ and $(x_2, u(x_2))$. It is known that a function $u \in C^2(a, b)$ is a subfunction on (a, b) for (1) if and only if

$$u''(x) \geqq f(x, u(x), u'(x)) \quad \text{for all } x \in (a, b). \tag{2}$$

In fact, this is true under weaker conditions than the existence and uniqueness of solutions to boundary value problems of (1) (for this and for the development of the theory of subfunctions, see [1], [2], [5], [9], [11], [12], [13], [14] and [15]).

The main purpose of this paper is to show that for certain types of second order differential equations the corresponding locally bounded subfunctions (although boundedness is not assumed in the linear case), in general, are locally Lipschitzian and they are characterized by continuity and the satisfaction (interpreted in the sense of distributions) of (2), where strict distributional inequality in (2) classifies the strict subfunctions. This extends the differential inequality test for C^2 subfunctions since a C^2 function u satisfies (2) if and only if it satisfies (2) in the distributional sense. However a C^2 function can satisfy strict inequality in (2) in the distributional sense but not in the ordinary sense.

In §§ 2 and 3 for the case that (1) is linear and homogenous or nonhomogenous, respectively, we obtain our distributional derivative test for subfunctions from the variational characterization of subfunctions of second order linear homogeneous differential equations given by Reid in [14]. Then in § 4, we establish (without assuming existence of solutions to boundary value problems) this test for subfunctions (and strict subfunctions) in the nonlinear case when (1) is of the form $y'' = f(x, y)$. The proofs in this case and in the linear case rely on the following well-known (see [17]) characterization of the convex functions (the subfunctions

* Received by the editors November 19, 1974, and in revised form July 24, 1975.
† Department of Mathematics, University of Missouri, Columbia, Missouri 65201.

of $y'' = 0$): a function u is convex on (a, b) if and only if u is continuous on (a, b) and the second distributional derivative of u is a positive distribution. We also use this fact to show that for these types of differential equations a function is a subfunction on (a, b) if it is locally a subfunction on (a, b) (sometimes called a subfunction in the small on (a, b)).

In §§ 2, 3 and 4, we also indicate the analogous results for superfunctions (functions which are "concave" with respect to solutions of (1)).

In § 5 we prove that if u and v are locally integrable, locally bounded functions on (a, b) with $u(x) \leqq v(x)$ a.e., and if $v''(x) \leqq f(x, v(x))$ and $u''(x) \geqq f(x, u(x))$ hold in the distributional sense, then on any interval $[c, d] \subset (a, b)$ there exists a solution y to $y'' = f(x, y)$ such that $u(x) \leqq y(x) \leqq v(x)$ a.e. on $[c, d]$.

1. Preliminaries. A test function in (a, b) is any infinitely differentiable function with compact support in (a, b). The vector space of all test functions in (a, b) is denoted by $C_0^\infty((a, b))$. The set of all distributions in (a, b) is denoted by $\mathscr{D}'((a, b))$. We denote by $\mathscr{D}((a, b))$ the space $C_0^\infty((a, b))$ equipped with a topology which makes $\mathscr{D}'((a, b))$ its dual space.

For u a locally Lebesgue integrable function on (a, b), we denote by T_u the distribution in (a, b) defined by $T_u(\phi) = \int_a^b u(x)\phi(x)\,dx$ for each $\phi \in \mathscr{D}((a, b))$ and the kth distributional derivative of u, denoted by $\mathscr{D}^k u$, is defined by

$$\mathscr{D}^k u(\phi) = (-1)^k T_u(\phi^{(k)})$$

(when $k = 1$, we write $\mathscr{D}u$ rather than $\mathscr{D}^1 u$).

We say that $T \in \mathscr{D}'((a, b))$ is positive (or strictly positive) and write $T \geqq 0$ (or $T > 0$), if $T(\phi) \geqq 0$ (or > 0) for all nonnegative test functions $\phi (\phi \not\equiv 0)$ in (a, b). For T_1 and T_2 in $\mathscr{D}'((a, b))$, $T_1 \geqq T_2$ (or $T_1 > T_2$) means that $T_1 - T_2$ is positive (or strictly positive).

$L^2(R)$ denotes the space of measurable functions with Lebesgue integrable squares on R. The space of functions η which are absolutely continuous on any compact subinterval $[c, d]$ of (a, b) and $\eta' \in L^2[c, d]$ is denoted by $A^{\text{loc}}(a, b)$. The subspace of $A^{\text{loc}}(a, b)$ of functions η with compact support in (a, b) is denoted by $A_0^2(a, b)$. A function u is said to be locally Lipschitzian on (a, b) if for each compact subinterval $[c, d]$ of (a, b) there exists a constant M (depending on $[c, d]$) such that $|u(x) - u(y)| \leqq M|x - y|$ for all x and y in $[c, d]$. The space of k-times continuously differentiable functions on (a, b) is denoted by $C^k(a, b)$, where for the continuous functions on (a, b) (when $k = 0$) we write $C(a, b)$.

2. The homogenous linear equation $y'' = p_1(x)y' + p_2(x)y$. Let r, p and q be continuous functions on (a, b) with $r > 0$ on (a, b). Consider the self-adjoint differential equation

$$(r(x)y' + q(x)y)' - (q(x)y' + p(x)y) = 0. \tag{2.1}$$

Here we shall be concerned with equations which possess the following property:

(P) If $a < x_1 < x_2 < b$ and y_1 and y_2 are arbitrary real numbers, then there exists a unique solution y of (2.1) such that $y(x_\alpha) = y_\alpha$ $(\alpha = 1, 2)$.

Let u be any function in $A^{\text{loc}}(a, b)$. Then u' is a locally square integrable function on (a, b). Let g and h be functions on (a, b) such that $g = ru' + qu$ a.e. and

$h = qu' + pu$ a.e. in (a, b), where p, q and r are continuous on (a, b).

Let $\mathscr{R}$ be the distribution defined in (a, b) by

$$\mathscr{R}(\phi) = \mathscr{D}g(\phi) - T_h(\phi) \quad \text{for } \phi \in \mathscr{D}((a, b)). \tag{2.2}$$

Let $\tilde{\mathscr{R}}$ be the natural extension of $\mathscr{R}$ to $A_0^2(a, b)$, i.e.,

$$\tilde{\mathscr{R}}(\eta) = -\int_a^b g(x)\eta'(x)\,dx - \int_a^b h(x)\eta(x)\,dx \quad \text{for } \eta \in A_0^2(a, b).$$

THEOREM 2.1 (Reid). *Let p, q and r be in $C(a, b)$ and suppose that* (2.1) *has property* (P). *Then u is a subfunction for* (2.1) *on (a, b) if and only if $u \in A^{\text{loc}}(a, b)$, and $\tilde{\mathscr{R}}(\eta) \geqq 0$ for all nonnegative $\eta \in A_0^2(a, b)$.*

Proof. This theorem is merely a restatement of Theorem 3.1 of [14, p. 575] where (3.3) of [14] is replaced by (3.3′) and where $\tilde{\mathscr{R}}$ is tested for each nonnegative (rather than nonpositive) $\eta \in \Gamma_0(c, d)$ ($\Gamma_0(c, d)$ as defined in [14]) by defining η to be zero outside of (c, d). This completes the proof.

Remark. Note that for each $\eta \in A_0^2(a, b)$, there exists an interval $[c, d] \subset (a, b)$ such that the restriction of η to $[c, d]$ is an element of $\Gamma_0(c, d)$.

THEOREM 2.2. *The distribution $\mathscr{R}$ in (a, b) (as defined in* (2.2)*) is positive if and only if $\tilde{\mathscr{R}}$ is a positive functional on $A_0^2(a, b)$ (i.e., $\tilde{\mathscr{R}}(\eta) \geqq 0$ for all nonnegative $\eta \in A_0^2(a, b)$).*

Proof. Since $\mathscr{R}$ is the restriction of $\tilde{\mathscr{R}}$ to $\mathscr{D}((a, b))$, we have that $\mathscr{R}$ is positive whenever $\tilde{\mathscr{R}}$ is a positive functional.

On the other hand, suppose that $\mathscr{R} \geqq 0$. Let η be any nonnegative element of $A_0^2(a, b)$ and let $\{\eta_\varepsilon : \varepsilon > 0\}$ be the family of regularizations of η (regularization is defined in [8, p. 3]). Then each η_ε is nonnegative since $\eta \geqq 0$, and for ε sufficiently small, $\eta_\varepsilon \in \mathscr{D}((a, b))$. When $\varepsilon \to 0$, we have that $\eta_\varepsilon \to \eta$ uniformly. From this it follows that $\tilde{\mathscr{R}}(\eta) \geqq 0$ since $\mathscr{R} \geqq 0$ and $\tilde{\mathscr{R}}(\eta) = \lim_{\varepsilon \to 0} \mathscr{R}(\eta_\varepsilon)$. The proof is complete.

From Theorems 2.1 and 2.2 we obtain the next theorem.

THEOREM 2.3. *Let p, q and r be in $C(a, b)$ and suppose that* (2.1) *has property* (P). *Then u is a subfunction for* (2.1) *on (a, b) if and only if $u \in A^{\text{loc}}(a, b)$ and $\mathscr{R} \geqq 0$.*

Suppose that the differential equation

$$y'' = p_1(x)y' + p_2(x)y \tag{2.3}$$

satisfies property (P) where p_1 and p_2 are in $C(a, b)$. Let $r_0 = \exp(-\int p_1)$ *and* $p_0 = r_0 p_2$. Then (2.3) can be written in the form $(r_0 y')' - p_0 y = 0$.

Set $r = r_0$, $q = 0$ and $p = p_0$ in (2.1). Then for any $u \in A^{\text{loc}}(a, b)$, we have that $\mathscr{R}(\phi) = -\int_a^b r_0 u'\phi' - \int_a^b p_0 u\phi$ for $\phi \in \mathscr{D}((a, b))$. Since $r_0^{-1}\eta \in A_0^2(a, b)$ whenever $\eta \in A_0^2(a, b)$, $\tilde{\mathscr{R}}(r_0^{-1}\eta)$ defines a linear functional $(1/r_0)\tilde{\mathscr{R}}$ (where $(1/r_0)\tilde{\mathscr{R}}(\eta) = \tilde{\mathscr{R}}(r_0^{-1}\eta)$) and

$$\begin{aligned}
\frac{1}{r_0}\tilde{\mathscr{R}}(\eta) = \tilde{\mathscr{R}}(r_0^{-1}\eta) &= -\int_a^b r_0 u'(r_0^{-1}\eta' + p_1 r_0^{-1}\eta) - \int_a^b p_0 u r_0^{-1}\eta \\
&= -\int_a^b u'\eta' - \int_a^b p_1 u'\eta - \int_a^b p_0 u r_0^{-1}\eta \\
&= \int_a^b u\eta'' - \int_a^b p_1 u'\eta - \int_a^b p_2 u\eta \quad \text{for } \eta \in A_0^2(a, b).
\end{aligned}$$

Hence the restriction of $(1/r_0)\tilde{\mathscr{R}}$ to $\mathscr{D}((a,b))$ is the distribution $\mathscr{D}^2u - T_{p_1u'} - T_{p_2u}$, where $T_{p_1u'}$ and T_{p_2u} are the distributions defined by p_1u' and p_2u, respectively. Now from Theorem 2.2, we know that $\mathscr{R}$ is positive if and only if $\tilde{\mathscr{R}}$ is a positive functional on $A_0^2(a,b)$. Since r_0 is positive, we have that for each nonnegative $\eta \in A_0^2(a,b)$, $r_0^{-1}\eta$ is nonnegative and in $A_0^2(a,b)$. Hence $\tilde{\mathscr{R}}$ is a positive functional on $A_0^2(a,b)$ if and only if $(1/r_0)\tilde{\mathscr{R}}$ is also. Thus $(1/r_0)\tilde{\mathscr{R}}$ is a positive functional on $A_0^2(a,b)$ if and only if the distribution $\mathscr{D}^2u - T_{p_1u'} - T_{p_2u}$ is positive in (a,b). This leads to Theorem 2.4.

THEOREM 2.4. *Let p_1 and p_2 be in $C(a,b)$ and suppose that* (2.3) *satisfies condition* (P). *Then u is a subfunction on (a,b) for* (2.3) *if and only if $u \in A^{\mathrm{loc}}(a,b)$, and $\mathscr{D}^2u \geqq T_{p_1u'} + T_{p_2u}$, i.e., $u \in A^{\mathrm{loc}}(a,b)$ and $u'' \geqq p_1(x)u' + p_2(x)u$ on (a,b) (in the sense of distributions).*

Proof. The result follows from the preceding remarks and Theorems 2.2 and 2.3.

DEFINITION 2.5. Suppose that (1) has property (P). Then a function $v : (a,b) \to R$ is called a *superfunction* (or *strict superfunction*) for (1) on (a,b) if for any interval $[x_1, x_2] \subset (a,b)$, the solution y passing through the points $(x_1, v(x_1))$ and $(x_2, v(x_2))$ satisfies

$$y(x) \leqq (\text{or } <)\, v(x) \quad \text{on } (x_1, x_2).$$

THEOREM 2.6. *Let p_1 and p_2 be in $C(a,b)$ and suppose that* (2.3) *has property* (P). *Then v is a superfunction for* (2.3) *on (a,b) if and only if*

$$v \in A^{\mathrm{loc}}(a,b) \quad \text{and} \quad \mathscr{D}^2v \leqq T_{p_1v'} + T_{p_2v}.$$

Proof. Note that v is a superfunction for (2.3) if and only if $-v$ is a subfunction for $y'' = p_1(x)y' + p_2(x)y$. The result now follows from Theorem 2.4.

THEOREM 2.7. *Let p_1 and p_2 be in $C(a,b)$ and suppose that* (2.3) *has property* (P). *Then*

(a) *u is a strict subfunction for* (2.3) *on (a,b) if and only if $u \in A^{\mathrm{loc}}(a,b)$, and*

$$\mathscr{D}^2u - T_{p_1u'} - T_{p_2u} > 0;$$

(b) *v is a strict superfunction for* (2.3) *on (a,b) if and only if $v \in A^{\mathrm{loc}}(a,b)$ and*

$$-\mathscr{D}^2v + T_{p_1v'} + T_{p_2v} > 0.$$

Proof. From the observation in the proof of Theorem 2.6, it suffices to prove part (a) only.

Suppose that $u \in A^{\mathrm{loc}}(a,b)$, and suppose that the distributional inequality in (a) holds. Then from Theorem 2.4, we have that u is a subfunction for (2.3) on (a,b). Let S be the distribution $\mathscr{D}^2u - T_{p_1u'} - T_{p_2u}$ and suppose that u were not a strict subfunction on (a,b). Then u would be a solution to (2.3) on some subinterval $(x_1, x_2) \subset (a,b)$ by the same proof as given later in Lemma 4.6 but using property (P) in place of Lemma 4.1. But this would imply that $S(\phi) = 0$ for all nonnegative ϕ in $\mathscr{D}((a,b))$ with support in (x_1, x_2), which is impossible since $S > 0$.

On the other hand, if u is a strict subfunction on (a,b), then by Theorem 2.4 we have that $u \in A^{\mathrm{loc}}(a,b)$ and $S \geqq 0$. Since any positive distribution is determined

by a positive measure (see [17] for a proof), there exists a positive measure λ such that

$$S(\phi) = \int_a^b \phi(x)\, d\lambda(x) \quad \text{for all } \phi \text{ in } \mathscr{D}((a, b)).$$

Now suppose that S were not strictly positive; i.e., suppose that $S(\phi) = 0$ for some nonnegative $\phi \not\equiv 0$ in $\mathscr{D}((a, b))$. Then it must be that $\lambda(\{x : \phi(x) > 0\}) = 0$. Let (x_3, x_4) be any subinterval of $\{x : \phi(x) > 0\}$. Then the restriction of S to $\mathscr{D}((x_3, x_4))$ would be the zero distribution in $\mathscr{D}((x_3, x_4))$. By Theorem 2.6, we would then have that u is a superfunction on (x_3, x_4), which would contradict the fact that u is a strict subfunction on (a, b). This completes the proof.

For any function u in $A^{\text{loc}}(a, b)$, let S be the distribution in (a, b), $\mathscr{D}^2 u - T_{p_1 u'} - T_{p_2 u}$, where p_1 and p_2 are in $C(a, b)$. Take functions H_1 and H_2 such that $H_1'' = p_1 u'$ and $H_2'' = p_2 u$ on (a, b). Then integrations by parts show that for all $\phi \in \mathscr{D}((a, b))$,

$$\begin{aligned} S(\phi) &= \int_a^b u(x)\phi''(x)\, dx - \int_a^b p_1(x)u'(x)\phi(x)\, dx - \int_a^b p_2(x)u(x)\phi(x)\, dx \\ &= \int_a^b (u(x) - H_1(x) - H_2(x))\phi''(x)\, dx = \mathscr{D}^2(u - H_1 - H_2)(\phi). \end{aligned} \tag{2.4}$$

Since H_1 and H_2 are in $C(a, b)$, it follows from [17] (or from Theorem 2.4 in the case of $y'' = 0$) that $u - H_1 - H_2$ is a convex function on (a, b) if and only if $u \in A^{\text{loc}}(a, b)$ and $\mathscr{D}^2(u - H_1 - H_2) \geqq 0$. We are now prepared to state the following theorem.

THEOREM 2.8. *Let p_1 and p_2 be in $C(a, b)$ and suppose that* (2.3) *has property* (P). *Let H_1 and H_2 be such that $H_1'' = p_1 u'$ and $H_2'' = p_2 u$ where u is in $A^{\text{loc}}(a, b)$. Then the following are equivalent*:

(a) *u is a subfunction (or strict subfunction) for* (2.3) *on* (a, b);

(b) *u is a local subfunction (or local strict subfunction) for* (2.3) *on* (a, b);

(c) *$u \in A^{\text{loc}}(a, b)$, and $\mathscr{D}^2(u - H_1 - H_2) \geqq 0$ (or > 0).*

(d) *$u - H_1 - H_2$ is a convex (or strictly convex) function on* (a, b).

Moreover, if u satisfies any of the conditions (a)–(d), *then u is locally Lipschitzian on* (a, b).

Proof. The result follows from Theorems 2.4 and 2.7, the remarks preceding this theorem, and the facts that any convex function on (a, b) is locally Lipschitzian on (a, b), and a locally convex function on (a, b) is convex on (a, b).

The statement of the analogous theorem for superfunctions (or strict superfunctions) is obtained from Theorem 2.8 by replacing "subfunction" by "superfunction", "convex" by "concave" throughout the statement and by reversing the inequalities.

3. The nonhomogenous linear equation $y'' = p_1(x)y' + p_2(x)y + p_3(x)$. Let p_1, p_2 and p_3 be in $C(a, b)$ and suppose that

$$y'' = p_1(x)y' + p_2(x)y + p_3(x) \tag{3.1}$$

has property (P) on (a, b). Then the corresponding homogeneous equation (2.3) has property (P). Then u is a subfunction (or strict subfunction) for (3.1) on (a, b) if and only if for any solution y to (3.1) on (a, b), $u - y$ is a subfunction (or strict subfunction) for (2.3) on (a, b).

THEOREM 3.1. *Let p_1, p_2 and p_3 be in $C(a, b)$ and suppose that* (3.1) *has property* (P). *Let H_1, H_2 and H_3 be such that $H_1'' = p_1 u'$, $H_2'' = p_2 u$ and $H_3'' = p_3$ on (a, b) where u is a locally integrable function on (a, b). Then the following are equivalent*:

(a) *u is a subfunction (or strict subfunction) for* (3.1) *on* (a, b);

(b) *u is a local subfunction (or local strict subfunction) for* (3.1) *on* (a, b);

(c) $u \in A^{\text{loc}}(a, b)$, *and* $\mathscr{D}^2(u - H_1 - H_2 - H_3) \geqq 0$ (*or* > 0).

(d) *$u - H_1 - H_2 - H_3$ is a convex (or strictly convex) function on* (a, b).

Moreover, if u satisfies any of the conditions (a)–(d), *then u is locally Lipschitzian on* (a, b).

Proof. From the remarks preceding the theorem, it suffices to show that (a) $\Leftrightarrow$ (c). But this follows from Theorem 2.8 since $\mathscr{D}^2(u - H_1 - H_2 - H_3) = \mathscr{D}^2(u - y - \tilde{H}_1 - \tilde{H}_2)$, where $\tilde{H}_1'' = p_1(u' - y')$ and $\tilde{H}_2'' = p_2(u - y)$. The proof is now complete.

Again we do not bother to state the obvious analogous theorem for superfunctions (and strict superfunctions).

COROLLARY 3.2. *Let p_1, p_2 and p_3 be in $C(a, b)$ and suppose that* (3.1) *has property* (P). *Then y is a solution to* (3.1) *on* (a, b) *if and only if* $y \in A^{\text{loc}}(a, b)$, *and* $\mathscr{D}^2 y = T_{p_1 y'} + T_{p_2 y + p_3}$, *where $T_{p_1 y'}$ and $T_{p_2 y + p_3}$ are the distributions in (a, b) determined by $p_1 y'$ and $p_2 y + p_3$, respectively.*

Proof. Suppose that $y \in A^{\text{loc}}(a, b)$ and that $\mathscr{D}^2 y = T_{p_1 y'} + T_{p_2 y + p_3}$ holds. By Theorem 3.1, we know that y and $-y$ are subfunctions and, hence that $\pm(y - H_1 - H_2 - H_3)$ is convex on (a, b). Thus the graph of $y - H_1 - H_2 - H_3$ is linear. Since $H_1 \in C^1(a, b)$ and H_2 and H_3 are in $C^2(a, b)$, then $y \in C^1(a, b)$. But this means that $H_1 \in C^2(a, b)$ since $H_1'' = p_1 y'$. Therefore $y \in C^2(a, b)$ and satisfies

$$\int_a^b y\phi'' - \int_a^b (p_1 y' - p_2 y - p_3)\phi = \int_a^b (y'' - p_1 y' - p_2 y - p_3)\phi = 0$$

for all $\phi \in \mathscr{D}((a, b))$. This implies that y is a solution to (3.1) on (a, b).

On the other hand, if y is a solution it is now clear that

$$\mathscr{D}^2 y - T_{p_1 y'} - T_{p_2 y + p_3} = 0.$$

This completes the proof.

We shall now show that when p_1 is locally absolutely continuous, the subfunctions of $y'' = p_1(x)y' + p_2(x)y + p_3(x)$ are characterized by continuity and the satisfaction (interpreted in the distributional sense for continuous functions u) of

$$u'' \geqq p_1(x)u' + p_2(x)u + p_3(x).$$

DEFINITION 3.3. For p_1 a locally absolutely continuous function on (a, b) and $u \in C(a, b)$, the product of p_1 and the distribution $\mathscr{D}u$ is the distribution $p_1 \mathscr{D}u$ defined by

$$p_1 \mathscr{D}u(\phi) = -\int_a^b u(p_1'\phi + p_1\phi'), \qquad \phi \in \mathscr{D}((a, b)).$$

LEMMA 3.4. *If p_1 and u are locally absolutely continuous functions on (a, b), then*

$$T_{p_1u'} = p_1\mathscr{D}u.$$

Proof.

$$T_{p_1u'}(\phi) = \int_a^b p_1u'\phi = -\int_a^b u(p_1\phi)' = p_1\mathscr{D}u(\phi)$$

for all $\phi \in \mathscr{D}((a, b))$ since u and $p_1\phi$ are absolutely continuous on the support of each ϕ.

THEOREM 3.5. *Suppose that p_1 is locally absolutely continuous on (a, b) and that p_2 and p_3 are in $C(a, b)$. Then the following are equivalent provided* (3.1) *has property* (P):

(a) *u is a subfunction (or strict subfunction) for* (3.1) *on (a, b);*

(b) *u is a local subfunction (or local strict subfunction) for* (3.1) *on (a, b);*

(c) *$u \in A^{\text{loc}}(a, b)$ and $\mathscr{D}^2(u - H_1 - H_2 - H_3) \geqq 0$ (or > 0) on (a, b) where $H_1'' = p_1u'$, $H_2'' = p_2u$ and $H_3'' = p_3$;*

(d) *$u \in C(a, b)$, and $\mathscr{D}^2(u - G_1 - G_2 - G_3) \geqq 0$ (or > 0) on (a, b) where $G_1' = p_1u$, $G_2'' = -up_1' + p_2u$ and $G_3'' = p_3$;*

(e) *$u \in A^{\text{loc}}(a, b)$, and $u - H_1 - H_2 - H_3$ is convex (or strictly convex) on (a, b) where H_1, H_2 and H_3 are as in* (c);

(f) *$u \in C(a, b)$, and $u - G_1 - G_2 - G_3$ is convex (or strictly convex) on (a, b) where G_1, G_2 and G_3 are as in* (d).

Moreover, if u satisfies any of the conditions (a)–(f), *then u is locally Lipschitzian on (a, b).*

Proof. If (d) holds, then $u - G_1 - G_2 - G_3$ must be convex by the characterization of convex functions given in [17, p. 54]. From this it follows that u is locally Lipschitzian on (a, b). Since Lipschitz functions are absolutely continuous, we have that $u \in A^{\text{loc}}(a, b)$. The remainder of the proof now follows from Theorem 3.1 and Lemma 3.4.

Again we omit the obvious analogue of Theorem 3.5 for superfunctions.

Remarks. In Theorem 3.5, if $p_1 \in C^1(a, b)$, then (f) is equivalent to: u is locally integrable on (a, b) and $u - G_1 - G_2 - G_3$ is convex (or strictly convex) on (a, b).

Note that the equivalence of (d) and (e) in Theorem 3.5 does not follow from Theorem 2.4 in the special case $y'' = 0$ since u is not required to be in $A^{\text{loc}}(a, b)$.

COROLLARY 3.6. *Let p_1 be locally absolutely continuous on (a, b), and let p_2 and p_3 be in $C(a, b)$. Then y is a solution to* (3.1) *on (a, b) if and only if $y \in C(a, b)$ and $\mathscr{D}^2y = p_1\mathscr{D}y + T_{p_2y+p_3}$.*

Proof. By (d) of Theorem 3.5 we can conclude that if $y \in C(a, b)$ and satisfies (3.1) in the distributional sense, then $y - G_1 - G_2 - G_3$ is both convex and concave; i.e., its graph is linear. This implies that $y \in C^2(a, b)$ and hence is a classical solution.

On the other hand, if y is a solution, an easy calculation shows that (3.1) holds in the distributional sense. This completes the proof.

Thus any continuous function on (a, b) that satisfies (3.1) in the distributional sense, where p_1 is locally absolutely continuous and p_2 and p_3 are in $C(a, b)$, is in fact in $C^2(a, b)$ and is a classical solution.

When the p_i's ($i = 1, 2, 3$) are in $C^\infty(a, b)$, Corollaries 3.2 and 3.6 follow from § 2.3 of [4, pp. 39–43] (also see [18, p. 53]).

4. The nonlinear equation $y'' = f(x, y)$. Now we consider the case that (1) does not depend on y'. We identify f with its restriction to $(a, b) \times R$ and write

$$y'' = f(x, y). \tag{4.1}$$

We shall always assume that condition (i) holds:

(i) $f : (a, b) \times R \to R$ is continuous.

We shall sometimes assume that (4.1) has property (ii):

(ii) If y_1 and y_2 are solutions of (4.1) on $[x_1, x_2] \subset (a, b)$ with $y_1(x_1) = y_2(x_1)$ and $y_1(x_2) = y_2(x_2)$, then $y_1 = y_2$ on $[x_1, x_2]$.

It is known [3, Thm. 2, p. 1256] that if (i) holds and if u is a bounded subfunction for (4.1) on (a, b), then $u \in C(a, b)$. Moreover, if (i) and (ii) hold and if $u \in C^2(a, b)$, then $u''(x) \geqq f(x, u(x))$ on (a, b) is a necessary and sufficient condition for u to be a subfunction with respect to solutions of (4.1) on (a, b) (see [15, Corollary, p. 1011]).

LEMMA 4.1 (see [9, Thm. 2.1, p. 309] and [3, Lemmas 1, 2, 3, p. 1252] for related results). *Assume that f satisfies* (i) *and that $M > 0$ and $[c, d] \subset (a, b)$ are given. Let q be the maximum of $|f(x, y)|$ on the compact set $\{(x, y) : c \leqq x \leqq d, |y| \leqq 2M\}$. Then, if $\delta = (8M/q)^{1/2}$, any boundary value problem*

$$y'' = f(x, y), \quad y(x_1) = y_1, \quad y(x_2) = y_2$$

with $[x_1, x_2] \subset [c, d]$, $x_2 - x_1 \leqq \delta$, $|y_1| \leqq M$ and $|y_2| \leqq M$ has a solution $y \in C^2[x_1, x_2]$. Furthermore, given $\varepsilon > 0$, the solution y described above will satisfy $|y(x) - \omega(x)| \leqq \varepsilon$ on $[x_1, x_2]$ provided $x_2 - x_1 \leqq (8\varepsilon/q)^{1/2}$ where ω is the linear function with $\omega(x_1) = y_1$ and $\omega(x_2) = y_2$.

Proof. The set

$$B[x_1, x_2] = \{z : z \in C[x_1, x_2], |z(x)| \leqq 2M \text{ for } x_1 \leqq x \leqq x_2\}$$

is a closed convex subset of the Banach space $C[x_1, x_2]$. The mapping

$$T : C[x_1, x_2] \to C[x_1, x_2]$$

defined by

$$(Tz)(x) = \int_{x_1}^{x_2} G(x, t) f(t, z(t))\, dt + \omega(x),$$

where $G(x, t)$ is the Green's function for the boundary value problem $y'' = 0$, $y(x_1) = y(x_2) = 0$, is completely continuous. For a $z \in B[x_1, x_2]$ we have

$$|(Tz)(x)| \leqq \tfrac{1}{8} q(x_2 - x_1)^2 + M$$

on $[x_1, x_2]$. Thus $x_2 - x_1 \leqq \delta$ implies T maps $B[x_1, x_2]$ into itself. It then follows from [6, Corollary 0.1, p. 405] that T has a fixed point in $B[x_1, x_2]$. The fixed point is a solution of the stated boundary value problem. If y is a solution of the boundary value problem with $y \in B[x_1, x_2]$, then

$$|y(x) - \omega(x)| \leqq \tfrac{1}{8} q(x_2 - x_1)^2$$

on $[x_1, x_2]$ and the last assertion of the theorem follows.

DEFINITION 4.2. Let u and v be locally integrable, locally bounded functions on (a, b) and assume that (i) holds. Then we let U and V be the distributions in (a, b) defined by

$$U(\phi) = \int_a^b f(x, u(x))\phi(x)\,dx,$$

$$V(\phi) = \int_a^b f(x, v(x))\phi(x)\,dx, \quad \text{where } \phi \in \mathscr{D}((a, b)).$$

Also, just as in § 1, we define the distributions $\mathscr{D}^2 u$ and $\mathscr{D}^2 v$ by

$$\mathscr{D}^2 u(\phi) = \int_a^b u(x)\phi''(x)\,dx \quad \text{and} \quad \mathscr{D}^2 v(\phi) = \int_a^b v(x)\phi''(x)\,dx.$$

LEMMA 4.3. *Let $[c, d]$ be any closed subinterval of (a, b) and let u be in $C(a, b)$. Suppose that* (i) *holds and suppose that the restriction of $\mathscr{D}^2 u - U$ to $\mathscr{D}((c, d))$ is a positive distribution in (c, d). Then there exists $\delta > 0$ such that if $[x_1, x_2] \subset [c, d]$ with $x_2 - x_1 \leqq \delta$, the boundary value problem*

$$y'' = f(x, y), \quad y(x_1) = u(x_1), \quad y(x_2) = u(x_2)$$

has a solution $y \in C^2[x_1, x_2]$ and

$$y(x) \geqq u(x) \quad \textit{for } x_1 \leqq x \leqq x_2.$$

Proof. Define F, $F : (a, b) \times R \to R$ by

$$F(x, y) = \begin{cases} f(x, y) & \text{for } y \geqq u(x), \\ f(x, u(x)) & \text{for } y < u(x). \end{cases}$$

By Lemma 4.1, there is a $\delta > 0$ such that if $x_2 - x_1 \leqq \delta$, then the boundary value problem

$$y'' = F(x, y), \quad y(x_1) = u(x_1), \quad y(x_2) = u(x_2)$$

has a solution $y \in C^2[x_1, x_2]$. We need only show that $y(x) \geqq u(x)$ for $x_1 \leqq x \leqq x_2$ to complete the proof.

If $y(x_0) < u(x_0)$ for some x_0 in (x_1, x_2), then we can find an interval $[x_3, x_4] \subset [x_1, x_2]$ such that $y(x) < u(x)$ for $x_3 < x < x_4$, $y(x_3) = u(x_3)$ and $y(x_4) = u(x_4)$. Let W be the distribution in (x_3, x_4) defined by

$$W(\phi) = \int_{x_3}^{x_4} (u(x) - y(x))\phi''(x)\,dx \quad \text{for } \phi \in \mathscr{D}((x_3, x_4)).$$

Since $y''(x) = f(x, u(x))$ for $x_3 < x < x_4$, we obtain after two integrations by parts that

$$\int_{x_3}^{x_4} y(x)\phi''(x)\,dx = U(\phi) \quad \text{for } \phi \in \mathscr{D}((x_3, x_4)).$$

Thus we have that $W = \mathscr{D}^2(u - y) = \mathscr{D}^2 u - U$ in (x_3, x_4). Now $W \geqq 0$ since $\mathscr{D}^2 u - U$ is positive in $(c, d) \supset (x_3, x_4)$. Then by the second distributional derivative, test, $u - y$ is convex on (x_3, x_4) since it is continuous. But $u - y$ being convex

on (x_3, x_4) contradicts the facts that $u - y > 0$ on (x_3, x_4) and $u(x_3) - y(x_3) = u(x_4) - y(x_4) = 0$. The proof is complete.

LEMMA 4.4. *Let* (i) *and* (ii) *hold, and let* $u \in C(a, b)$ *be a subfunction with respect to solutions of* (4.1) *on* (a, b). *Then for any* $[c, d] \subset (a, b)$ *and for any* μ, $0 < \mu < \delta$ *(where* δ *comes from Lemma* 4.1 *with* $M = \max\{|u(x)| : c \leqq x \leqq d\}$, $y_1 = u(x_1)$ *and* $y_2 = u(x_2)$*) there exists a partition of* $[c, d]$, $c = x_0 < x_1 < \cdots < x_n = d$ *with* $x_i - x_{i-1} \leqq \mu$ $(i = 1, 2, \cdots, n)$, *and a solution* y_i *of* (4.1) *on* $[x_{i-1}, x_i]$ *which satisfies* $y_i(x_{i-1}) = u(x_{i-1})$, $y_i(x_i) = u(x_i)$ *and* $y_i(x) \geqq u(x)$ *for* $x_{i-1} \leqq x \leqq x_i$ *provided* $i = 1, 2, \cdots, n$. *Moreover,* y_i *satisfies* $y_i(x) \leqq u(x)$ *on* $[x_i, c_i)$ *for some* $c_i > x_i$, $i = 1, 2, \cdots, n - 1$ (if $n = 1$ *this condition is vacuously satisfied*).

Proof. The proof is by induction on the smallest positive integer m such that $c + m\mu \geqq d$. If $c + \mu \geqq d$, then let $x_0 = c$, $x_1 = d$ and y_1 be the solution of (4.1) (which exists by Lemma 4.1) satisfying $y_1(x) \geqq u(x)$ for $c \leqq x \leqq d$, where equality holds at $x = c$ and at $x = d$. Thus the lemma is true for $m = 1$. Now assume that the lemma is correct for some fixed value of m, $m \geqq 1$. We shall prove it for $m + 1$.

Let $c + m\mu < d$ and $c + (m + 1)\mu \geqq d$ where μ, $0 < \mu < \delta$, is fixed and δ is determined by Lemma 4.1 applied to $[c, d]$ with $M = \max\{|u(x)| : c \leqq x \leqq d\}$. Then let $\delta_1, \delta_1 \geqq \delta$, be determined by Lemma 4.1 applied to the interval $[c, d - \mu]$. Thus $0 < \mu < \delta \leqq \delta_1$, so by the induction hypotheses applied to the interval $[c, d - \mu]$, there exists a partition of $[c, d - \mu]$, $c = t_0 < t_1 < \cdots < t_n = d - \mu$ with $t_i - t_{i-1} \leqq \mu$ for $i = 1, 2, \cdots, n$ and solutions z_i of (4.1) on $[t_{i-1}, t_i]$ which satisfy $z_i(t_{i-1}) = u(t_{i-1})$, $z_i(t_i) = u(t_i)$ and $z_i(x) \geqq u(x)$ for $t_{i-1} \leqq x \leqq t_i$ provided $i = 1, 2, \cdots, n$. Moreover, z_i satisfies $z_i(x) \leqq u(x)$ on $[t_i, d_i)$ for some $d_i > t_i$ for $i = 1, 2, \ldots, n - 1$ (vacuously if $n = 1$). Let $x_i = t_i$, $c_i = d_i$ and $y_i = z_i$ for $i = 0, 1, \cdots, n - 1$.

We need to consider two possibilities. If there is a point b_0 satisfying $t_{n-1} < b_0 < t_n$, $t_n - b_0 < \mu/16$ and $z_n(b_0) > u(b_0)$, then let y_n be any solution on a maximal interval of existence $(\omega_1 -, \omega_1 +)$ of the boundary value problem

$$y'' = f(x, y), \quad y(t_{n-1}) = u(t_{n-1}), \quad y(b_0) = u(b_0)$$

which exists on $[t_{n-1}, b_0] \subset (\omega_1 -, \omega_1 +)$ by Lemma 4.1 and which satisfies $y_n(x) \geqq u(x)$ for $t_{n-1} \leqq x \leqq b_0$ since u is a subfunction. Let a_0, $b_0 \leqq a_0 < t_n$, be the largest value of x satisfying $b_0 \leqq x < t_n$ and $y_n(x) = u(x)$. Then let $x_n = a_0$ and c_n be any point in $(a_0, t_n) \cap (\omega_1 -, \omega_1 +)$. If no such number b_0 exists, then let $x_n = t_n - \mu/32$, $c_n = t_n - \mu/64$ and $y_n = z_n$.

Let w_1 be the solution of $y'' = f(x, y)$, $y(x_n) = u(x_n)$, $y(x_n + \mu) = u(x_n + \mu)$. If there is a point b_1 satisfying $x_n < b_1 < x_n + \mu$, $x_n + \mu - b_1 < \mu/16$ and $w_1(b_1) > u(b_1)$, then let y_{n+1} be any solution on a maximal interval of existence $(\omega_2 -, \omega_2 +)$ of the boundary value problem

$$y'' = f(x, y), \quad y(x_n) = u(x_n), \quad y(b_1) = u(b_1)$$

which exists on $[x_n, b_1] \subset (\omega_2 -, \omega_2 +)$ by Lemma 4.1 and which satisfies $y_{n+1}(x) \geqq u(x)$ for $x_n \leqq x \leqq b_1$ since u is a subfunction. Let $a_1, b_1 \leqq a_1 < x_n + \mu$, be the largest value of x satisfying $b_1 \leqq x < x_n + \mu$ and $y_{n+1}(x) = u(x)$. Then let $x_{n+1} = a_1$ and c_{n+1} be any point in $(a_1, x_n + \mu) \cap (\omega_2 -, \omega_2 +)$. If no such point b_1 exists, then let $x_{n+1} = x_n + \mu - \mu/32$, $c_{n+1} = x_n + \mu - \mu/64$ and $y_{n+1} = w_1$.

Now let y_{n+2} be the solution of the boundary value problem

$$y'' = f(x, y), \quad y(x_{n+1}) = u(x_{n+1}), \quad y(d) = u(d)$$

which exists by Lemma 4.1 and which satisfies $y_{n+2}(x) \geqq u(x)$ since u is a subfunction. This completes the proof.

THEOREM 4.5. *Suppose that* (i) *and* (ii) *hold. Then u is a locally bounded subfunction for* (4.1) *on* (a, b) *if and only if* $u \in C(a, b)$, *and* $\mathscr{D}^2 u - U$ *is a positive distribution in* (a, b). *Moreover, v is a locally bounded superfunction for* (4.1) *on* (a, b) *if and only if* $v \in C(a, b)$ *and* $V - \mathscr{D}^2 v$ *is a positive distribution in* (a, b).

Proof. We only prove the first equivalence since the proof of the second is similar.

Let $u \in C(a, b)$ be such that $\mathscr{D}^2 u \geqq U$. Now suppose that u is not a subfunction on (a, b). Then there exists an interval $[c, d] \subset (a, b)$ and a solution z of (4.1) with $z(c) = u(c)$, $z(d) = u(d)$ and $z(x) < u(x)$ for $c < x < d$. For each positive integer n, we let P(n) be the proposition that there exists an interval $[a_n, b_n] \subset [c, d]$ with $0 < b_n - a_n \leqq d - c - (n-1)\delta/2$ (where δ comes from Lemma 4.3) and a solution z_n of (4.1) on $[a_n, b_n]$ such that $z_n(a_n) = u(a_n)$, $z_n(b_n) = u(b_n)$ and $z_n(x) < u(x)$ for $a_n < x < b_n$. We will show that under our assumption that u is not a subfunction on (a, b) relative to solutions of (4.1), it follows that P(n) holds for each positive integer n. This gives a contradiction since it is not possible to have $0 < d - c - (n-1)\delta/2$ for every positive integer n.

The fact that P(1) is true follows by letting $a_1 = c$, $b_1 = d$ and $z_1 = z$. We assume that P(k) is true and will show that this implies P($k+1$) is true. If $b_k - a_k \leqq \delta$, then we get a contradiction from Lemma 4.3, so we have $b_k - a_k > \delta$. Let y_1 be any solution defined on a maximal interval of existence, (ω_1-, ω_1+), to the boundary value problem

$$y'' = f(x, y), \quad y(a_k) = u(a_k), \quad y(a_k + \delta) = u(a_k + \delta)$$

which exists on $[a_k, a_k + \delta] \subset (\omega_1-, \omega_1+)$ by Lemma 4.3. If P($k+1$) is not true, then $y_1(x) \geqq u(x)$ for $a_k \leqq x < \min\{b_k, \omega_1+\}$. Also, if P($k+1$) is not true, then $u(x) = y_1(x)$ for $a_k + \delta/2 \leqq x \leqq a_k + \delta$, for otherwise the solution y, given by Lemma 4.3 of the boundary value problem

$$y'' = f(x, y), \quad y(a_k) = u(a_k), \quad y(b_0) = u(b_0)$$

(where b_0 is chosen so that $a_k + \delta/2 \leqq b_0 \leqq a_k + \delta$ and $u(b_0) < y_1(b_0)$) when extended to a maximal interval of existence to the right would contradict P($k+1$) not being true.

If $b_k - a_k - \delta \leqq \delta/2$, we let y_2 be the solution given by Lemma 4.3 to the boundary value problem

$$y'' = f(x, y), \quad y(a_k + \delta/2) = u(a_k + \delta/2), \quad y(b_k) = u(b_k).$$

Then $y_2(x) \geqq u(x)$ for $a_k + \delta/2 \leqq x \leqq b_k$ and $y_2(x) \leqq y_1(x)$ for $a_k + \delta/2 \leqq x < \min\{\omega_1+, b_k\}$. This implies that $u(x) = y_1(x) = y_2(x)$ for $a_k + \delta/2 \leqq x \leqq a_k + \delta$. But now (ii) is violated since the function l defined by

$$l(x) = \begin{cases} y_1(x) & \text{for } a_k \leqq x \leqq a_k + \delta, \\ y_2(x) & \text{for } a_k + \delta < x \leqq b_k, \end{cases}$$

is a solution of (4.1) and is a solution to the same boundary value problem that z_k is, but l and z_k are not identical on $[a_k, b_k]$. We conclude that $b_k - a_k - \delta > \delta/2$. Now let y_2 be any solution on a maximal interval of existence (ω_2-, ω_2+) of the boundary value problem

$$y'' = f(x, y), \quad y(a_k + \delta/2) = u(a_k + \delta/2), \quad y\left(a_k + \frac{3\delta}{2}\right) = u\left(a_k + \frac{3\delta}{2}\right)$$

which exists on $[a_k + \delta/2, a_k + 3\delta/2] \subset (\omega_2-, \omega_2+)$ by Lemma 4.3. Note that $y_2(x) \geqq u(x)$ for $a_k + \delta/2 \leqq x < \min\{\omega_2+, b_k\}$ for otherwise the function m defined by

$$m(x) = \begin{cases} y_1(x) & \text{for } a_k \leqq x \leqq a_k + \delta, \\ y_2(x) & \text{for } a_k + \delta < x < \omega_2+, \end{cases}$$

is a solution of (4.1) that would imply P($k + 1$) was true. Continuing in this way we construct $y_3, y_4, \cdots$ until we have worked our way across the interval $[a_k, b_k]$ which implies P(k) is not true. Thus P($k + 1$) is true. So by the induction principle we obtain a contradiction. This proves that u is a subfunction on (a, b) whenever $u \in C(a, b)$ and $\mathscr{D}^2 u \geqq U$.

Suppose on the other hand, that u is a locally bounded subfunction for (4.1) on (a, b). It follows from [3, Thm. 2, p. 1256] that $u \in C(a, b)$.

We shall now show that $\mathscr{D}^2 u \geqq U$. Let $[c, d]$ be any compact subinterval of (a, b). Take any nonnegative $\phi \in \mathscr{D}((a, b))$ with support contained in $[c, d]$.

For each n such that $1/n < \delta$ (where δ comes from Lemma 4.1 with

$$M = \max\{|u(x)| : c \leqq x \leqq d\})$$

let $\mu = 1/n$ and apply Lemma 4.4 on the interval $[c, d]$. Then define $u_n : [c, d] \to R$ by

$$u_n(x) = y_i(x) \quad \text{for } x_{i-1} \leqq x \leqq x_i, \qquad i = 1, 2, \cdots, k.$$

It follows from the last sentence in the statement of Lemma 4.1 and from the continuity of u that $u_n(x)$ converges uniformly to $u(x)$ on the interval $[c, d]$. If we can show that

$$\int_a^b u_n(t)\phi''(t)\,dt \geqq \int_a^b f(t, u_n(t))\phi(t)\,dt,$$

then by taking the limit of each side as $n \to +\infty$, we would have that

$$\int_a^b u(t)\phi''(t)\,dt \geqq \int_a^b f(t, u(t))\phi(t)\,dt,$$

and since $[c, d]$ is arbitrary, we could conclude that $\mathscr{D}^2 u \geqq U$.

Now

$$\int_a^b u_n(t)\phi''(t)\,dt = \sum_{i=1}^{k}\left(\int_{x_{i-1}}^{xi} y_i(t)\phi''(t)\,dt\right)$$

$$= \sum_{i=1}^{k}(y_i(t)\phi'(t) - y_i'(t)\phi(t))\Big|_{x_{i-1}}^{x_i} + \sum_{i=1}^{k}\left(\int_{x_{i-1}}^{x_i} y_i''(t)\phi(t)\,dt\right)$$

$$= \sum_{i=1}^{k}(y_i(x_i)\phi'(x_i) - y_i'(x_i)\phi(x_i)) - \sum_{i=1}^{k}(y_i(x_{i-1})\phi'(x_{i-1}) - y_i'(x_{i-1})\phi(x_{i-1}))$$

$$+ \sum_{i=1}^{k}\int_{x_{i-1}}^{x_i} y_i''(t)\phi(t)\,dt$$

$$= \sum_{i=1}^{k-1}(y_i(x_i)\phi'(x_i) - y_i'(x_i)\phi(x_i)) - \sum_{i=2}^{k}(y_i(x_{i-1})\phi'(x_{i-1}) - y_i'(x_{i-1})\phi(x_{i-1}))$$

$$+ \sum_{i=1}^{k}\int_{x_{i-1}}^{x_i} f(t, y_i(t))\phi(t)\,dt$$

$$= \sum_{j=2}^{k}(y_{j-1}(x_{j-1})\phi'(x_{j-1}) - y_{j-1}'(x_{j-1})\phi(x_{j-1}))$$

$$- \sum_{j=2}^{k}(y_j(x_{j-1})\phi'(x_{j-1}) - y_j'(x_{j-1})\phi(x_{j-1}))$$

$$+ \int_a^b f(t, u_n(t))\phi(t)\,dt$$

$$= \sum_{j=2}^{k}(y_{j-1}(x_{j-1}) - y_j(x_{j-1}))\phi'(x_{j-1}) + \sum_{j=2}^{k}(y_j'(x_{j-1}) - y_{j-1}'(x_{j-1}))\phi(x_{j-1})$$

$$+ \int_a^b f(t, u_n(t))\phi(t)\,dt$$

$$= \sum_{j=2}^{k}(y_j'(x_{j-1}) - y_{j-1}'(x_{j-1}))\phi(x_{j-1}) + \int_a^b f(t, u_n(t))\phi(t)\,dt$$

$$\geqq \int_a^b f(t, u_n(t))\phi(t)\,dt.$$

The last inequality holds because it follows from Lemma 4.4 that $y_j'(x_{j-1}) \geqq y_{j-1}'(x_{j-1})$ for $j = 2, 3, \cdots, k$ since $y_j(x) \geqq u(x) \geqq y_{j-1}(x)$ holds for $x_{j-1} \leqq x \leqq c_{j-1}$. The proof is now complete.

LEMMA 4.6. *Assume that* (i) *and* (ii) *hold and that solutions to initial value problems for* (4.1) *are unique. If* $u \in C(a, b)$ *is a subfunction for* (4.1) *on* (a, b), *but not a strict subfunction, then* u *is a solution to* (4.1) *on some subinterval of* (a, b).

Proof. Since u is a continuous subfunction, but is not a strict subfunction, there exists a solution y on some interval $(c, d) \subset (a, b)$ such that $y(c) = u(c)$, $y(d) = u(d)$, $y(x) \geqq u(x)$ on $[c, d]$ and $y(x_0) = u(x_0)$ for some $x_0 \in (c, d)$. Suppose that u were not a solution on any subinterval of (a, b). Then by Lemma 4.1, there

exists an interval (x_1, x_2) containing x_0 and a solution y_0 on $[x_1, x_2] \subset (c, d)$ such that $u(x) \leqq y_0(x)$ for $x \in [x_1, x_2]$, $y_0(x_1) = y(x_1) > u(x_1)$ and $y_0(x_2) = u(x_2) < y(x_2)$. Since $y_0(x) \leqq y(x)$ on $[x_1, x_2]$ by (ii), it follows that $y_0(x_0) = y(x_0)$, and hence $y_0(x) = y(x)$ for $x_1 \leqq x \leqq x_0$. Since $y_0(x_2) < y(x_2)$, y_0 and y are different solutions of the same initial value problem which is a contradiction. Therefore u must be a solution to (4.1) on some subinterval of (a, b) containing x_0.

THEOREM 4.7. *If* (i) *and* (ii) *hold, and if u is a locally bounded strict subfunction on (a, b) for* (4.1), *then $u \in C(a, b)$ and $\mathscr{D}^2 u > U$. Furthermore, if* (i) *and* (ii) *hold, and if solutions to initial value problems of* (4.1) *are unique, then u is a strict subfunction on (a, b) whenever $u \in C(a, b)$ and $\mathscr{D}^2 u > U$.*

Proof. We can essentially apply the proof of Theorem 2.7, using Theorem 4.5 and Lemma 4.6. The details will not be repeated.

THEOREM 4.8. *Let u be locally bounded and locally integrable on (a, b). Assume that* (i) *and* (ii) *hold (and in the strict case, assume also that solutions to initial value problems for* (4.1) *are unique), and let H be such that $H''(x) = f(x, u(x))$ for all $x \in (a, b)$. Then the following are equivalent:*

(a) *u is a subfunction (or strict subfunction) for* (4.1) *on (a, b);*

(b) *u is a local subfunction (or local strict subfunction) for* (4.1) *on (a, b);*

(c) *$u \in C(a, b)$ and $\mathscr{D}^2(u - H) = \mathscr{D}^2 u - U \geqq 0$ (or > 0);*

(d) *$u - H$ is a convex (or strictly convex) function on (a, b).*

Moreover, if u is locally bounded on (a, b) and satisfies any of the conditions (a)–(d), *then u is locally Lipschitzian.*

Proof. We can apply the proof of Theorem 3.1, using H instead of $H_1 + H_2 + H_3$ and Theorems 4.5 and 4.7. The details will not be repeated.

Remark. Note that u is a subfunction (or strict subfunction) for (4.1) if and only if $u - H$ is a subfunction (or strict subfunction) for

$$y'' = f(x, y + H) - f(x, u(x)).$$

We will not give the analogues of Theorems 4.7 and 4.8 for superfunctions.

COROLLARY 4.9. *Suppose that* (i) *and* (ii) *hold. Let $u \in C(a, b)$ be a distributional solution of* (4.1) *on (a, b) (i.e., $\mathscr{D}^2 u - U$ is the zero distribution in (a, b)). Then $u \in C^2(a, b)$ and u is a classical solution on (a, b).*

Proof. By Theorem 4.8 we obtain that the graph of $u - H$ is linear. The result now follows.

5. Boundary value problems. In this section, we give necessary and sufficient conditions for the boundary value problem

$$y'' = f(x, y), \quad y(x_1) = y_1, \quad y(x_2) = y_2 \tag{5.1}$$

to have a solution. This result is a useful improvement of the result in [16, Cor. 3.1] making it easier to get close bounds on the solution as the later example will show.

THEOREM 5.1. *Assume that* (i) *holds and that $[x_1, x_2] \subset (a, b)$. Then* (5.1) *has a solution if and only if there exist locally bounded, locally integrable functions u and v on (a, b) such that $\mathscr{D}^2 u \geqq U$, $\mathscr{D}^2 v \leqq V$, $u(x) \leqq v(x)$ a.e.,*

$$\lim_{x \to x_1^+} \inf v(x) \geqq y_1 \geqq \lim_{x \to x_1^+} \sup u(x), \tag{5.2}$$

and

$$\lim_{x \to x_2^-} \inf v(x) \geqq y_2 \geqq \lim_{x \to x_2^-} \sup u(x). \tag{5.3}$$

Proof. If (5.1) has a solution y, then choose $u = v = y$. On the other hand, if u and v as described in the theorem exist, then let H_u and H_v be such that $H_u''(x) = f(x, u(x))$ for all $x \in (a, b)$, and $H_v''(x) = f(x, v(x))$ for all $x \in (a, b)$. Since $\mathscr{D}^2(u - H_u) = \mathscr{D}^2 u - U \geqq 0$, and $\mathscr{D}^2(v - H_v) = \mathscr{D}^2 v - V \leqq 0$, the distributional characterization of convex functions (see [17]) gives that $u - H_u$ and $-v + H_v$ each equals a.e. a convex function on (a, b). Since convex functions are continuous, it follows from the continuity of H_u and H_v that there exist functions u_c and v_c in $C(a, b)$ such that $u = u_c$ a.e. and $v = v_c$ a.e. Thus we have that $u_c \leqq v_c$, in particular, (5.2) and (5.3) hold. We also have that $u_c - H_{u_c}$ is convex on (a, b) and $v_c - H_{v_c}$ is concave on (a, b). Define $F : [x_1, x_2] \times R \to R$ by

$$F(x, y) = \begin{cases} f(x, v_c(x)) & \text{for } y(x) \geqq v_c, \\ f(x, y) & \text{for } u_c(x) < y(x) < v_c(x), \\ f(x, u_c(x)) & \text{for } y(x) \leqq u_c(x), \end{cases}$$

and consider the boundary value problem

$$y'' = F(x, y), \quad y(x_1) = y_1, \quad y(x_2) = y_2.$$

It follows from [10, Lemma 2.2, p. 250] that this problem has a solution $y \in C^2[x_1, x_2]$. If we can shown that $u_c(x) \leqq y(x) \leqq v_c(x)$, we are done, for then y is also a solution to (5.1). The proof that $y(x) \geqq u_c(x)$ is essentially the same as the second paragraph of the proof of Lemma 4.3. The proof that $y(x) \leqq v_c(x)$ is similar so is omitted.

The next result and its analogue provide us with a practical method of constructing continuous functions u and v for use in Theorem 5.1.

THEOREM 5.2. *Assume that* (i) *holds and that* $[x_1, x_2] \subset (a, b)$. *Suppose there exist a partition of* $[x_1, x_2]$, $x_1 = \alpha_1 < \alpha_2 < \cdots < \alpha_k = x_2$, *and a function* $u \in C(a, b)$ *having left-hand and right-hand derivatives at each* α_i $(i = 1, \cdots, k)$ *which satisfy for each i,*

$$u'_-(\alpha_i) \leqq u'_+(\alpha_i). \tag{5.4}$$

Suppose also that the restrictions of $\mathscr{D}^2 u - U$ *to each* (α_i, α_{i+1}) *for* $i = 1, \cdots, k - 1$ *are positive. Then the restriction of* $\mathscr{D}^2 u - U$ *to* (x_1, x_2) *is positive.*

Proof. Let H be such that $H''(x) = f(x, u(x))$ for all $x \in (x_1, x_2)$. Then the restrictions of $u - H$ to (α_i, α_{i+1}), $i = 1, \cdots, k - 1$, are convex. This and the fact that (5.4) holds imply by [7, (18.43), p. 300] that $u' - H'$ exists a.e. in (x_1, x_2) and is nondecreasing where it is defined. From the same result in [7], we obtain that $u - H$ is convex in (x_1, x_2). This means, as we have seen before, that $\mathscr{D}^2(u - H) = \mathscr{D}^2 u - U$ restricted to (x_1, x_2) is positive, which completes the proof.

Remark. It is now clear that if $u \in C(a, b)$ is such that $v'_-(\alpha_i) \geqq v'_+(\alpha_i)$, $i = 1, \cdots, k$, (α_i as in Theorem 5.2), and if $\mathscr{D}^2 v \leqq V$ in each (α_i, α_{i+1}), then $\mathscr{D}^2 v \leqq V$ in (x_1, x_2).

To illustrate the use of the results in this section to establish the existence of solutions to two point boundary value problems and to obtain closer bounds on the solution than is easily possible by known methods we investigate one example.

Consider the boundary value problem

$$y'' = -4xy + y^3,$$
$$y(-1) = -2, \qquad y(2) = 1.$$

Define

$$u_1(x) = -3 \qquad \text{for } -1 \leqq x \leqq 2,$$
$$u_2(x) = -2\sqrt{2} \qquad \text{for } -1 \leqq x \leqq 2,$$
$$u_3(x) = \tfrac{2}{3}(1 - \sqrt{2})x + \tfrac{2}{3}(1 - \sqrt{2}) - 2 \qquad \text{for } -1 \leqq x \leqq 2,$$
$$v_1(x) = 3 \qquad \text{for } -1 \leqq x \leqq 2,$$
$$v_2(x) = 2\sqrt{2} \qquad \text{for } -1 \leqq x \leqq 2,$$
$$v_3(x) = \begin{cases} 2\sqrt{2}x + 2\sqrt{2} & \text{for } -1 \leqq x \leqq 0, \\ 2\sqrt{2} & \text{for } 0 < x \leqq 2, \end{cases}$$

and

$$v_4(x) = \begin{cases} x + 1 & \text{for } -1 \leqq x \leqq 1, \\ 2x^{1/2} & \text{for } 1 < x \leqq 2. \end{cases}$$

Then $u_i(x) \leqq u_{i+1}(x) \leqq v_{i+2}(x) \leqq v_{i+1}(x) \leqq v_i(x)$ for all $x \in [-1, 2]$, $i = 1, 2$, and we also have that $u_3(-1) \leqq -2 \leqq v_4(-1)$ and $u_3(2) \leqq 1 \leqq v_4(2)$. It is easy to check that for each u_i, $\mathscr{D}^2 u_i \geqq U_i$, where U_i is the distribution in the interval $(-1, 2)$ defined by $-4xu_i(x) + (u_i(x))^3$, and for each v_i, $\mathscr{D}^2 v_i \leqq V_i$, where V_i is the distribution in the interval $(-1, 2)$ defined by $-4xv_i(x) + (v_i(x))^3$. It now follows from Theorem 5.1 using any u_i and any v_i that the boundary value problem has a solution y satisfying

$$u_i(x) \leqq y(x) \leqq v_i(x) \quad \text{for } x \in [-1, 2].$$

Clearly the best choice of the given functions would be u_3 and v_4, which gives that

$$u_3(x) \leqq y(x) \leqq v_4(x) \quad \text{for } x \in [-1, 2].$$

REFERENCES

[1] E. F. Beckenbach, *Generalized convex functions*, Bull. Amer. Math. Soc., 43 (1937), pp. 363–371.

[2] F. F. Bonsall, *The characterization of generalized convex functions*, Quart. J. Math. Oxford Ser., 1 (1950), pp. 100–111.

[3] L. D. Fountain and L. K. Jackson, *A generalized solution of the boundary value problem for* $y'' = f(x, y, y')$, Pacific J. Math., 12 (1962), pp. 1251–1272.

[4] I. M. Gel'fand and G. E. Shilov, *Generalized Functions*, vol. 1, Academic Press, New York, 1964.

[5] J. W. Green, *Support, convergence and differentiability properties of generalized convex functions*, Proc. Amer. Math. Soc., 4 (1953), pp. 391–396.

[6] P. Hartman, *Ordinary Differential Equations*, John Wiley, New York, 1964.

[7] E. HEWITT AND K. STROMBERG, *Real and Abstract Analysis*, Springer-Verlag, New York, 1965.
[8] L. HÖRMANDER, *Linear Partial Differential Operators*, Springer-Verlag, New York, 1969.
[9] L. K. JACKSON, *Subfunctions and second order differential inequalities*, Advances in Math., 2 (1968), pp. 307–363.
[10] L. K. JACKSON AND K. W. SCHRADER, *Comparison theorems for nonlinear differential equations*, J. Differential Equations, 3 (1967), pp. 248–255.
[11] ———, *On second order differential inequalities*, Proc. Amer. Math. Soc., 5 (1966), pp. 1023–1027.
[12] M. M. PEIXOTO, *Generalized convex functions and second order differential inequalities*, Bull. Amer. Math. Soc., 55 (1949), pp. 563–572.
[13] ———, *On the existence of derivative of generalized convex functions*, Summa Brasil. Math., 2 (1948), pp. 35–42.
[14] W. T. REID, *Variational aspects of generalized convex functions*, Pacific J. Math., 9 (1959), pp. 571–581.
[15] K. W. SCHRADER, *A note on second order differential inequalities*, Proc. Amer. Math. Soc., 19 (1968), pp. 1007–1012.
[16] ———, *Boundary-value problems for second-order ordinary differential equations*, J. Differential Equations, 3 (1967), pp. 403–413.
[17] L. SCHWARTZ, *Théorie des Distributions*, Hermann, Paris, 1966.
[18] I. STAKGOLD, *Boundary Value Problems of Mathematical Physics*, vol. 1, Macmillan, New York, 1967.

SIAM J. MATH. ANAL.
Vol. 8, No. 1, February 1977

EXISTENCE THEOREMS AND A SOLUTION ALGORITHM FOR PIECEWISE-LINEAR RESISTOR NETWORKS*

T. OHTSUKI†, T. FUJISAWA‡ AND S. KUMAGAI‡

Abstract. This paper deals with nonlinear networks which can be characterized by the equation $\mathbf{f}(\mathbf{x})=\mathbf{y}$, where $\mathbf{f}$ is a continuous piecewise-linear mapping from R^n into itself. The main theorem asserts that the existence of solutions $\mathbf{x}\in R^n$ of $\mathbf{f}(\mathbf{x})=\mathbf{y}$ for an arbitrary given $\mathbf{y}\in R^n$ is guaranteed by fairly general conditions based on the theory of the degree of mapping. Then it is shown that an iterative algorithm (generalized Katzenelson algorithm) leads to a solution in a finite number of iteration steps. Finally, a comprehensive study of physical nonlinear elements demonstrates that the theory can be applied to most of the currently used nonlinear networks.

1. Introduction. The problem of analyzing large scale nonlinear resistor networks is becoming of widespread interest. In particular, the theory of networks composed of resistors with continuous, piecewise-linear characteristics has been developed rapidly [1]–[6]. This is partly due to the ease of numerical computation.

In 1965, Katzenelson gave an algorithm for solving networks which contain uncoupled, piecewise-linear resistors of the strictly increasing type [1]. The piecewise-linear model was then extended to include coupled resistor networks [3]. Following these works, Kuh and Hajj made an attempt to modify the Katzenelson algorithm to deal with networks having multiple solutions, yet there remained many theoretical problems to be investigated [4].

In 1972, Fujisawa and Kuh presented a fairly general theory of piecewise-linear mappings and some sufficient conditions for such a mapping to be a homeomorphism [5]. They also showed that the application of the Katzenelson algorithm leads to the unique solution whenever the mapping is a homeomorphism.

Recently, Fujisawa, Kuh and Ohtsuki further extended the theory in two directions [6]. First, they gave a sufficient condition which guarantees the existence of solutions, and also verified the applicability of the original Katzenelson algorithm to this case. Secondly, they formulated an efficient computational method which exploits not only the very nature of piecewise-linear mappings but also the sparsity of Jacobian matrices.

In spite of the development mentioned above, there still remains a serious question on availability of a convergent algorithm for solving a more general class of resistor networks which one often encounters in practice, i.e., those containing so-called "active" elements such as transistors and tunnel diodes. Common to almost all existing iterative algorithms, including the original Katzenelson algorithm and the Newton–Raphson algorithm, to solve $\mathbf{f}(\mathbf{x})=\mathbf{y}$ for a given $\mathbf{y}$, is that a sequence $\{\mathbf{x}^{(k)};\ k=1, 2, \cdots\}$ of approximate solutions has to be generated in such a way that the value of $\|\mathbf{y}-\mathbf{f}(\mathbf{x}^{(k)})\|$ decreases monotonically as k increases. Because of this restriction, one often encounters the case where a solution cannot be found even if it exists.

* Received by the editors February 23, 1975.

† Central Research Laboratories, Nippon Electron Co., Ltd., Kawasaki 213, Japan.
‡ School of Engineering Sciences, Osaka University, Toyonaka 560, Japan.

It is the purpose of this paper to give an affirmative answer to the question posed in the above paragraph. To be more precise, the main theorem asserts the existence of a solution of resistor networks of a fairly general class, like those containing "active" elements. It is then shown that the application of an algorithm (*generalized Katzenelson algorithm*) leads to a solution in a finite number of iteration steps if the condition of the main theorem is satisfied. The condition for a resistor network to have at least one solution, or for the solution algorithm to converge, is imposed only on the characteristics of component resistors, i.e., independent of the network topology. This type of condition is very favorable in implementing a general circuit analysis program [7].

The solution algorithm presented here is designed, by taking advantage of the simplicity of piecewise-linear mappings, so that it converges to a solution of $\mathbf{y}=\mathbf{f}(\mathbf{x})$ for a given $\mathbf{y}$ even if a number of local minimums of $\|\mathbf{y}-\mathbf{f}(\mathbf{x})\|$ exist. In other words, the original Katzenelson algorithm is generalized so that it treats piecewise-linear mappings which possess regions with both positive and negative Jacobian determinants and even those with singular Jacobian matrices. It is only fair to mention here that the generalization of the Katzenelson algorithm in this direction is motivated by the work of Kuh and Hajj [4].

After the introduction of fundamental propositions in § 2, fairly general theorems on the existence of solutions based on the theory of the degree of mapping [8], [12] are given in § 3. Then it is shown in § 4 that the generalized Katzenelson algorithm leads to a solution if the given piecewise-linear mappings satisfy a certain condition (*1-degree condition*). In § 5, it is concluded that a very general class of nonlinear networks are covered by the theory described in the previous sections. More specifically, it is shown that if each component resistor possesses a certain property (*Property* U), the network equation satisfies the 1-degree condition independent of the interconnection of the resistors. In Appendix A, it is shown that the definition of the degree of piecewise-linear mappings given in this paper can be derived from the one in [8, p. 154]. Appendix B shows that typical physical elements possess Property U under a pertinent piecewise-linear approximation of their characteristics.

2. Preliminaries. The following notations are used throughout this paper. The n-dimensional Euclidean space is denoted by R^n, each element of which is an n-column vector. If $\mathbf{x}\in R^n$, then x_i denotes the ith component of $\mathbf{x}$ for $i=1, 2, \cdots, n$. For any $\mathbf{x}, \mathbf{y}\in R^n$, the inner product $\sum_{i=1}^{n} x_i y_i$ is denoted by $\langle \mathbf{x}, \mathbf{y}\rangle$. The norm of $\mathbf{x}$ is denoted by $\|\mathbf{x}\|=\langle \mathbf{x}, \mathbf{x}\rangle^{1/2}$. The superscript T stands for the transpose of a vector or a matrix. The determinant of a square matrix $\mathbf{J}$ is denoted by $\det \mathbf{J}$. For any subset $S\subset R^n$, the closure of S is denoted by $\bar{S}$ and the boundary by ∂S.

A continuous mapping $\mathbf{f}: R^n \to R^n$ is said to be (i) *norm-coercive* if

$$\|\mathbf{f}(\mathbf{x})\|\to\infty \quad \text{as } \|\mathbf{x}\|\to\infty, \tag{1}$$

(ii) *weakly coercive* if there exists a $\mathbf{u}\in R^n$ such that

$$\langle \mathbf{f}(\mathbf{x}), \mathbf{x}-\mathbf{u}\rangle/\|\mathbf{x}-\mathbf{u}\|\to\infty \quad \text{as } \|\mathbf{x}\|\to\infty, \tag{2}$$

and (iii) *strongly coercive* if (2) holds for all $\mathbf{u}\in R^n$. It is easy to see from the Schwarz inequality that weak coerciveness implies norm-coerciveness.

A continuous mapping $\mathbf{f}: R^n \to R^n$ is said to be (i) *monotone* if

$$\langle \mathbf{f}(\mathbf{u}) - \mathbf{f}(\mathbf{v}), \mathbf{u} - \mathbf{v} \rangle \geqq 0 \tag{3}$$

for all $\mathbf{u}, \mathbf{v} \in R^n$, (ii) *strictly monotone* if the strict inequality holds in (3) whenever $\mathbf{u} \neq \mathbf{v}$, and (iii) *uniformly monotone* if there exists a $\gamma > 0$ such that

$$\langle \mathbf{f}(\mathbf{u}) - \mathbf{f}(\mathbf{v}), \mathbf{u} - \mathbf{v} \rangle \geqq \gamma \|\mathbf{u} - \mathbf{v}\|^2 \tag{4}$$

for all $\mathbf{u}, \mathbf{v} \in R^n$.

A continuous mapping $\mathbf{f}: R^n \to R^n$ is said to be (i) *passive* on $\mathbf{u}$ if there exists a $\mathbf{u} \in R^n$ such that

$$\langle \mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{u}), \mathbf{x} - \mathbf{u} \rangle \geqq 0 \tag{5}$$

for all $\mathbf{x} \in R^n$, (ii) *strictly passive* on $\mathbf{u}$ if the strict inequality holds in (5) for all $\mathbf{x} \neq \mathbf{u}$, and (iii) *uniformly passive* on $\mathbf{u}$ if there exists a $\gamma > 0$ such that

$$\langle \mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{u}), \mathbf{x} - \mathbf{u} \rangle \geqq \gamma \|\mathbf{x} - \mathbf{u}\|^2 \tag{6}$$

for all $\mathbf{x} \in R^n$. It is clear that if $\mathbf{f}$ is monotone, strictly monotone or uniformly monotone, then $\mathbf{f}$ is passive, strictly passive or uniformly passive on any $\mathbf{u} \in R^n$, respectively. It is also clear that if $\mathbf{f}$ is uniformly monotone or uniformly passive on $\mathbf{u}$, then $\mathbf{f}$ is strongly coercive or weakly coercive, respectively.

Remark 1. A physical interpretation of passivity is as follows. Let $\mathbf{f}$ be the voltage-current characteristics of an n-port resistor. Thus if x_i represents the voltage (current) of the ith port, then $(\mathbf{f}(\mathbf{x}))_i$ represents the corresponding current (voltage). Therefore $\langle \mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{u}), \mathbf{x} - \mathbf{u} \rangle$ expresses the net power flow into the resistor when it is operated on the point $\mathbf{u}$. If this inner product is positive, it means that the resistor consumes energy [9].

A continuous mapping $\mathbf{f}: R^n \to R^n$ is said to be *piecewise-linear* if the whole space R^n is devided into a finite number of convex polyhedral regions by a finite number of $(n-1)$-dimensional hyperplanes so that, in each region, $\mathbf{f}$ is an affine mapping. Namely, in each region, say R_k, $\mathbf{f}$ is represented by

$$\mathbf{f}(\mathbf{x}) = \mathbf{J}^{(k)}\mathbf{x} + \mathbf{b}^{(k)} \tag{7}$$

for any $\mathbf{x} \in R_k$, where $\mathbf{J}^{(k)}$ is a constant $n \times n$ matrix called a *Jacobian matrix* and $\mathbf{b}^{(k)}$ is a constant n-vector. For a piecewise-linear mapping $\mathbf{f}: R^n \to R^n$, the term *boundary hyperplane* is, henceforth, used to mean an $(n-1)$-dimensional hyperplane which separates two neighboring regions. Throughout the arguments hereafter in this section, the mappings $\mathbf{f}: R^n \to R^n$ under consideration are assumed to be continuous and piecewise-linear.

The continuity of a piecewise-linear mapping on boundary hyperplanes leads to the following property, which plays a key role in the solution algorithm [5], [6].

PROPOSITION 1. *If two regions R_1 and R_2 with Jacobian matrices $\mathbf{J}^{(1)}$ and $\mathbf{J}^{(2)}$, respectively, have a common $(n-1)$-dimensional boundary hyperplane H (see Fig. 1), then there exists a constant n-vector $\mathbf{c}$ such that*

$$\mathbf{J}^{(2)} - \mathbf{J}^{(1)} = \mathbf{c}\mathbf{r}^{\mathrm{T}}, \tag{8}$$

where $\mathbf{r}$ is the normal vector of H.

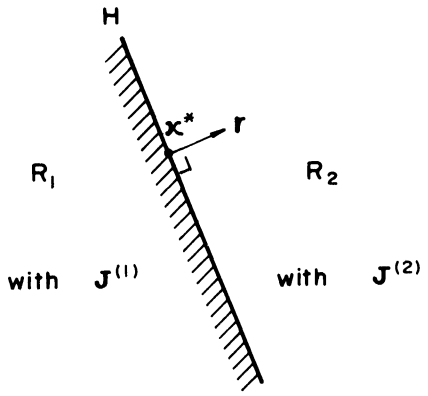


FIG. 1. *Two neighboring regions having a common boundary hyperplane*

As an immediate consequence of the above relation, one obtains the following important property of piecewise-linear mappings [5].

PROPOSITION 2. *Let* $\mathbf{J}^{(1)}$ *and* $\mathbf{J}^{(2)}$ *be as in Proposition* 1. *Then the ranks of* $\mathbf{J}^{(1)}$ *and* $\mathbf{J}^{(2)}$ *differ at most by one.*

Since the number of regions is finite, one easily obtains the following proposition [5].

PROPOSITION 3. *There exists a* $\gamma > 0$ *such that*

$$\|\mathbf{f}(\mathbf{u}) - \mathbf{f}(\mathbf{v})\| \leqq \gamma \|\mathbf{u} - \mathbf{v}\| \tag{9}$$

for all $\mathbf{u}, \mathbf{v} \in R^n$, *i.e.*, $\mathbf{f}$ *is Lipschitzian.*

Remark 2. For the value of γ it suffices to take the maximum of matrix norms of all the Jacobian matrices [8].

The following property is an immediate consequence of Proposition 3 and the definition of piecewise-linear mapping.

PROPOSITION 4. *If* $\mathbf{f}$ *is weakly coercive, then it implies that* $\mathbf{f}$ *is strongly coercive, too.*

When a continuous mapping under consideration is piecewise-linear, the term *coercive* is, henceforth, used to mean either weakly coercive or strongly coercive.

If a continuous piecewise-linear mapping is strictly monotone, then all the Jacobian matrices must be positive definite.[1] Therefore one can easily derive the inequality (4), where γ is the minimum of the eigenvalues of symmetric parts of all the Jacobian matrices [3], [5]. This proves the following proposition.

PROPOSITION 5. *If* $\mathbf{f}$ *is strictly monotone, then it implies that* $\mathbf{f}$ *is uniformly monotone, too.*

Remark 3. If a continuous mapping $\mathbf{f}: R^n \to R^n$ is strictly monotone, then $\mathbf{f}$ is one-to-one. In addition, if $\mathbf{f}$ is piecewise-linear, i.e., uniformly monotone due to Proposition 5, then $\mathbf{f}$ is coercive. Thus $\mathbf{f}$ maps R^n onto R^n (see § 5.4, Theorem A). Therefore a continuous, strictly monotone, piecewise-linear mapping is a homeomorphism.

[1] In this paper a real square matrix is said to be *positive definite* if its symmetric part is positive definite.

The norm-coerciveness of a piecewise-linear mapping, which plays an essential role in the degree invariance property discussed in the next section, implies the following property.

PROPOSITION 6. *If* $\mathbf{f}$ *is a norm-coercive, then*

$$\mathbf{f}(E) \subset \mathbf{f}(B), \tag{10}$$

where $E \subset R^n$ *is the set of all points in regions with singular Jacobian matrices, and* $B \subset R^n$ *is the set of all points on boundary hyperplanes.*

It suffices for the proof to show that, for any interior point $\mathbf{x}$ of a region with a singular Jacobian matrix, there exists a point $\mathbf{x}' \in B$ such that $\mathbf{f}(\mathbf{x}') = \mathbf{f}(\mathbf{x})$. Let R_k be the region in which $\mathbf{x}$ lies. Then the singularity of the Jacobian matrix $\mathbf{J}^{(k)}$ implies the existence of a unit vector such that $\mathbf{J}^{(k)}\boldsymbol{\alpha} = \mathbf{0}$. Hence $\mathbf{f}(\mathbf{w}(\lambda)) = \mathbf{f}(\mathbf{x})$ for any real λ whenever $\mathbf{w}(\lambda) = \mathbf{x} + \lambda\boldsymbol{\alpha} \in R_k$. If the straight line $\{\mathbf{w}(\lambda) | \lambda \in (-\infty, \infty)\}$ wholly lies in region R_k, it contradicts the norm-coerciveness assumption. Therefore the line must meet a boundary hyperplane of R_k at a point, say $\mathbf{x}'$. This completes the proof.

3. Degree of mapping and existence of solutions. The problem considered in this section is whether

$$\mathbf{f}(\mathbf{x}) = \mathbf{y} \tag{11}$$

has a solution for an arbitrary given $\mathbf{y} \in R^n$, where $\mathbf{f}$: $R^n \to R^n$ is a continuous, piecewise-linear mapping. Throughout this section, $B \subset R^n$ denotes the set of points on boundary hyperplanes and $E \subset R^n$ the set of points in regions with singular Jacobian matrices.

Let $C \subset R^n$ be an open bounded set and assume that $\mathbf{y} \notin \mathbf{f}(B) \cup \mathbf{f}(E) \cup \mathbf{f}(\partial C)$. Then there are, at most, finitely many solutions of (11) in C. Let $\{\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(m)}\} = \{\mathbf{x} \in C | \mathbf{f}(\mathbf{x}) = \mathbf{y}\}$ be the set of solutions and $\mathbf{J}^{(i)}$, $i = 1, \cdots, m$, be the Jacobian matrix of the region in which $\mathbf{x}^{(i)}$ lies. Then the integer

$$\deg(\mathbf{f}, C, \mathbf{y}) = \sum_{i=1}^{m} \operatorname{sgn} \det \mathbf{J}^{(i)} \tag{12}$$

is called the *degree* of $\mathbf{f}$ at $\mathbf{y}$ with respect to C.

An analytic definition of the degree of continuous mappings and its properties are found in [8]. The definition (12) for $\mathbf{y} \notin \mathbf{f}(B) \cup \mathbf{f}(E)$ coincides with the one given in [8]. (See Appendix A for the proof.)

The following invariance property of the degree [8] is needed in the proof of Theorem 1: Let C be an open bounded set and assume that $\mathbf{y}^{(1)}, \mathbf{y}^{(2)} \notin \mathbf{f}(\partial C)$. If the two points can be connected by a continuous path without passing through $\mathbf{f}(\partial C)$, then

$$\deg(\mathbf{f}, C, \mathbf{y}^{(1)}) = \deg(\mathbf{f}, C, \mathbf{y}^{(2)}). \tag{13}$$

If $\mathbf{f}$ is norm-coercive, then for any $\mathbf{x} \in R^n$, $\mathbf{f}^{-1}(\mathbf{f}(\mathbf{x}))$ is a compact set. Therefore one can determine the maximal connected subset of $\mathbf{f}^{-1}(\mathbf{f}(\mathbf{x}))$ containing $\mathbf{x}$ [10, p. 54] which, henceforth, is denoted by $X(\mathbf{f}, \mathbf{x})$.

THEOREM 1. *Let* $\mathbf{f}$: $R^n \to R^n$ *be a continuous, norm-coercive, piecewise-linear mapping and* $\mathbf{x}^* \in R^n$ *be a point. Furthermore, let* C *be an open bounded set*

such that $X(\mathbf{f}, \mathbf{x}^*) \subset C$ *and* $\bar{C} \cap [\mathbf{f}^{-1}(\mathbf{f}(\mathbf{x}^*)) - X(\mathbf{f}, \mathbf{x}^*)] = \varnothing$. *Then there exists an open neighborhood* D *of* $\mathbf{f}(\mathbf{x}^*)$ such that (13) holds for any $\mathbf{y}^{(1)}, \mathbf{y}^{(2)} \in D - \mathbf{f}(B)$.

Proof. First, note that the degrees at $\mathbf{y}^{(1)}$ and $\mathbf{y}^{(2)}$ are well-defined by (12) if $\mathbf{f}$ is norm-coercive (see Proposition 6). From the continuity of $\mathbf{f}$, the compactness of ∂C and $\mathbf{f}(\mathbf{x}^*) \notin \mathbf{f}(\partial C)$, it is easily seen that $\mathbf{f}(\partial C) \cap D = \varnothing$ for a sufficiently small open neighborhood D of $\mathbf{f}(\mathbf{x}^*)$. Hence the theorem follows from the degree invariance property.

Remark 4. The connected, compact set $X(\mathbf{f}, \mathbf{x})$ is complex [11, p. 13] as shown in Fig. 2 since $\mathbf{f}$ is piecewise-linear. Each one-dimensional part is in a region with Jacobian matrix of rank $n-1$, each two-dimensional part is in a region with that of rank $n-2$, and so forth.

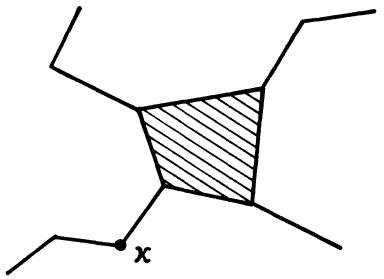


FIG. 2. *An illustrative example of* $X(\mathbf{f}, \mathbf{x})$

For a point $\mathbf{y} \in R^n$ and unit vector $\boldsymbol{\alpha} \in R^n$, consider the straight line

$$L(\mathbf{y}, \boldsymbol{\alpha}) = \{\mathbf{z} \in R^n \mid \mathbf{z} = \mathbf{y} + \lambda \boldsymbol{\alpha}, \lambda \in (-\infty, \infty)\} \tag{14}$$

THEOREM 2. *Let* $\mathbf{f}: R^n \to R^n$ *be a continuous, norm-coercive, piecewise-linear mapping,* $\mathbf{x}^* \in R^n$ *be a point and* $\boldsymbol{\alpha} \in R^n$ *be a unit vector. Furthermore, assume that* $\mathbf{f}(\mathbf{x}^*) + \lambda \boldsymbol{\alpha} \notin \mathbf{f}(B)$ for all sufficiently small $\lambda \neq 0$. *Then there exist an even number of unit vectors* $\boldsymbol{\beta}$ *such that, for all sufficiently small* $\nu > 0$,

$$\mathbf{v} + \nu \boldsymbol{\beta} \notin X(\mathbf{f}, \mathbf{x}^*) \tag{15}$$

and

$$\mathbf{f}(\mathbf{v} + \nu \boldsymbol{\beta}) \in L(\mathbf{f}(\mathbf{x}^*), \alpha), \tag{16}$$

where $\mathbf{v} \in \partial X(\mathbf{f}, \mathbf{x}^*)$ *is a point uniquely determined for each* $\boldsymbol{\beta}$.

Proof. Let C be an open bounded set as in Theorem 1. Then there exists an open neighborhood D of $\mathbf{f}(\mathbf{x}^*)$ such that (13) holds for any $\mathbf{y}^{(1)}, \mathbf{y}^{(2)} \in D - f(B)$. Namely, there exists a $\bar{\lambda} > 0$ such that $\mathbf{f}(\mathbf{x}^*) \pm \lambda \boldsymbol{\alpha} \in D - \mathbf{f}(B)$ for all $\lambda \in (0, \bar{\lambda}]$, and the degree preserves a constant value, say d. Let $\mathbf{y}^{(1)} = \mathbf{f}(\mathbf{x}^*) + \bar{\lambda} \boldsymbol{\alpha}$ and $\mathbf{y}^{(2)} = \mathbf{f}(\mathbf{x}) - \bar{\lambda} \boldsymbol{\alpha}$.

Suppose $\mathbf{f}(\mathbf{x}) = \mathbf{y}^{(1)}$ has $n^{(1)}$ solutions in C, $n_+^{(1)}$ (resp., $n_-^{(1)}$) of which are in regions with positive (resp., negative) Jacobian determinants. Then $d = n_+^{(1)} - n_-^{(1)}$ and $n^{(1)} = n_+^{(1)} + n_-^{(1)}$. Integers $n^{(2)}$, $n_+^{(2)}$ and $n_-^{(2)}$ are defined in the same way for $\mathbf{f}(\mathbf{x}) = \mathbf{y}^{(2)}$. Then $d = n_+^{(2)} - n_-^{(2)}$ and $n^{(2)} = n_+^{(2)} + n_-^{(2)}$. Therefore the total number of

solutions of $\mathbf{f}(\mathbf{x})=\mathbf{y}^{(1)}$ and $\mathbf{f}(\mathbf{x})=\mathbf{y}^{(2)}$ is equal to $2(d+n_{-}^{(1)}+n_{-}^{(2)})$, which is an even number.

First, consider the case where $\mathbf{x}^* \notin B \cup E$ and hence $\mathbf{x}^*$ is an interior point of a region, say R_k, with nonsingular Jacobian matrix. In this case, the single point $\mathbf{x}^*$ constitutes whole $X(\mathbf{f}, \mathbf{x}^*)$, and hence $\mathbf{v}=\mathbf{x}^* \in \partial X(\mathbf{f}, \mathbf{x}^*)$. Then it is clear from the definition of $\bar{\lambda}$ that $\mathbf{f}(\mathbf{x})=\mathbf{y}^{(i)}$ has a unique solution, say $\mathbf{x}^{(i)}$, in R_k for $i=1, 2$. Thus exactly two unit vectors $\boldsymbol{\beta}^{(1)}=(\mathbf{x}^{(1)}-\mathbf{x}^*)/\|\mathbf{x}^{(1)}-\mathbf{x}^*\|$ and $\boldsymbol{\beta}^{(2)}=(\mathbf{x}^{(2)}-\mathbf{x}^*)/\|\mathbf{x}^{(2)}-\mathbf{x}\|=-\boldsymbol{\beta}^{(1)}$ are found so that (15) and (16) hold for $\boldsymbol{\beta}=\boldsymbol{\beta}^{(i)}$, $i=1, 2$.

Next consider the case where $\mathbf{x}^* \in B \cup E$. It is clear that no two solutions of $\mathbf{f}(\mathbf{x})=\mathbf{y}^{(1)}$ or $\mathbf{f}(\mathbf{x})=\mathbf{y}^{(2)}$ can be found in the same region. Let $\mathbf{x}^{(1)}$ be a solution of $\mathbf{f}(\mathbf{x})=\mathbf{y}^{(1)}$ and R_k be the region containing $\mathbf{x}^{(1)}$ as an interior point. Since the Jacobian matrix $\mathbf{J}^{(k)}$ of R_k is nonsingular, a unit vector $\boldsymbol{\beta}$ such that $\mathbf{J}^{(k)}\boldsymbol{\beta}=\boldsymbol{\alpha}$ is uniquely determined. Consider the set $\{\mathbf{x} \in R^n \mid \mathbf{x}=\mathbf{x}^{(1)}-\mu\boldsymbol{\beta}\}$ which is contained in R_k for sufficiently small $\mu \geqq 0$. Then for such a μ, $\mathbf{f}(\mathbf{x}^{(1)}-\mu\boldsymbol{\beta})=\mathbf{f}(\mathbf{x}^*)+(\bar{\lambda}-\mu)\boldsymbol{\alpha}$. It should be noted here that $\mathbf{f}(\mathbf{x}^{(1)}-\mu\boldsymbol{\beta}) \notin \mathbf{f}(B)$, i.e., $\mathbf{x}^{(1)}-\mu\boldsymbol{\beta}$ is an interior point of R_k for any $\mu \in [0, \bar{\lambda})$. It follows from the continuity of $\mathbf{f}$ that $\mathbf{f}(\mathbf{x}^{(1)}-\bar{\lambda}\boldsymbol{\beta})=\mathbf{f}(\mathbf{x}^*)$ and that the point $\mathbf{v}=\mathbf{x}^{(1)}-\bar{\lambda}\boldsymbol{\beta}$ lies on the boundary of R_k, i.e., $\mathbf{v} \in \partial X(\mathbf{f}, \mathbf{x}^*)$. Let $\nu=\bar{\lambda}-\mu$; then $\mathbf{v}+\nu\boldsymbol{\beta} \notin X(\mathbf{f}, \mathbf{x}^*)$ and $\mathbf{f}(\mathbf{v}+\nu\boldsymbol{\beta})=\mathbf{f}(\mathbf{x}^*)+\nu\boldsymbol{\alpha}$ for all $\mu \in [0, \bar{\lambda})$, i.e., for all $\nu \in (0, \bar{\lambda}]$. Hence (15) and (16) hold for all sufficiently small $\nu>0$. Since the same argument is valid for any solution of $\mathbf{f}(\mathbf{x})=\mathbf{y}^{(1)}$ or $\mathbf{f}(\mathbf{x})=\mathbf{y}^{(2)}$, the proof has been completed.

Remark 5. This theorem may be viewed as a generalization of Theorem 4 of [6].

Remark 6. When $\mathbf{x}^* \in B \cup E$, each $\mathbf{v}$ must be found on the boundary of a region having a nonsingular Jacobian matrix.

Remark 7. If $\mathbf{x}^*$ lies on a single boundary hyperplane H which separates two regions with nonsingular Jacobian matrices as shown in Fig. 1, then for any unit vector $\boldsymbol{\alpha}$, the identity

$$\langle(\mathbf{J}^{(1)})^{-1}\boldsymbol{\alpha}, \mathbf{r}\rangle \det \mathbf{J}^{(1)}=\langle(\mathbf{J}^{(2)})^{-1}\boldsymbol{\alpha}, \mathbf{r}\rangle \det \mathbf{J}^{(2)} \tag{17}$$

holds [5]. Suppose the determinants have the same sign. Then if $\langle(\mathbf{J}^{(1)})^{-1}\boldsymbol{\alpha}, \mathbf{r}\rangle>0$, $\boldsymbol{\beta}^{(1)}=(\mathbf{J}^{(1)})^{-1}\boldsymbol{\alpha}/\|(\mathbf{J}^{(1)})^{-1}\boldsymbol{\alpha}\|$ is only one unit vector such that $\mathbf{f}(\mathbf{x}^*+\nu\boldsymbol{\beta}^{(1)})=\mathbf{f}(\mathbf{x}^*)+\nu\boldsymbol{\alpha}$ for all sufficiently small $\nu>0$. Also $\boldsymbol{\beta}^{(2)}=-(\mathbf{J}^{(2)})^{-1}\boldsymbol{\alpha}/\|(\mathbf{J}^{(2)})^{-1}\boldsymbol{\alpha}\|$ is only one unit vector such that $\mathbf{f}(\mathbf{x}^*+\nu\boldsymbol{\beta}^{(2)})=\mathbf{f}(\mathbf{x}^*)-\nu\boldsymbol{\alpha}$ for all sufficiently small $\nu>0$. The case $\langle(\mathbf{J}^{(1)})^{-1}\boldsymbol{\alpha}, \mathbf{r}\rangle<0$ can be treated in the same way. Next consider the case where the two determinants have opposite signs. Then, if $\langle(\mathbf{J}^{(1)})^{-1}\boldsymbol{\alpha}, \mathbf{r}\rangle>0$, $\mathbf{f}(\mathbf{x}^*+\nu\boldsymbol{\beta})=\mathbf{f}(\mathbf{x}^*)+\nu\boldsymbol{\alpha}$ for all sufficiently small $\nu>0$ with $\boldsymbol{\beta}=\boldsymbol{\beta}^{(1)}$ or $\boldsymbol{\beta}=-\boldsymbol{\beta}^{(2)}$ and no other unit vector possesses this property. It is important to note that there exists no vector $\boldsymbol{\beta}$ such that $\mathbf{f}(\mathbf{x}^*+\nu\boldsymbol{\beta})=\mathbf{f}(\mathbf{x}^*)-\nu\boldsymbol{\alpha}$ for any sufficiently small $\nu>0$. If $\langle(\mathbf{J}^{(1)})^{-1}\boldsymbol{\alpha}, \mathbf{r}\rangle<0$, then $\mathbf{f}(\mathbf{x}^*+\nu\boldsymbol{\beta})=\mathbf{f}(\mathbf{x}^*)-\nu\boldsymbol{\alpha}$ for all sufficiently small $\nu>0$ with $\boldsymbol{\beta}=\boldsymbol{\beta}^{(1)}$ or $\boldsymbol{\beta}=-\boldsymbol{\beta}^{(2)}$, and there exists no vector $\boldsymbol{\beta}$ such that $\mathbf{f}(\mathbf{x}^*+\nu\boldsymbol{\beta})=\mathbf{f}(\mathbf{x}^*)+\nu\boldsymbol{\alpha}$ for any sufficiently small $\nu>0$. Therefore if $\mathbf{x}^*$ lies on a single boundary hyperplane separating two regions with nonsingular Jacobian matrices, there are exactly two such $\boldsymbol{\beta}$'s that satisfy (16) with $\mathbf{v}=\mathbf{x}^*$.

Remark 8. The existence of an even number of $\boldsymbol{\beta}$'s possibly holds even if $\mathbf{f}(\mathbf{x}^*)+\lambda\boldsymbol{\alpha} \in \mathbf{f}(B)$ for all sufficiently small λ. However, in this case, the interpretation of the number of $\boldsymbol{\beta}$'s is slightly different. Since $\mathbf{f}(B)$ is a set union of a finite

number of at most $(n-1)$-dimensional convex polyhedral domains, there exist a unit vector $\varepsilon \in R^n$ and $\bar{\lambda}, \bar{\mu} > 0$ such that $\mathbf{f}(\mathbf{x}^*) + \lambda(\alpha + \mu\varepsilon) \notin \mathbf{f}(B)$ for all $\lambda \in (0, \bar{\lambda}]$ and $\mu \in (0, \bar{\mu}]$. Thus for any fixed $\mu \in (0, \bar{\mu}]$, there exist an even number of unit vectors $\beta^{(i)}(\mu)$, $i = 1, 2, \cdots, 2m$, and corresponding points $\mathbf{v}^{(i)} \in \partial X(\mathbf{f}, \mathbf{x}^*)$, $i = 1, 2, \cdots, 2m$, such that $\mathbf{v}^{(i)} + \nu\beta^{(i)}(\mu) \notin X(\mathbf{f}, \mathbf{x}^*)$ and $\mathbf{f}(\mathbf{v}^{(i)} + \nu\beta^{(i)}(\mu)) \in L(\mathbf{f}(\mathbf{x}^*), \alpha + \mu\varepsilon)$ for all sufficiently small $\nu > 0$. Let R_i be the region in which $\mathbf{v}^{(i)} + \nu\beta^{(i)}(\mu)$ lies. Then since the Jacobian matrix of R_i is nonsingular, $\nu\beta^{(i)}(0) = \lim_{\mu \to 0} \nu\beta^{(i)}(\mu)$ exists in R_i for any fixed, sufficiently small $\nu > 0$, preserving the relations $\mathbf{v}^{(i)} + \nu\beta^{(i)}(0) \notin X(\mathbf{f}, \mathbf{x}^*)$ and $\mathbf{f}(\mathbf{v}^{(i)} + \nu\beta^{(i)}(0)) \in L(\mathbf{f}(\mathbf{x}^*), \alpha)$. Since $\mathbf{v}^{(i)} + \nu\beta^{(i)}(0)$ may lie on the boundary of R_i, it is possible that two or more $\beta^{(i)}(\mu)$'s merge into a single vector as $\mu \to 0$. In this case, such β's should be considered to have the corresponding multiplicity. It should also be noted that, as $\mu \to 0$, some new β's might appear. Such β's should be excluded in counting the number. In this way of counting, the even number property still holds. Note that the number of β's depends on which ε is taken. The following example demonstrates what happens in such a degenerate case.

Example 1. The whole space R^2 is divided into four regions by the x_1-axis and x_2-axis as shown in Fig. 3. Let $\mathbf{f}: R^2 \to R^2$ be defined as follows:

$$R_1: y_1 = x_1 + x_2, \qquad y_2 = x_2,$$

$$R_2: y_1 = x_1, \qquad y_2 = -x_2,$$

$$R_3: y_1 = -x_1, \qquad y_2 = -2x_1 - x_2,$$

$$R_4: y_1 = -x_1 + x_2, \qquad y_2 = -2x_1 + x_2.$$

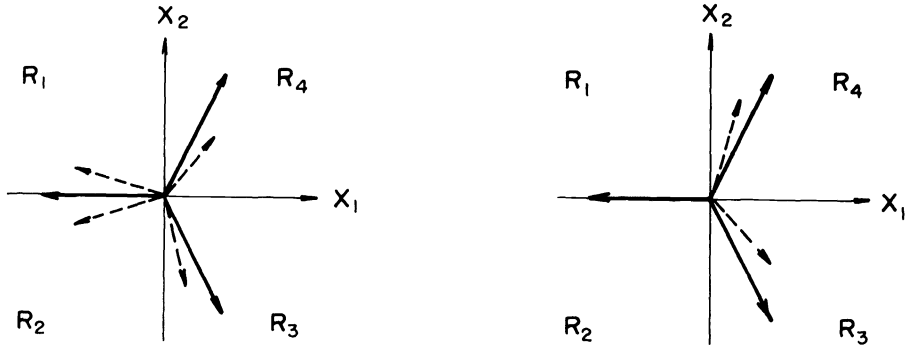


FIG. 3. *An illustrative example of degenerate case for Theorem* 2

Note that $\mathbf{f}$ is norm-coercive. Let $\mathbf{x}^* = (0, 0)^T$ and $\mathbf{a} = (-1, 0)^T$; then $X(\mathbf{f}, \mathbf{x}^*) = \{\mathbf{x}^*\}$ and $L(\mathbf{f}(\mathbf{x}^*), \alpha)$ runs on the y_1-axis of the y-space. As is shown in Fig. 3, three linearly-independent β's $(-1, 0)^T$, $(1, -2)^T$ and $(1, 2)^T$ satisfy (16) with $\mathbf{v} = \mathbf{x}^*$. If one considers the perturbed straight line $L(\mathbf{f}(\mathbf{x}^*), \alpha + \varepsilon)$, where $\varepsilon = (0, \varepsilon)^T$ and $0 < \varepsilon \ll 1$, there are four β's, $(-1-\varepsilon, \varepsilon)^T$, $(-1, -\varepsilon)^T$, $(1, -2-\varepsilon)^T$ and $(1+\varepsilon, 2+\varepsilon)^T$ as illustrated by the broken lines in Fig. 3(a). In this case, the first two vectors merge into the single vector $(-1, 0)^T$ as $\varepsilon \to 0$. Thus the vector $(-1, 0)^T$ may be considered to have multiplicity 2. If one considers the perturbed line $L(\mathbf{f}(\mathbf{x}^*),$

$\alpha-\varepsilon$), on the other hand, there are only two β's $(1, -2+\varepsilon)^T$ and $(1-\varepsilon, 2-\varepsilon)^T$ as illustrated by the broken lines in Fig. 3(b). At the limit as $\varepsilon \to 0$, the new vector $(-1, 0)^T$ suddenly appears. Therefore the vector $(-1, 0)^T$ can be excluded in counting the number of β's. In this example, the number changes depending on ε. Note, however, that the even number property holds no matter in which direction the straight line is perturbed.

For a continuous, norm-coercive, piecewise-linear mapping $\mathbf{f}: R^n \to R^n$, let $\mathbf{y} \notin \mathbf{f}(B)$. Then the number of solutions of (11) in R^n is finite. This leads to the following definition of the degree in the large. Let $\{\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(m)}\}$ be the set of solutions of (11) and $\mathbf{J}^{(i)}$, $i = 1, 2, \cdots, m$, be the Jacobian matrix of the region in which $\mathbf{x}^{(i)}$ lies. Then the integer

$$\deg(\mathbf{f}, R^n, \mathbf{y}) = \sum_{i=1}^{m} \operatorname{sgn} \det \mathbf{J}^{(i)} \tag{18}$$

is called the *degree* of $\mathbf{f}$ at $\mathbf{y}$ *in the large.*

THEOREM 3. *Let* $\mathbf{f}: R^n \to R^n$ *be a continuous, norm-coercive, piecewise-linear mapping. Then*

$$\deg(\mathbf{f}, R^n, \mathbf{y}^{(1)}) = \deg(\mathbf{f}, R^n, \mathbf{y}^{(2)}) \tag{19}$$

holds for any $\mathbf{y}^{(1)}, \mathbf{y}^{(2)} \in R^n - \mathbf{f}(B)$.

Proof. The norm-coerciveness of $\mathbf{f}$ implies that there exists a sufficiently large open ball C such that all the solutions of $\mathbf{f}(\mathbf{x}) = \mathbf{y}^{(i)}$, $i = 1, 2$, are contained in C and that the line segment $\{\mathbf{x} \in R^n | \mathbf{y} = \lambda \mathbf{y}^{(1)} + (1-\lambda)\mathbf{y}^{(2)}, \lambda \in [0, 1]\}$ does not intersect $\mathbf{f}(\partial C)$. Then the degree invariance property (13) yields (19).

In what follows the Kronecker theorem [8, p. 161] plays an essential role in asserting the existence of solutions of (11). The Kronecker theorem says the following: Let $\mathbf{f}: R^n \to R^n$ be continuous and C be an open bounded set. If $\mathbf{y} \notin \mathbf{f}(\partial C)$ and $\deg(\mathbf{f}, C, \mathbf{y}) \neq 0$, then $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ has a solution in C.

THEOREM 4. *Let* $\mathbf{f}: R^n \to R^n$ *be a continuous, norm-coercive, piecewise-linear mapping. If, for all unbounded regions, the Jacobian determinants are nonnegative and, for at least one unbounded region, the Jacobian determinant is positive, then* $\mathbf{f}$ *maps* R^n *onto itself.*

Proof. Let R_i be an unbounded region with $\det \mathbf{J}^{(i)} > 0$. Then it follows from the norm-coerciveness of $\mathbf{f}$ that there exists a $\mathbf{u} \in R_i$ such that $\mathbf{f}(\mathbf{u}) \notin \mathbf{f}(B)$ and no solution of $\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{u})$ exists in any bounded region. For such a point $\mathbf{u}$, it is clear that $\deg(\mathbf{f}, R^n, \mathbf{f}(\mathbf{u})) > 0$. Then it follows from Theorem 3 and the Kronecker theorem that (11) has at least one solution for any $\mathbf{y} \in R^n - \mathbf{f}(B)$. If $\mathbf{y} \in \mathbf{f}(B)$, then there exists a point $\mathbf{x} \in B$ which is a solution of (11). This completes the proof.

COROLLARY. *Let* $\mathbf{f}: R^n \to R^n$ *be a continuous, piecewise-linear mapping and assume that all the unbounded regions have positive Jacobian determinants. Then* $\mathbf{f}$ *maps* R^n *onto itself.*

Proof. This assumption made above implies the norm-coerciveness of $\mathbf{f}$, from which the corollary follows.

Remark 9. This corollary is an extension of Theorem 5 of [6].

The following definition is important for the purpose of finding a good initial point with which one can start the iterative algorithm described in the next section.

DEFINITION 1. Let $\mathbf{f}: R^n \to R^n$ be a continuous, piecewise-linear mapping. A point $\mathbf{u} \in R^n$ is said to satisfy 1-*degree condition* with respect to $\mathbf{f}$ if (i) $\mathbf{u} \notin B$, (ii)

$\mathbf{f}(\mathbf{x}) \neq \mathbf{f}(\mathbf{u})$ for any $\mathbf{x} \neq \mathbf{u}$, and (iii) $\det \mathbf{J} > 0$, where $\mathbf{J}$ is the Jacobian matrix of the region in which $\mathbf{u}$ lies.

Remark 10. If $\mathbf{u}$ satisfies the 1-degree condition, then $\deg(\mathbf{f}, R^n, \mathbf{f}(\mathbf{u})) = 1$. However, the converse, in general, is not true.

The following lemma is needed to derive the solution algorithm for some degenerate cases in the next section.

LEMMA 1. *Let $\mathbf{f}: R^n \to R^n$ be a continuous, norm-coercive, piecewise-linear mapping and $\mathbf{u} \in R^n$ be a point satisfying the 1-degree condition. Then there exists an open neighborhood C of $\mathbf{u}$ such that any point in C satisfies the 1-degree condition.*

Proof. Let $\mathbf{J}^{(i)}$ be the Jacobian matrix of the region R_i in which $\mathbf{u}$ lies. Then the definition of the 1-degree condition implies that $\mathbf{f}(\mathbf{u}) \notin \mathbf{f}(B)$ and that $\det \mathbf{J}^{(i)} > 0$. Thus there exists an open neighborhood U of $\mathbf{u}$ and an open neighborhood V of $\mathbf{f}(\mathbf{u})$ such that $U \subset R_i$, $V \cap \mathbf{f}(B) = \varnothing$ and $\mathbf{f}$ is a homeomorphism of U onto V. Let $W = \mathbf{f}^{-1}(V)$; then W is an open bounded set due to the norm-coerciveness of $\mathbf{f}$. Let $\{\mathbf{y}^{(m)}\} \subset V$ be any sequence which converges to $\mathbf{f}(\mathbf{u})$. Then for each $\mathbf{y}^{(m)}$, there exists a unique point $\mathbf{x}^{(m)} \in U$ such that $\mathbf{f}(\mathbf{x}^{(m)}) = \mathbf{y}^{(m)}$. Suppose, for each $m = 1, 2, \cdots$, $\mathbf{f}^{-1}(\mathbf{y}^{(m)})$ contains a point $\mathbf{z}^{(m)} \neq \mathbf{x}^{(m)}$. Then the sequence $\{\mathbf{z}^{(m)}\}$ is wholly contained in the compact set $\bar{W} - U$, and hence there exists a point of accumulation $\mathbf{z} \in \bar{W} - U$ for which $\mathbf{f}(\mathbf{z}) = \lim \mathbf{f}(\mathbf{z}^{(m)}) = \mathbf{f}(\mathbf{u})$. This contradicts the definition of the 1-degree condition. Therefore there exists an open neighborhood $C \subset U$ of $\mathbf{u}$ and an open neighborhood $D \subset V$ of $\mathbf{f}(\mathbf{u})$ such that $\mathbf{f}$ is a homeomorphism of C onto D and that, for any $\mathbf{y} \in D$, no solution of $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ exists in $R^n - C$. This completes the proof.

The following theorem is an existence theorem on which the solution algorithm described in the next section is based.

THEOREM 5. *Let $\mathbf{f}: R^n \to R^n$ be a continuous, norm-coercive, piecewise-linear mapping and assume that there exists a point $\mathbf{u} \in R^n$ satisfying the 1-degree condition. Then $\mathbf{f}$ maps R^n onto itself.*

This theorem follows from Theorem 3 and the Kronecker theorem by noting that $\deg(\mathbf{f}, R^n, \mathbf{f}(\mathbf{u})) = 1$ and $\mathbf{f}(\mathbf{u}) \notin \mathbf{f}(B)$.

Remark 11. The existence of solutions is guaranteed even if the 1-degree condition is replaced by "there exists a point $\mathbf{u}$ such that $\mathbf{f}(\mathbf{u}) \notin \mathbf{f}(B)$ and $\deg(\mathbf{f}, R^n, \mathbf{f}(\mathbf{u})) \neq 0$". However, as will be seen in § 5, the equations of many kinds of nonlinear networks satisfy the assumption of Theorem 5. Furthermore, the existence of a point satisfying the 1-degree condition is essential to let the algorithm successfully converge to a solution.

COROLLARY. *Let $\mathbf{f}: R^n \to R^n$ be a continuous, piecewise-linear mapping, and assume that* (i) *all the unbounded regions have positive definite Jacobian matrices and* (ii) *all the bounded regions have positive Jacobian determinants. Then $\mathbf{f}$ is a homeomorphism of R^n onto itself.*

Proof. Let r be a positive number such that the ball $S = \{\mathbf{x} \in R^n \mid \|\mathbf{x}\| \leqq r\}$ contains all the bounded regions. Let $M > 0$ be the maximum of the norms of Jacobian matrices of the bounded regions, and $m > 0$ be the minimum of the eigenvalues of symmetric parts of Jacobian matrices of the unbounded regions. It is readily seen that there exists an interior point $\mathbf{u}$ of an unbounded region R_0 such that the distance between $\mathbf{u}$ and ∂R_0 is not less than $3Mr/m$. Suppose there exists a $\mathbf{v} \in R^n$ such that $\mathbf{v} \neq \mathbf{u}$ and $\mathbf{f}(\mathbf{v}) = \mathbf{f}(\mathbf{u})$. Then $\mathbf{v} \notin R_0$. Thus the line segment $L =$

$\{\mathbf{x}\in R^n \mid \mathbf{x}=(1-\lambda)\mathbf{u}+\lambda\mathbf{v},\ \lambda\in[0,1]\}$ intersects ∂R_0 at a point, say $\mathbf{w}$. Let $\theta_1=\langle\mathbf{f}(\mathbf{u})-\mathbf{f}(\mathbf{w}),\mathbf{u}-\mathbf{v}\rangle$ and $\mathbf{J}_0$ be the positive definite Jacobian matrix of region R_0. Then $\theta_1=\langle\mathbf{J}_0(\mathbf{u}-\mathbf{v}),\mathbf{u}-\mathbf{v}\rangle\cdot\|\mathbf{u}-\mathbf{w}\|/\|\mathbf{u}-\mathbf{v}\|\geqq m\|\mathbf{u}-\mathbf{v}\|\cdot\|\mathbf{u}-\mathbf{w}\|\geqq 3Mr\|\mathbf{u}-\mathbf{v}\|$. Since the length of the subset $L'\subset L$ contained in S is less than $2r$, $\theta_2=\langle\mathbf{f}(\mathbf{w})-\mathbf{f}(\mathbf{v}),\mathbf{u}-\mathbf{v}\rangle\geqq -2Mr\|\mathbf{u}-\mathbf{v}\|$. Let $\theta=\langle\mathbf{f}(\mathbf{u})-\mathbf{f}(\mathbf{v}),\mathbf{u}-\mathbf{v}\rangle$; then $\theta=\theta_1+\theta_2\geqq Mr\|\mathbf{u}-\mathbf{v}\|$. Hence $\mathbf{u}\neq\mathbf{v}$ implies $\mathbf{f}(\mathbf{u})\neq\mathbf{f}(\mathbf{v})$, which is a contradiction. Thus it has been shown that $\deg(\mathbf{f},R^n,\mathbf{f}(\mathbf{u}))=1$. Furthermore, the norm-coerciveness of $\mathbf{f}$ implies that, for any $\mathbf{y}\in R^n-\mathbf{f}(B)$, $\deg(\mathbf{f},R^n,\mathbf{y})=1$ due to Theorem 3. Since all the Jacobian determinants are positive, (11) has one and only one solution for any $\mathbf{y}\in R^n-\mathbf{f}(B)$.

It remains to show that for any $\mathbf{y}\in\mathbf{f}(B)$ there exists a unique solution of (11). Suppose (11) has two solutions, say $\mathbf{u}$ and $\mathbf{v}$, for some $\mathbf{y}\in\mathbf{f}(B)$. Then due to [6, Thm. 4], there exists an open neighborhood U (resp., V) of $\mathbf{u}$ (resp., $\mathbf{v}$) and an open neighborhood Y of $\mathbf{y}$ such that $U\cup V=\varnothing$ and $\mathbf{f}$ maps U (resp., V) onto Y. This means that, for a point $\mathbf{y}'\in Y-\mathbf{f}(B)$, $\deg(\mathbf{f},R^n,\mathbf{y}')\geqq 2$, which is a contradiction. Therefore $\mathbf{f}$ is a homeomorphism of R^n onto itself.

4. Solution algorithm. The problem under consideration is to find a solution of

$$\mathbf{f}(\mathbf{x})=\mathbf{y} \tag{11}$$

for a piecewise-linear mapping $\mathbf{f}: R^n\to R^n$ and a given input $\mathbf{y}\in R^n$. Initially, an inner point $\mathbf{x}^{(1)}$ of a region, say R_1, is selected as the starting point. It is assumed that the determinant of the Jacobian matrix $\mathbf{J}^{(1)}$ of region R_1 is positive.[2] Let L_y be the line segment joining $\mathbf{y}^{(1)}=\mathbf{f}(\mathbf{x}^{(1)})$ and $\mathbf{y}$ in the y-space. The problem is then to trace a continuous polygonal curve L_x such that $\mathbf{f}(L_x)=L_y$, starting with $\mathbf{x}^{(1)}$ in the x-space. Thus the other endpoint of L_x is a solution of (11).

The portion of the solution curve L_x which lies in R_1 is indicated by

$$\mathbf{w}^{(1)}(\lambda)=\mathbf{x}^{(1)}+\lambda(\mathbf{J}^{(1)})^{-1}(\mathbf{y}-\mathbf{y}^{(1)}), \tag{20}$$

where $\lambda>0$ is a parameter. If $\mathbf{w}^{(1)}$ (1) happens to be in R_1, then it is the desired solution. Otherwise, the value of $\lambda\in(0,1)$ has to be determined in such a way that $\mathbf{w}^{(1)}(\lambda)$ lies on a boundary hyperplane, say H_1, of R_1. Let λ_1 be such value of λ, and define $\mathbf{x}^{(2)}=\mathbf{w}^{(1)}(\lambda_1)$ and $\mathbf{y}^{(2)}=\mathbf{f}(\mathbf{x}^{(2)})$. The line segment joining $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ is thus the first portion of the solution curve.

The next step is to extend the solution curve beyond $\mathbf{x}^{(2)}$. For simplicity, to describe the essence of the algorithm, it is assumed, for the time being, that $\mathbf{x}^{(2)}$ lies on a single boundary hyperplane which separates R_1 from one and only one neighboring region, say R_2. Then the portion of L_x lying in R_2 is indicated by

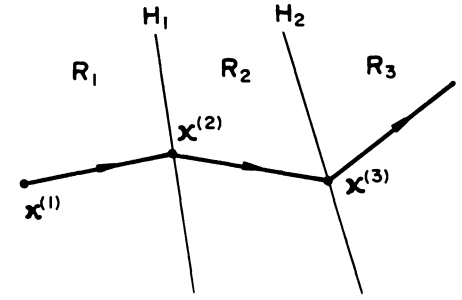
$$\mathbf{w}^{(2)}(\lambda)=\mathbf{x}^{(2)}+\lambda(\mathbf{J}^{(2)})^{-1}(\mathbf{y}-\mathbf{y}^{(2)}) \tag{21}$$

unless $\det\mathbf{J}^{(2)}=0$.

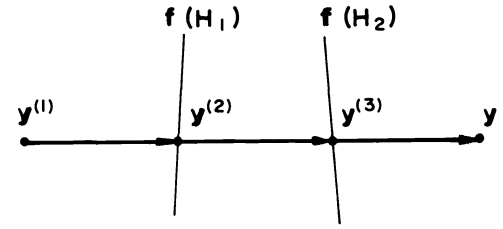
If $\det\mathbf{J}^{(2)}>0$, $\mathbf{w}^{(2)}(\lambda)$ lies in R_2 for $\lambda\geqq 0$. As in the initial step, if $\mathbf{w}^{(2)}(1)$ is a solution, the algorithm terminates here. Otherwise, the value of $\lambda\in(0,1)$ is

[2] As will be seen in what follows, the starting point is selected so as to satisfy the 1-degree condition, which implies that the region has a positive Jacobian determinant.

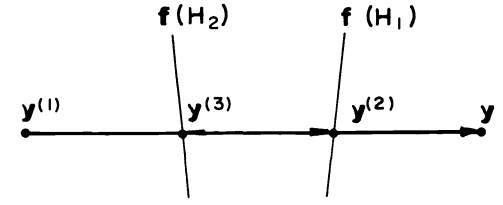
determined in such a way that $\mathbf{w}^{(2)}(\lambda)$ lies on another boundary hyperplane, say H_2, of R_2. Let λ_2 be such a value of λ, and define $\mathbf{x}^{(3)} = \mathbf{w}^{(2)}(\lambda_2)$ and $\mathbf{y}^{(3)} = \mathbf{f}(\mathbf{x}^{(3)})$. The line segment joining $\mathbf{x}^{(2)}$ and $\mathbf{x}^{(3)}$ is thus the second portion of the solution curve. Continuing this way, we extend solution curve as shown in Fig. 4(a). Note that, in this case, the sequence $\{\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \mathbf{y}^{(3)}\}$ approaches $\mathbf{y}$ monotonically as shown in Fig. 4(b). Generally speaking, the algorithm described here is no different from the original Katzenelson algorithm [1] as long as the solution curve traverses only regions with positive Jacobian determinants.



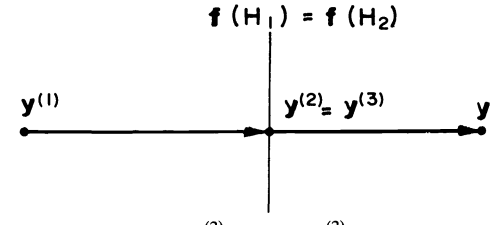
(a) L_x: $\det \mathbf{J}^{(1)} > 0$



(b) L_y: $\det \mathbf{J}^{(2)} > 0$



(c) L_y: $\det \mathbf{J}^{(2)} < 0$



(d) L_y: $\det \mathbf{J}^{(2)} = 0$, $\operatorname{rank} \mathbf{J}^{(2)} = n - 1$

FIG. 4. *Construction of the solution curve*

If $\det \mathbf{J}^{(2)} < 0$, $\mathbf{w}^{(2)}(\lambda)$ lies in R_2 for $\lambda \leqq 0$. In this case, too, the value of $\lambda \in (-\infty, 0)$ is determined in such a way that $\mathbf{w}^{(2)}(\lambda)$ lies on another boundary hyperplane, say H_2, of R_2. Suppose such a value, say λ_2, of λ exists. Then the line segment joining $\mathbf{x}^{(2)}$ and $\mathbf{x}^{(3)} = \mathbf{w}^{(2)}(\lambda_2)$ is regarded as the second portion of L_x. It should be noted here that the vector $\mathbf{y}^{(3)} - \mathbf{y}^{(2)}$; $\mathbf{y}^{(3)} = \mathbf{f}(\mathbf{x}^{(3)})$ is directed away from $\mathbf{y}$ as shown in Fig. 4(c) if $\det \mathbf{J}^{(2)} < 0$. In other words, the solution curve is temporarily extended away from the solution.

Even if $\det \mathbf{J}^{(2)} = 0$, the solution curve can be extended into R_2. Since $\det \mathbf{J}^{(1)} > 0$, it follows from Proposition 2 that $\mathbf{J}^{(2)}$ is of rank $n-1$. Therefore the nonzero vector $\boldsymbol{\beta}$ such that

$$\mathbf{J}^{(2)}\boldsymbol{\beta} = \mathbf{0} \tag{22}$$

is uniquely determined within constant multipliers, and the portion of L_x lying in R_2 is indicated by

$$\mathbf{w}^{(2)}(\lambda) = x^{(2)} + \lambda \boldsymbol{\beta}. \tag{23}$$

Then the value of $\lambda \in (-\infty, \infty)$ is determined so that $\mathbf{w}^{(2)}(\lambda)$ lies on another boundary hyperplane, say H_2, of R_2. If such a value λ_2 is found, the line segment joining $\mathbf{x}^{(2)}$ and $\mathbf{x}^{(3)} = \mathbf{w}^{(2)}(\lambda_2)$ constitutes the second portion of L_x. In this case, the whole line segment is mapped into the single point $\mathbf{y}^{(2)}$; i.e., the whole region R_2 is mapped into hyperplane $\mathbf{f}(H_1)$ as shown in Fig. 4(d). It should be noted here that $\mathbf{w}^{(2)}(\lambda)$ in (21) or (23) never runs along hyperplane H_1 for $\lambda \neq 0$. If this were the case, then $\mathbf{x}^{(1)} \notin H_1$ and two or more points on H_1 would be mapped on L_y, which would contradict the nonsingularity of $\mathbf{J}^{(1)}$.

Next, consider the special case where the next region R_3 also has a singular Jacobian matrix $\mathbf{J}^{(3)}$. If $\mathbf{J}^{(2)}$ is of rank $n-1$, it follows from Proposition 2 that the rank of $\mathbf{J}^{(3)}$ is $n-2$ or more. Suppose $\mathbf{J}^{(3)}$ is of rank $n-2$. Then there exists a two-dimensional plane H^*, containing $\mathbf{x}^{(3)}$, such that $\mathbf{f}(H^* \cap R_3) = \{\mathbf{y}^{(3)}\}$, i.e., $\mathbf{f}(H^* \cap H_2) = \{\mathbf{y}^{(3)}\}$. As long as H_2 is the only boundary hyperplane on which $\mathbf{x}^{(3)}$ lies, the hyperplane $H^* \cap H_2$ containing $\mathbf{x}^{(3)}$ is at least one-dimensional. Hence let $\mathbf{x}^* \neq \mathbf{x}^{(3)}$ be a point of $H^* \cap H_2$; then $\mathbf{f}(\mathbf{x}^*) = \mathbf{f}(\mathbf{x}^{(3)})$ and, for the two linearly-independent vectors $\mathbf{x}^* - \mathbf{x}^{(2)}$ and $\mathbf{x}^{(3)} - \mathbf{x}^{(2)}$, $\mathbf{J}^{(2)}(\mathbf{x}^* - \mathbf{x}^{(2)}) = \mathbf{J}^{(2)}(\mathbf{x}^{(3)} - \mathbf{x}^{(2)})$. Thus the rank of $\mathbf{J}^{(2)}$ is at most $n-2$, which is a contradiction. Therefore the rank of $\mathbf{J}^{(3)}$ must be $n-1$. Furthermore, the extension of the solution curve beyond $\mathbf{x}^{(3)}$, in this case, never runs along H_2, because, if this were the case, then the rank of $\mathbf{J}^{(2)}$ would be just as before, less than $n-1$, which would be a contradiction. Since the above argument is valid for any iteration step of the algorithm, one obtains the following key property of solution curves.

LEMMA 2. *Let L_x be a solution curve starting from an interior point of a region with nonsingular Jacobian matrix, and assume that L_x crosses only one boundary hyperplane at a time. Then* (i) *L_x never enters a region with Jacobian matrix of rank $n-2$ or less and* (ii) *for any boundary hyperplane H which L_x crosses, L_x intersects H at a single point.*

It has been shown that the solution curve can be extended into a new region whenever it crosses a single boundary hyperplane. When the solution curve hits two or more boundary hyperplanes simultaneously, the way to extend it into a new region is confirmed by means of Theorem 2.

Without loss of generality, the solution curve is assumed to hit a corner[3] at $\mathbf{x}^{(2)}$ in the first iteration step. Let $X(\mathbf{f}, \mathbf{x})$ be the maximal connected subset of $\mathbf{f}^{-1}(\mathbf{f}(\mathbf{x}))$ containing $\mathbf{x}$. If $\mathbf{f}$ is norm-coercive, then $X(\mathbf{f}, \mathbf{x})$ is a compact set. Since the line segment joining $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ is a portion of L_x, it is clear that $\mathbf{f}(\mathbf{x}^{(2)}+\nu(\mathbf{x}^{(1)}-\mathbf{x}^{(2)}))\in L_y$ for all $\nu\in[0, 1]$. Therefore it follows from Theorem 2 that the solution curve has at least one and, in general, an odd number of extensions outside $X(\mathbf{f}, \mathbf{x}^{(2)})$, excluding that lying in R_1. Note that if L_y runs along $\mathbf{f}(B)$ in the neighborhood of $\mathbf{y}^{(2)}$, the number of extensions of the solution curve is counted in the sense of perturbation as described in Remark 8. Thus each extension can be considered as lying in a region with nonsingular Jacobian matrix. Let R_2 be the region in which one such extension lies and $\mathbf{J}^{(2)}$ be the Jacobian matrix of R_2. Then due to Theorem 2,

$$\mathbf{w}^{(2)}(\lambda)=\mathbf{v}^{(2)}+\lambda(\mathbf{J}^{(2)})^{-1}(\mathbf{y}-\mathbf{y}^{(2)}) \tag{24}$$

indicates the portion of L_x lying in R_2, where

$$\mathbf{v}^{(2)}\in R_2\cap\partial X(\mathbf{f}, \mathbf{x}^{(2)}) \tag{25}$$

is the point from which the solution curve is extended.[4]

When the solution curve comes back to $X(\mathbf{f}, \mathbf{x}^{(2)})$ again, it must be extended into a region which has not been traversed. This is always possible since L_x has an even number of branches incident with each corner and two of them are traversed whenever the solution curve passes it.

From the above arguments, it is clear that the solution curve can always be extended into a new region regardless of whether it hits a corner. Furthermore, the piecewise-linearity of $\mathbf{f}$ implies that no two line segments in a region possessing a nonsingular Jacobian matrix can be portions of a solution curve. Since there exist at most a finite number of regions, a solution can be found in a finite number of iterations unless one of the following unfavorable phenomena occurs.

Case A. The solution curve, in some region, might be extended infinitely without crossing any other boundary hyperplane and without obtaining a solution.

Case B. The solution curve might reenter region R_1 in which starting point $\mathbf{x}^{(1)}$ lies without obtaining a solution.

In the previous section, it has been shown in an unconstructive way that (11) has at least one solution for any $\mathbf{y}$ if $\mathbf{f}$ satisfies the condition of Theorem 5. Now it shall be shown in what follows that, under this condition, the algorithm described above always leads to a solution; i.e., one never encounters Case A or Case B.

As long as $\mathbf{f}$ is norm-coercive, Case A never occurs in any region with a singular Jacobian matrix. Suppose the solution curve enters region R_m at $\mathbf{x}^{(m)}$ and it is extended infinitely in R_m. Since no solution has been obtained, it is clear that $\langle\mathbf{y}-\mathbf{y}^{(1)}, \mathbf{y}-\mathbf{y}^{(m)}\rangle>0$. If $\mathbf{y}^{(m)}$ is located on the extension of L_y as shown in Fig. 5(b), it means that the solution curve has passed a solution, say $\hat{\mathbf{x}}^{(1)}$ of

$$\mathbf{f}(\mathbf{x})=\mathbf{y}^{(1)} \tag{26}$$

---

[3] A *corner* is a closed, connected set determined by the intersection of at least two boundary hyperplanes, or determined by the union of regions with Jacobian matrices of rank at most $n-2$.

[4] In many cases, $X(\mathbf{f}, \mathbf{x}^{(2)})$ consists of the single point $\mathbf{x}^{(2)}=\mathbf{v}^{(2)}$. In general, $\mathbf{v}^{(2)}$ and $\mathbf{x}^{(2)}$ are vertices of the complex $X(\mathbf{f}, \mathbf{x}^{(2)})$.

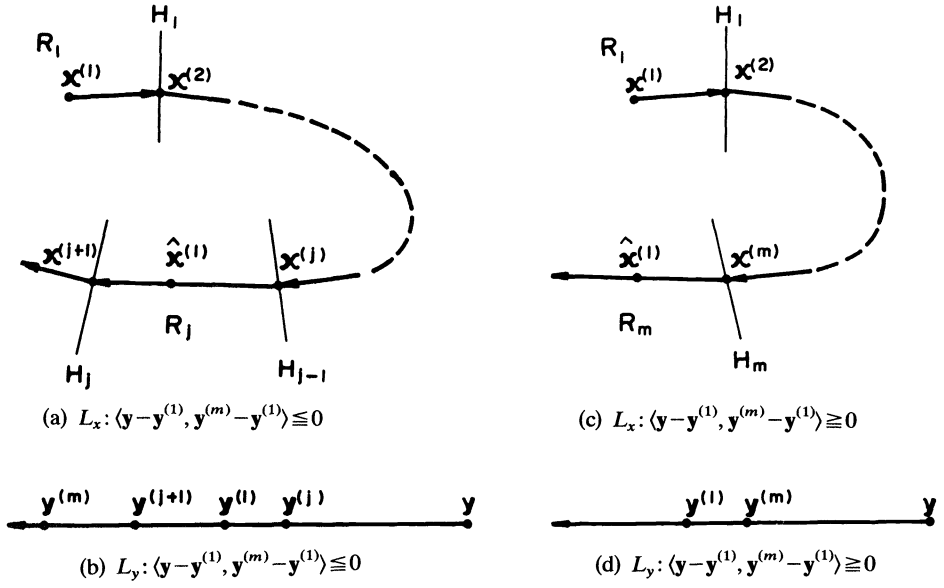


FIG. 5. *Illustration of Case* A

in some region R_j, $2 \leqq j \leqq m$, as shown in Fig. 5(a). If $\mathbf{y}^{(m)}$ is located on L_y as shown in Fig. 5(d), it means that the extension of the solution curve passes a solution, say $\hat{\mathbf{x}}^{(1)}$, of (26) in region R_m as shown in Fig. 5(c).

Next Case B shall be analyzed. Suppose the solution curve reenters region R_1 at $\mathbf{x}^{(m)}$ as shown in Fig. 6(a). Then $\mathbf{y}^{(m)}$ should be located on the extension of L_y as shown in Fig. 6(b), since $\mathbf{y}^{(2)}$ lies on L_y. It means that the solution curve has passed a solution, say $\hat{\mathbf{x}}^{(1)}$, of (26) in some region R_j; $2 \leqq j \leqq m$.

Now the above arguments lead to a strategy to get rid of Cases A and B; that is, to impose the 1-degree condition on the starting point $\mathbf{x}^{(1)}$. Under this assumption, $\mathbf{x}^{(1)}$ is the unique solution of (26), which excludes the occurrence of Cases A and B. In summary, the following theorem has been proved.

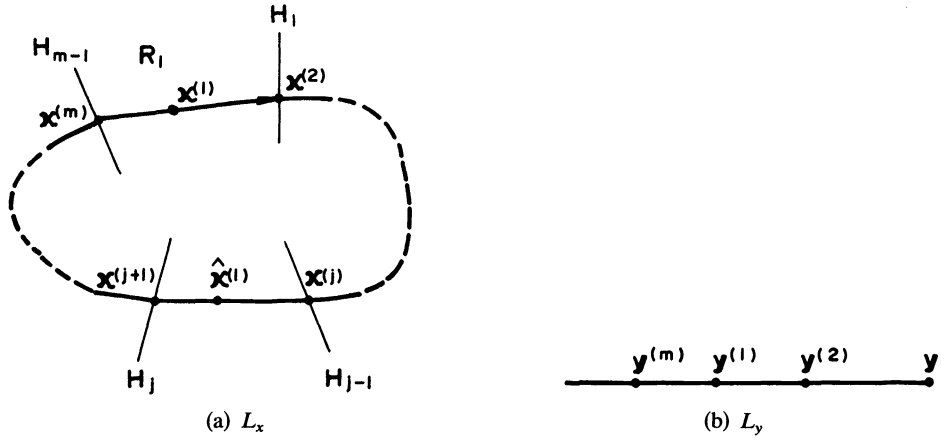


FIG. 6. *Illustration of Case* B

THEOREM 6. *Under the same condition as Theorem* 5, *the generalized Katzenelson algorithm, starting from* $\mathbf{u}$, *converges to a solution of* (11) *in a finite number of iteration steps for any* $\mathbf{y} \in R^n$.

There remains a computational problem to be investigated; the *corner problem.* Whenever the solution curve hits a corner, it can theoretically be extended into a new region. The problem from a computational point of view, however, is how to find such a new region.

A corner includes no interior point of the regions with Jacobian matrices of rank $n-1$ or n. Thus let Γ be the set of all points lying in corners; then $\mathbf{f}(\Gamma)$ is the finite set union of convex polyhydral sets $K_1, K_2, \cdots, K_q$, each of which lies in an $(n-2)$-dimensional hyperplane in the y-space. Let $\mathbf{x}^{(1)}$ be the starting point satisfying the 1-degree condition and $\mathbf{y}^{(1)} = \mathbf{f}(\mathbf{z}^{(1)})$; then $\mathbf{y}^{(1)} \notin \mathbf{f}(B)$ and, obviously, $\mathbf{f}(\Gamma) \subset \mathbf{f}(B)$. Let $\mathbf{y}$ be the given input and L_y be the line segment joining $\mathbf{y}^{(1)}$ and $\mathbf{y}$. It is clear that the solution curve never hits a corner if L_y does not meet $\mathbf{f}(\Gamma)$ except at $\mathbf{y}$. Due to Lemma 1, there exists a $\delta > 0$ such that any point of the open ball C with center $\mathbf{x}^{(1)}$ and radius δ satisfies the 1-degree condition and C is wholly contained in region R_1 in which $\mathbf{x}^{(1)}$ lies. Since the Jacobian matrix $\mathbf{J}^{(1)}$ of R_1 is nonsingular, there exists an $\varepsilon > 0$ such that $\|\mathbf{z} - \mathbf{y}^{(1)}\| < \varepsilon$ implies $\mathbf{z} = \mathbf{f}(\mathbf{w})$ for some $\mathbf{w} \in C$. In what follows, it shall be shown that there exists a point $\hat{\mathbf{x}}^{(1)} \in C$ such that the line segment joining $\hat{\mathbf{y}}^{(1)} = \mathbf{f}(\hat{\mathbf{x}}^{(1)})$ and $\mathbf{y}$ never meets $\mathbf{f}(\Gamma)$ except at $\mathbf{y}$.

Suppose L_y meets $K_1 \subset \mathbf{f}(\Gamma)$ at a point $\mathbf{y}^* = \mathbf{y}^{(1)} + \lambda^*(\mathbf{y} - \mathbf{y}^{(1)})$; $\lambda^* \in (0, 1)$ and let H^* be an $(n-2)$-dimensional hyperplane including K_1. First, consider the case where $L_y \cap H^*$ is the single point $\mathbf{y}^*$. Let $\boldsymbol{\alpha}^{(i)}$, $i = 1, 2, \cdots, n-2$, be a basis of the vector space $\{\mathbf{v} \in R^n | \mathbf{v} = \mathbf{w} - \mathbf{y}^*, \mathbf{w} \in H^*\}$. Then $\mathbf{y}^{(1)} - \mathbf{y}$ and another vector $\boldsymbol{\alpha}^{(n)}$ can be added so that $\boldsymbol{\alpha}^{(1)}, \cdots, \boldsymbol{\alpha}^{(n-2)}, \mathbf{y}^{(1)} - \mathbf{y}, \boldsymbol{\alpha}^{(n)}$ form a basis of the n-dimensional vector space R^n. Let μ be a sufficiently small positive number. Then it is easy to show that the line segment joining $\mathbf{y}$ and $\mathbf{y}^{(1)} = \mathbf{y}^{(1)} + \mu\boldsymbol{\alpha}^{(n)}$ does not meet H^* except at y. Secondly, consider the case where a portion of L_y with a positive length is contained in H^*. Let $\boldsymbol{\alpha}^{(1)}, \cdots, \boldsymbol{\alpha}^{(n-3)}, \mathbf{y}^{(1)} - \mathbf{y}$ be a basis of the vector space $\{\mathbf{v} \in R^n | \mathbf{v} = \mathbf{w} - \mathbf{y}^*, \mathbf{w} \in H^*\}$. Then two vectors $\boldsymbol{\alpha}^{(n-1)}, \boldsymbol{\alpha}^{(n)}$ can be added so that $\boldsymbol{\alpha}^{(1)}, \cdots, \boldsymbol{\alpha}^{(n-3)}, \mathbf{y}^{(1)} - \mathbf{y}, \boldsymbol{\alpha}^{(n-1)}, \boldsymbol{\alpha}^{(n)}$ form a basis of R^n. Again, $\hat{\mathbf{y}}^{(1)} = \mathbf{y}^{(1)} + \mu\boldsymbol{\alpha}^{(n)}$ does not meet H^* except at $\mathbf{y}$ for sufficiently small $\mu > 0$. Since this modification can be made small enough to guarantee $\|\hat{\mathbf{y}}^{(1)} - \mathbf{y}^{(1)}\| < \varepsilon$, one can repeat this process so as to remove L_y, except at $\mathbf{y}$, away from $K_1, \cdots, K_q$.

Remark 12. With this modification of the starting point, the solution curve does not hit a corner except at a solution. This implies, due to Lemma 2, that the modified solution curve never traverses along a boundary hyperplane.

In numerical computation, this modification cannot be determined a priori. There are two practical methods to perturb the solution curve as described in the following two paragraphs.

The first method is the following. If the solution curve hits a corner, one selects a new starting point, which is close enough to the solution curve and is an interior point of the region just traversed, so that a new solution curve does not hit the corner. This modification has to be small enough so that the solution curve can be included in another solution curve starting with a point in a neighborhood of $x^{(1)}$. The difficulty in this method is to judge what is small enough.

The second method is based on the following principle. What is necessary to do, when the solution curve hits a corner, is to determine a region into which it can be extended. Let $\mathbf{y}$, $\mathbf{y}^{(1)}$, μ and $\boldsymbol{\alpha}^{(n)}$ be as before. Then it is easily seen that the line segment joining $\mathbf{y}+\mu\boldsymbol{\alpha}^{(n)}$ and $\mathbf{y}^{(1)}+\mu\boldsymbol{\alpha}^{(n)}$ does not meet the image of the corner. Thus one selects a new starting point, which is an interior point of the region just traversed, so that the new solution curve does not hit the corner. It should be noted that the new solution curve is parallel with the old one in any region. Once the new solution curve moves away from the boundary of the corner, the perturbation is removed; i.e., the original solution curve is again traced. This is illustrated in the examples in this section.

Example 2. Figure 7(a) illustrates a twin tunnel diode circuit. Its input-output characteristics are governed by the two-dimensional equation

$$\begin{aligned} g(x_1)+G(x_1+x_2)&=y_1,\\ g(x_2)+G(x_1+x_2)&=y_2, \end{aligned} \tag{27}$$

where $\mathbf{y}=(y_1, y_2)^{\mathrm{T}}$ represents the values of given current sources and $\mathbf{x}=(x_1, x_2)^{\mathrm{T}}$ represents the voltages across the diodes both characterized by Fig. 7(b). Then the x-space is divided into nine regions as shown in Fig. 8(a), where Jacobian determinants in regions R_1, R_3, R_5, R_6 and R_8 are positive and those in R_2, R_4, R_7 and R_9 are negative.

Let the value of given input be $\mathbf{y}=(69, 75)^{\mathrm{T}}$ and the starting point be $\mathbf{x}^{(1)}=(0, 1)^{\mathrm{T}}\in R_1$. Note that $\mathbf{x}^{(1)}$ satsifies the 1-degree condition. Calculate $\mathbf{y}^{(1)}=\mathbf{f}(\mathbf{x}^{(1)})=(1, 7)^{\mathrm{T}}$ and consider the line segment L_y joining $\mathbf{y}^{(1)}$ and $\mathbf{y}$. Then the inverse image $\mathbf{f}^{-1}(L_y)$ consists of an open polygonal curve $L_x^{(a)}$ and a closed polygonal curve $L_x^{(b)}$ as shown in Fig. 8(a). The solution curve starting from $\mathbf{x}^{(1)}$ traverses through $L_x^{(a)}$ and finds the solution $\mathbf{x}=(16, 17)^{\mathrm{T}}\in R_5$ of (27) in five steps. Figure 8(b) illustrates how the sequence of approximate solutions approaches the solution in the y-space.

Next replace the starting point by $\mathbf{x}^{(6)}=(13, 4)^{\mathrm{T}}\in R_6$ with the same input, keeping $\mathbf{y}^{(6)}=\mathbf{f}(\mathbf{x}^{(6)})=(35, 41)^{\mathrm{T}}$ on L_y. Note that $\mathbf{x}^{(6)}$ does not satisfy the 1-degree condition. The line segment L_y' joining $\mathbf{y}^{(6)}$ and $\mathbf{y}$ is a proper subset of L_y, i.e., $\|\mathbf{y}^{(6)}-\mathbf{y}\|<\|\mathbf{y}^{(1)}-\mathbf{y}\|$. However, the solution curve starting from $\mathbf{x}^{(6)}$ traverses through $L_x^{(b)}$ and comes back to region R_6 without finding a solution. Figure 8(c) illustrates the behavior of the solution curve in the y-space.

Example 3. In the same circuit as in Example 1, let the given input be $\mathbf{y}=(68, 68)^{\mathrm{T}}$ and the starting point be $\mathbf{x}^{(1)}=(0, 0)^{\mathrm{T}}\in R_1$; $\mathbf{y}^{(1)}=(0, 0)^{\mathrm{T}}$. Then the solution curve starting from $\mathbf{x}^{(1)}$ hits a corner at $\mathbf{x}^2=(6, 6)^{\mathrm{T}}$. In the neighborhood of $\mathbf{x}^{(2)}$, the inverse image of the line segment L_y joining $\mathbf{y}^{(1)}$ and $\mathbf{y}$ has four branches as shown in Fig. 9(a). By means of the perturbation technique (the second method), it is seen that the solution curve can be extended into region R_2 or R_9. Independent of the region to which the solution curve is extended, it finally reaches the solution $\mathbf{x}=(16, 16)^{\mathrm{T}}\in R_5$.

$$\begin{aligned} g(x_1)&=y_1,\\ g(x_2)&=y_2, \end{aligned} \tag{28}$$

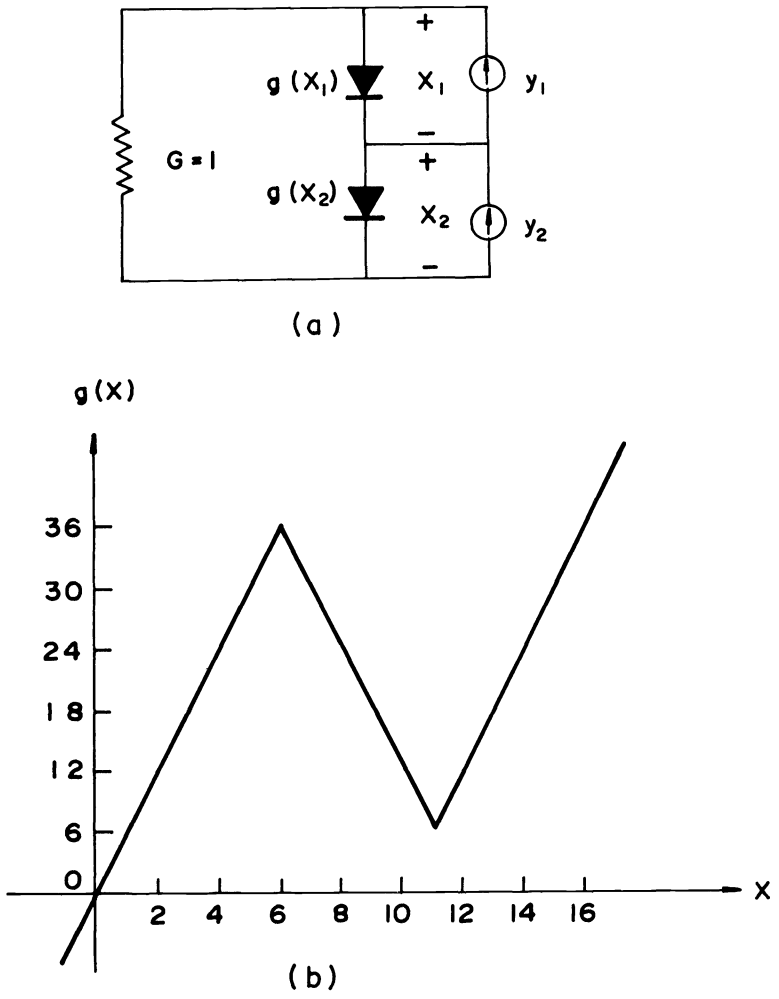


FIG. 7. *Twin tunnel diode circuit*

where

$$g(x)=\begin{cases} x, & x\in(-\infty, 1], \\ 1, & x\in[1, 2], \\ x-1, & x\in[2, \infty). \end{cases} \tag{29}$$

The x-space is divided into nine regions as shown in Fig. 9(b), where Jacobian determinants in regions R_1, R_3, R_5 and R_6 are positive and those in the other regions are all zero. More precisely, regions R_2, R_4, R_7 and R_9 are of rank 1 and region R_8 is of rank 0.

Let the value of a given input be $\mathbf{y}=(2, 2)^{\mathrm{T}}$ and the starting point be $\mathbf{x}^{(1)}=(0, 0)^{\mathrm{T}}\in R_1$. Note that $\mathbf{x}^{(1)}$ satisfies the 1-degree condition. Let L_y be the line segment joining $\mathbf{y}^{(1)}=(0, 0)^{\mathrm{T}}$ and $\mathbf{y}$. Then the inverse $\mathbf{f}^{-1}(L_y)$ includes the whole

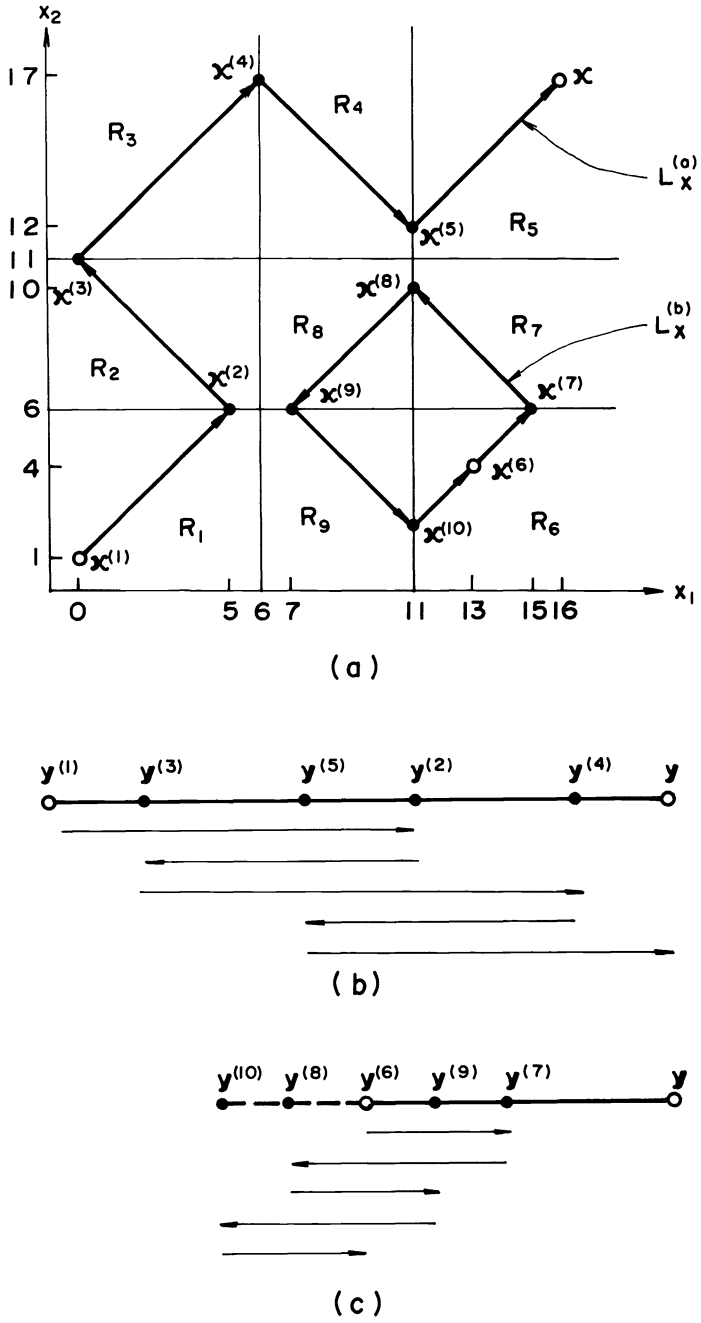


FIG. 8. *Illustration of the solution curve*

region R_8. Even in this case, the solution can be obtained as follows. The solution curve starting from $\mathbf{x}^{(1)}$ hits a corner at $\mathbf{x}^{(2)} = (1, 1)^T$. Then by means of the perturbation technique, it is extended through the boundary between R_8 and R_9.

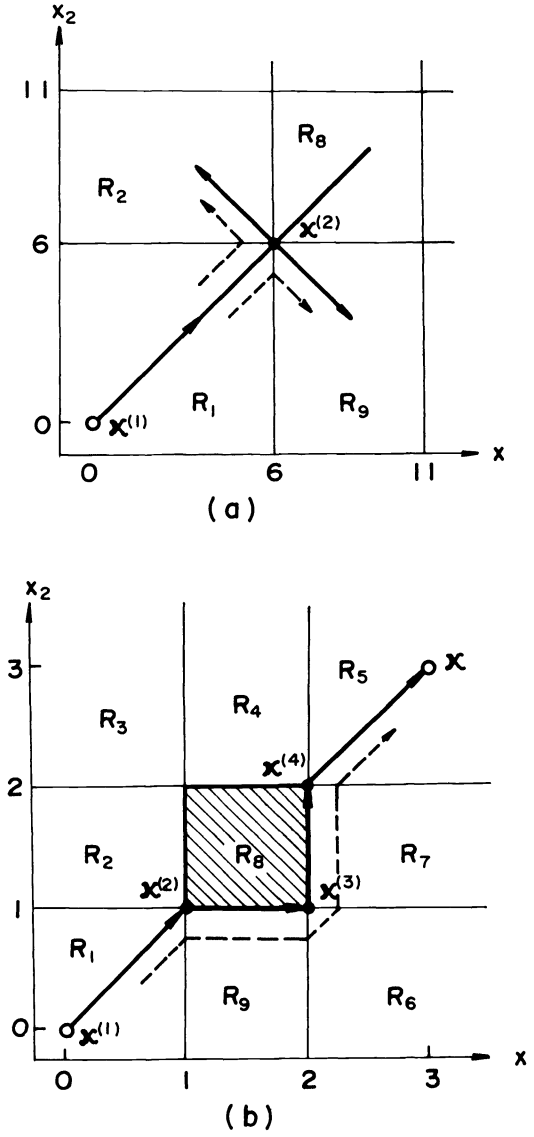


FIG. 9. *Solution curve hitting a corner*

Finally, it reaches the solution $\mathbf{x} = (3, 3)^T \in R_5$ of (9), passing through vertices $\mathbf{x}^{(2)}$, $\mathbf{x}^{(3)}$ and $\mathbf{x}^{(4)}$ of the square R_8.

5. Resistor networks.

5.1. Branch characteristics. In this paper, a *resistor* is regarded, in general, as a multiport element characterized by

$$\mathbf{z}^{(k)} = \mathbf{h}^{(k)}(\mathbf{x}^{(k)}), \tag{30}$$

where $\mathbf{h}^{(k)}$ is a continuous mapping from n_k-dimensional Euclidean space R^{n_k}

into itself. The underlying restriction imposed on the expression (30) is that when an entry, say $x_i^{(k)}$, of $\mathbf{x}^{(k)}$ is the voltage (resp., current), then the corresponding entry $z_i^{(k)}$ of $\mathbf{z}^{(k)}$ is the current (resp., voltage). According to the dimension n_k of $\mathbf{h}^{(k)}$, the resistor is called an *n_k-port resistor.*

The branch constitutive relation for all the resistors, say r resistors, is represented by

$$\mathbf{z} = \mathbf{h}(\mathbf{x}), \tag{31}$$

where $\mathbf{h}$: $R^n \to R^n$ is a block-diagonal mapping such that each block characterizes a resistor. In other words, $\mathbf{h}$ is the direct sum of resistor characteristics, i.e.,

$$\mathbf{h}(\mathbf{x}) = \mathbf{h}^{(1)}(\mathbf{x}^{(1)}) \dot{+} \mathbf{h}^{(2)}(\mathbf{x}^{(2)}) \dot{+} \cdots \dot{+} \mathbf{h}^{(r)}(\mathbf{x}^{(r)}). \tag{32}$$

The dimension n is equal to the total number of resistor ports.

5.2. Network topology. Consider the graph associated with a resistor network such that each edge corresponds to either a resistor port or an independent source. The graph is assumed to be connected, without loss of generality. Choose a tree which contains all the voltage sources, the maximum number of voltage controlled resistor ports and no current source. Based on the tree, subscript t (resp., l) is attached to represent the variable associated with a tree branch (resp., link). According to the hybrid characterization (31), subscript v (resp., c) is attached to represent the variable associated with a voltage controlled (resp., current controlled) resistor port. In this way, the set of resistor ports is partitioned into four disjoint subsets, and the Kirchhoff's laws are thus expressed by [13], [14]

$$\begin{aligned}
\mathbf{v}_{vl} \quad &+ \mathbf{F}_{vv}\mathbf{v}_{vt} \quad = \mathbf{e}_v, \\
\mathbf{i}_{ct} + \quad &- \mathbf{F}_{cc}^{\mathrm{T}}\mathbf{i}_{cl} = \mathbf{j}_c, \\
-\mathbf{F}_{vv}^{\mathrm{T}}\mathbf{i}_{vl} \quad &+ \quad \mathbf{i}_{vt} - \mathbf{F}_{cv}^{\mathrm{T}}\mathbf{i}_{cl} = \mathbf{j}_v, \\
\mathbf{F}_{cc}\mathbf{v}_{ct} + \mathbf{F}_{cv}\mathbf{v}_{vt} + \quad &\mathbf{v}_{cl} = \mathbf{e}_c,
\end{aligned} \tag{33}$$

where

$$\begin{array}{c} \\ vl \\ cl \end{array} \begin{array}{c} \begin{array}{cccc} vl & ct & vt & cl \end{array} \\ \left[\begin{array}{c|c|c|c} \mathbf{I} & \mathbf{0} & \mathbf{F}_{vv} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{F}_{cc} & \mathbf{F}_{cv} & \mathbf{I} \end{array}\right] \end{array}$$

is the fundamental loop matrix with respect to the tree, and $\mathbf{e}_v$ and $\mathbf{e}_c$ (resp., $\mathbf{j}_c$ and $\mathbf{j}_v$) stand for the voltage source (resp., current source) vectors around fundamental loops (resp., across fundamental cutsets) determined by the links (resp., tree branches).

To simplify the presentation, let

$$\text{(34)}\qquad \mathbf{x}=\begin{bmatrix}\mathbf{v}_{vl}\\ \hline \mathbf{i}_{ct}\\ \hline \mathbf{v}_{vt}\\ \hline \mathbf{i}_{cl}\end{bmatrix}\in R^n,\qquad \mathbf{y}=\begin{bmatrix}\mathbf{e}_{v}\\ \hline \mathbf{j}_{c}\\ \hline \mathbf{j}_{v}\\ \hline \mathbf{e}_{c}\end{bmatrix}\in R^n,\qquad \mathbf{z}=\begin{bmatrix}\mathbf{i}_{vl}\\ \hline \mathbf{v}_{ct}\\ \hline \mathbf{i}_{vt}\\ \hline \mathbf{v}_{cl}\end{bmatrix}\in R^n.$$

Then (33) can be written in the form,

$$\text{(35)}\qquad \mathbf{Pz}+\mathbf{Qx}=\mathbf{y},$$

where

$$\text{(36)}\qquad \mathbf{P}=\left[\begin{array}{c|c}\mathbf{0} & \mathbf{0}\\ \hline \mathbf{A}^{\mathrm{T}} & \mathbf{I}\end{array}\right],\qquad \mathbf{Q}=\left[\begin{array}{c|c}\mathbf{I} & -\mathbf{A}\\ \hline \mathbf{0} & \mathbf{B}\end{array}\right],$$

$$\mathbf{A}=\begin{array}{c}\\ vl\\ ct\end{array}\begin{array}{c}\begin{array}{cc}vt & cl\end{array}\\ \left[\begin{array}{c|c}-\mathbf{F}_{vv} & \mathbf{0}\\ \hline \mathbf{0} & \mathbf{F}_{cc}^{\mathrm{T}}\end{array}\right]\end{array},\qquad \mathbf{B}=\begin{array}{c}\\ vt\\ cl\end{array}\begin{array}{c}\begin{array}{cc}vt & cl\end{array}\\ \left[\begin{array}{c|c}\mathbf{0} & -\mathbf{F}_{cv}^{\mathrm{T}}\\ \hline \mathbf{F}_{cv} & \mathbf{0}\end{array}\right]\end{array}.$$

The pair $(\mathbf{P}, \mathbf{Q})$ of $n\times n$ square matrices represented by (36), with $\mathbf{B}$ being skew-symmetric, possesses the following properties.

LEMMA 3. $\mathbf{Pz}+\mathbf{Qx}=\mathbf{0}$ *implies that* $\langle \mathbf{z}, \mathbf{x}\rangle = 0$.

Proof. According to (35), partition $\mathbf{z}$ and $\mathbf{x}$ as

$$\mathbf{z}=\begin{bmatrix}\mathbf{z}^{(1)}\\ \hline \mathbf{z}^{(2)}\end{bmatrix},\qquad \mathbf{x}=\begin{bmatrix}\mathbf{x}^{(1)}\\ \hline \mathbf{x}^{(2)}\end{bmatrix}.$$

Then $\mathbf{Pz}+\mathbf{Qx}=\mathbf{0}$ implies that

$$\mathbf{x}^{(1)}=\mathbf{A}\mathbf{x}^{(2)}$$

and

$$\mathbf{z}^{(2)}=-\mathbf{A}^{\mathrm{T}}\mathbf{z}^{(1)}-\mathbf{B}\mathbf{x}^{(2)}.$$

Since $\mathbf{B}$ is a skew-symmetric matrix,

$$\begin{aligned}\langle \mathbf{z}, \mathbf{x}\rangle &= \langle \mathbf{z}^{(1)}, \mathbf{x}^{(1)}\rangle + \langle \mathbf{z}^{(2)}, \mathbf{x}^{(2)}\rangle\\ &= \langle \mathbf{z}^{(1)}, \mathbf{A}\mathbf{x}^{(2)}\rangle + \langle -\mathbf{A}^{\mathrm{T}}\mathbf{x}^{(1)}-\mathbf{B}\mathbf{x}^{(2)}, \mathbf{x}^{(1)}\rangle\\ &= -\langle \mathbf{B}\mathbf{x}^{(2)}, \mathbf{x}^{(2)}\rangle = 0.\end{aligned}$$

Remark 13. This property can be viewed as an alternative expression of Tellegen's theorem.

LEMMA 4. *If* $\mathbf{H}$: $n \times n$ *is positive definite, then* $\det(\mathbf{PH}+\mathbf{Q})>0$.

Proof. According to (36), partition $\mathbf{H}$ as

$$\mathbf{H}=\left[\begin{array}{c|c} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \hline \mathbf{H}_{21} & \mathbf{H}_{22} \end{array}\right].$$

Then $\mathbf{PH}+\mathbf{Q}$ is of the form,

$$\mathbf{PH}+\mathbf{Q}=\left[\begin{array}{c|c} \mathbf{I} & -\mathbf{A} \\ \hline \mathbf{A}^{\mathrm{T}}\mathbf{H}_{11}+\mathbf{H}_{21} & \mathbf{H}_{22}+\mathbf{A}^{\mathrm{T}}\mathbf{H}_{12}+\mathbf{B} \end{array}\right]$$

$$=\left[\begin{array}{c|c} \mathbf{I} & \mathbf{0} \\ \hline \mathbf{A}^{\mathrm{T}}\mathbf{H}_{11}+\mathbf{H}_{21} & \mathbf{I} \end{array}\right]\cdot\left[\begin{array}{c|c} \mathbf{I} & -\mathbf{A} \\ \hline \mathbf{0} & \mathbf{C} \end{array}\right],$$

where

$$\mathbf{C}=[\mathbf{A}^{\mathrm{T}}\,|\,\mathbf{I}]\cdot\left[\begin{array}{c|c} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \hline \mathbf{H}_{21} & \mathbf{H}_{22} \end{array}\right]\left[\begin{array}{c} \mathbf{A} \\ \hline \mathbf{I} \end{array}\right]+\mathbf{B}.$$

Since $\mathbf{H}$ is positive definite and $\mathbf{B}$ is skew-symmetric, $\mathbf{C}$ is also positive definite. Hence $\det(\mathbf{PH}+\mathbf{Q})=\det\mathbf{C}>0$.

5.3. Formulation of the network equation. By rearranging $\mathbf{x}$ and $\mathbf{z}$ in (31) according to the partition (34) and then substituting (31) into (35), one obtains the network equation in the form

$$\mathbf{f}(\mathbf{x})=\mathbf{Ph}(\mathbf{x})+\mathbf{Qx}=\mathbf{y}, \tag{37}$$

where $\mathbf{f}: R^n \to R^n$ and the dimension n is equal to the total number of resistor ports.

A *solution* of a network is a set of branch voltages and branch currents which satisfy both Kirchhoff's laws and the branch constitutive relations. Thus to obtain a solution of a resistor network is nothing but to obtain a point $\mathbf{x}\in R^n$ which satisfies (37) for a given input $\mathbf{y}\in R^n$.

When there are some linear (passive) resistors in the given network, their variables can be eliminated in (37). If all such variables are eliminated, the dimension of the reduced equation is equal to the total number of nonlinear resistor ports. The reduced equation can also be described in the form of (37), although $\mathbf{P}$ and $\mathbf{Q}$ are not topological matrices any more. In the reduced equation, Lemma 3 should be replaced by[5]

LEMMA 3′. $\mathbf{Pz}+\mathbf{Qx}=\mathbf{0}$ *implies that* $\langle \mathbf{z}, \mathbf{x}\rangle \leqq 0$.

Remark 14. Although Lemma 3 has to be replaced by Lemma 3′, the solution algorithm presented here works as it stands even if it is applied to the reduced equation.

[5] The reduction of linear resistor variables yields the formulation due to Sandberg and Willson [15], and the pair $(\mathbf{P}, \mathbf{Q})$ possessing the property of Lemma 3′ is called a *passive pair.*

5.4. Existence and uniqueness of solutions. Regarding existence and uniqueness of solutions of $\mathbf{f}(\mathbf{x}) = \mathbf{y}$, the following theorem is known [8, p. 167].

THEOREM A. *Let* $\mathbf{f}: R^n \to R^n$ *be continuous. Then* (i) $\mathbf{f}$ *maps* R^n *onto itself if* $\mathbf{f}$ *is weakly coercive and* (ii) $\mathbf{f}$ *is a one-to-one mapping from* R^n *into itself if* $\mathbf{f}$ *is strictly monotone.*

In contrast to this theorem, it is interesting to give the following theorem regarding existence and uniqueness of solutions of resistor networks.

THEOREM 7. *Let* $\mathbf{h}: R^n \to R^n$ *be continuous and* $\mathbf{f}: R^n \to R^n$ *be defined by* (37). *Then* (i) $\mathbf{f}$ *maps* R^n *onto itself if* $\mathbf{h}$ *is strongly coercive and* (ii) $\mathbf{f}$ *is a one-to-one mapping from* R^n *into itself if* $\mathbf{h}$ *is strictly monotone.*

Proof. According to (36), partition (37) as

$$\begin{aligned} &\mathbf{x}^{(1)} - \mathbf{A}\mathbf{x}^{(2)} = \mathbf{y}^{(1)}, \\ &\mathbf{A}^{\mathrm{T}}\mathbf{h}^{(1)}(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) + \mathbf{h}^{(2)}(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) + \mathbf{B}\mathbf{x}^{(2)} = \mathbf{y}^{(2)}. \end{aligned} \tag{38}$$

Eliminating $\mathbf{x}^{(1)}$ from (38), we have

$$\begin{aligned} \mathbf{g}(\mathbf{x}^{(2)}, \mathbf{y}^{(1)}) = {} & \mathbf{A}^{\mathrm{T}}\mathbf{h}^{(1)}(\mathbf{y}^{(1)} + \mathbf{A}\mathbf{x}^{(2)}, \mathbf{x}^{(2)}) \\ & + \mathbf{h}^{(2)}(\mathbf{y}^{(1)} + \mathbf{A}\mathbf{x}^{(2)}, \mathbf{x}^{(2)}) + \mathbf{B}\mathbf{x}^{(2)} = \mathbf{y}^{(2)}. \end{aligned} \tag{39}$$

Thus the problem of solving (38) for given $\mathbf{y}$ is equivalent to solving (39) for given $\mathbf{y}^{(1)}$ and $\mathbf{y}^{(2)}$. Define a scalar function

$$\theta(\mathbf{x}^{(2)}, \mathbf{y}^{(1)}) = \langle \mathbf{g}(\mathbf{x}^{(2)}, \mathbf{y}^{(1)}), \mathbf{x}^{(2)} \rangle / \|\mathbf{x}^{(2)}\|$$

and n-vectors

$$\mathbf{x} = \begin{bmatrix} \mathbf{y}^{(1)} + \mathbf{A}\mathbf{x}^{(2)} \\ \hline \mathbf{x}^{(2)} \end{bmatrix}, \qquad \mathbf{u} = \begin{bmatrix} \mathbf{y}^{(1)} \\ \hline \mathbf{0} \end{bmatrix}$$

for given $\mathbf{x}^{(2)}$ and $\mathbf{y}^{(1)}$. Then it follows from the skew-symmetricity of $\mathbf{B}$ that

$$\theta(\mathbf{x}^{(2)}, \mathbf{y}^{(1)}) = \frac{\langle \mathbf{h}(\mathbf{x}), \mathbf{x} - \mathbf{u} \rangle}{\|\mathbf{x} - \mathbf{u}\|} \cdot \frac{\|\mathbf{x} - \mathbf{u}\|}{\|\mathbf{x}^{(2)}\|}.$$

Now observe that $\|\mathbf{x}\| \to \infty$ whenever $\|\mathbf{x}^{(2)}\| \to \infty$ and that

$$\|\mathbf{x} - \mathbf{u}\| = \|\mathbf{x}^{(2)}\| \cdot \left(1 + \frac{\|\mathbf{A}\mathbf{x}^{(2)}\|^2}{\|\mathbf{x}^{(2)}\|^2}\right)^{1/2}.$$

Thus strong coerciveness of $\mathbf{h}$ implies weak coerciveness of $\mathbf{g}$ for each fixed $\mathbf{y}^{(1)}$. Therefore it follows from Theorem A (i) that $\mathbf{g}$ is onto for each fixed $\mathbf{y}^{(1)}$, which implies that $\mathbf{f}$ is also onto; part (i) of Theorem 7 has thus been proved.

Let $\mathbf{u}, \mathbf{v} \in R^n$ be two distinct points. Then as long as $\mathbf{h}$ is strictly monotone,

$$\langle \mathbf{h}(\mathbf{u}) - \mathbf{h}(\mathbf{v}), \mathbf{u} - \mathbf{v} \rangle > 0.$$

On the other hand, the contraposition of Lemma 3 yields

$$\mathbf{f}(\mathbf{u}) - \mathbf{f}(\mathbf{v}) = \mathbf{P}(\mathbf{h}(\mathbf{u}) - \mathbf{h}(\mathbf{v})) + \mathbf{Q}(\mathbf{u} - \mathbf{v}) \neq \mathbf{0}.$$

Thus $\mathbf{f}$ has to be one-to-one; part (ii) has also been proved.

COROLLARY. *Let* $\mathbf{h}$ *and* $\mathbf{f}$ *be as in Theorem* 7. *Then* (i) $\mathbf{f}$ *maps* R^n *onto itself if* $\mathbf{h}$ *is uniformly passive and Lipschitzian and* (ii) $\mathbf{f}$ *is a homeomorphism of* R^n *onto itself if* $\mathbf{h}$ *is uniformly monotone.*

Proof. Due to the definitions in § 2, uniform passivity of $\mathbf{h}$ implies its weak coerciveness and, if $\mathbf{h}$ is Lipschitzian in addition, then $\mathbf{h}$ is strongly coercive. Thus part (ii) follows from Theorem 7 (i). Part (ii) is obvious since uniform monotonicity, by definition, implies strict monotonicity and strong coerciveness simultaneously.

Remark 15. In the corollary, "Lipschitzian" can be replaced by "piecewise-linear" due to Proposition 3.

Remark 16. Since $\mathbf{h}(\cdot)$ stands for the direct sum of resistor characteristics, uniform monotonicity (resp., uniform passivity with Lipschitz condition) imposed on the characteristics of each resistor guarantees existence of a unique solution (resp., at least one solution) of a resistor network independent of the interconnection.

5.5. Starting point for the solution algorithm. In the solution algorithm described in the previous section, the norm-coerciveness of a given piecewise-linear mapping is an underlying assumption.

LEMMA 5. *Let* $\mathbf{h}: R^n \to R^n$ *be a piecewise-linear, coercive mapping. Then* $\mathbf{f}: R^n \to R^n$ *in* (37) *is norm-coercive.*

Proof. According to (36), partition $\mathbf{f}$ as

$$\mathbf{f}^{(1)}(\mathbf{x}) = \mathbf{x}^{(1)} - \mathbf{A}\mathbf{x}^{(2)},$$
$$\mathbf{f}^{(2)}(\mathbf{x}) = \mathbf{A}^T\mathbf{h}^{(1)}(\mathbf{x}) + \mathbf{h}^{(2)}(\mathbf{x}) + \mathbf{B}\mathbf{x}^{(2)}.$$

Consider the inner product,

$$\langle \mathbf{f}^{(1)}(\mathbf{x}) - \mathbf{f}^{(1)}(\mathbf{0}), \mathbf{h}^{(1)}(\mathbf{x}) - \mathbf{h}^{(1)}(\mathbf{0}) \rangle + \langle \mathbf{f}^{(2)}(\mathbf{x}) - \mathbf{f}^{(2)}(\mathbf{0}), \mathbf{x}^{(2)} \rangle$$
$$= \langle \mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{0}), \mathbf{x} \rangle,$$

which yields

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{0})\| \cdot (\|\mathbf{h}^{(1)}(\mathbf{x}) - \mathbf{h}^{(1)}(\mathbf{0})\|^2 + \|\mathbf{x}^{(2)}\|^2)^{1/2} \geqq \langle \mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{0}), \mathbf{x} \rangle.$$

Since $\mathbf{h}^{(1)}$ is piecewise-linear, there exists a $\gamma > 0$ such that (see Proposition 3)

$$\|\mathbf{h}^{(1)}(\mathbf{x}) - \mathbf{h}^{(1)}(\mathbf{0})\| \leqq \gamma \|\mathbf{x}\|.$$

Noting that $\|\mathbf{x}^{(2)}\| \leqq \|\mathbf{x}\|$, it follows that

$$\|\mathbf{f}(\mathbf{x})\| \geqq \frac{\langle \mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{0}), \mathbf{x} \rangle}{\sqrt{1+\gamma^2}\|\mathbf{x}\|} - \|\mathbf{f}(\mathbf{0})\|.$$

Now it is seen that $\mathbf{f}$ is norm-coercive if $\mathbf{h}$ is coercive.

It is obvious that if $\mathbf{h}$ is piecewise-linear, then so is $\mathbf{f}$ defined by (37). Furthermore, the domain R^n of $\mathbf{f}$ is divided into the same set of polyhedral regions as that of $\mathbf{h}$, so that $\mathbf{f}$ and $\mathbf{h}$ are linear in each region. Let $\mathbf{H}_i$ be the Jacobian matrix

for $\mathbf{h}$ of a region, say R_i. Then the corresponding Jacobian matrix $\mathbf{J}_i$ for $\mathbf{f}$ is given by

$$\mathbf{J}_i = \mathbf{P}\mathbf{H}_i + \mathbf{Q}. \tag{40}$$

Let $\mathbf{u} \in R^n$ be an interior point of the region R_i, and assume that, for all $\mathbf{x} \neq \mathbf{u}$,

$$\langle \mathbf{h}(\mathbf{x}) - \mathbf{h}(\mathbf{u}), \mathbf{x} - \mathbf{u} \rangle > 0.$$

Then $\mathbf{H}_i$ must be positive definite, which implies by Lemma 4 that $\det \mathbf{J}_i > 0$. Furthermore, the contraposition of Lemma 3 yields $\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{u}) \neq 0$ for all $\mathbf{x} \neq \mathbf{u}$. Thus the following lemma has been proved.

LEMMA 6. *Let $\mathbf{u} \in R^n$ be an interior point of a region of a piecewise-linear mapping $\mathbf{h}$: $R^n \to R^n$. Then $\mathbf{u}$ satisfies the 1-degree condition with respect to $\mathbf{f}$ defined by* (37), *if $\mathbf{h}$ is strictly passive on the point $\mathbf{u}$.*

The following definition provides a criterion to determine a starting point which guarantees the convergence of the generalized Katzenelson algorithm applied to resistor network equations.

DEFINITION 2. A point $\mathbf{u} \in R^n$ is said to possess *Property* U with respect to a piecewise-linear mapping $\mathbf{h}$: $R^n \to R^n$ if $\mathbf{h}$ is uniformly passive on $\mathbf{u}$ and $\mathbf{u}$ does not lie on any boundary hyperplane in the domain of $\mathbf{h}$.

Suppose $\mathbf{u} \in R^n$ possesses Property U with respect to $\mathbf{h}$: $R^n \to R^n$. Then it implies, due to Lemma 5, that $\mathbf{f}$ in (37) is norm-coercive. It also implies, due to Lemma 6, that $\mathbf{u}$ satisfies the 1-degree condition with respect to $\mathbf{f}$. Thus the following theorem has been proved.

THEOREM 8. *Let $\mathbf{h}$: $R^n \to R^n$ be a piecewise-linear mapping and $\mathbf{f}$: $R^n \to R^n$ be defined by* (37), *and assume that there exists a point $\mathbf{u} \in R^n$ possessing Property* U *with respect to $\mathbf{h}$. Then the generalized Katzenelson algorithm, starting with $\mathbf{u}$, leads to a solution of* (37) *in a finite number of iteration steps for any $\mathbf{y} \in R^n$.*

This theorem suggests a scheme to give a starting point for the generalized Katzenelson algorithm. For the characteristics of each resistor, a point possessing Property U is chosen. Then the collection of such points for all the resistors constitutes an n-vector which also possesses Property U with respect to $\mathbf{h}$ defined by (32). In the special case where all the resistors are uniformly monotone, the original Katzenelson algorithm converges to the unique solution starting from any boundary-free point.

The piecewise-linear models of the static characteristics of almost all practical elements have points possessing Property U. This is demonstrated by two typical examples in Appendix B.

Appendix A. It is shown in this appendix that the definition of degree in (12), § 3 coincides with the one in [8]. Let $\mathbf{f}$: $R^n \to R^n$ be continuously differentiable and $C \subset R^n$ be an open bounded set. Then for any $\mathbf{y} \notin \mathbf{f}(\partial C)$, the degree of $\mathbf{f}$ at $\mathbf{y}$ with respect to C is defined by means of an integral [8, p. 149]. In a special case where $\mathbf{f}'(\mathbf{x})$ is nonsingular for all $\mathbf{x} \in \Gamma = \{\mathbf{x} \in C | \mathbf{f}(\mathbf{x}) = \mathbf{y}\}$, Γ consists of at most finitely many points $\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(m)}$ and the degree is defined as follows [8, p. 150]:

$$\deg(\mathbf{f}, C, \mathbf{y}) = \sum_{j=1}^{m} \operatorname{sgn} \det \mathbf{f}'(\mathbf{x}^{(j)}). \tag{A.1}$$

When a continuous mapping $\mathbf{f}: R^n \to R^n$ is not differentiable, the degree is defined as follows [8, p. 154]:

$$\deg(\mathbf{f}, C, \mathbf{y}) = \lim_{k\to\infty} \deg(\mathbf{f}^{(k)}, C, \mathbf{y}), \tag{A.2}$$

where $\mathbf{f}^{(k)}: R^n \to R^n$ is any sequence of continuously differentiable mappings such that $\lim_{k\to\infty} \|\mathbf{f}^{(k)} - \mathbf{f}\|_C = 0$. Here $\|\cdot\|_C$ denotes the L_∞-norm on C.

Now let $\mathbf{f}: R^n \to R^n$ be a continuous, piecewise-linear mapping and $C \subset R^n$ be an open bounded set. Then for any $\mathbf{y} \notin \mathbf{f}(\partial C) \cup \mathbf{f}(B \cup E)$, $\Gamma = \{\mathbf{x} \in C \mid \mathbf{f}(\mathbf{x}) = \mathbf{y}\}$ is a finite set. Let $\{\mathbf{f}^{(k)}\}$ be a sequence as defined above.

If $\Gamma = \varnothing$, it is clear from the compactness of $\bar{C}$ that $\{\mathbf{x} \in C \mid \mathbf{f}^{(k)}(\mathbf{x}) = \mathbf{y}\}$ is empty for all sufficiently large k. Therefore it follows from (A.1) and (A.2) that $\deg(\mathbf{f}, C, \mathbf{y}) = 0$.

If Γ is a nonempty finite set $\{\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(m)}\}$, each $\mathbf{x}^{(j)}$ is an interior point of a region R_j with nonsingular Jacobian matrix $\mathbf{J}^{(j)}$ for $j = 1, \cdots, m$. Note that R_j, $j = 1, \cdots, m$, are mutually distinct. Let p, q, $p < q$, be positive numbers such that the closed ball with center $\mathbf{x}^{(j)}$ and radius q is wholly contained in the interior of R_j for all $j = 1, \cdots, m$. Let $P_j = \{\mathbf{x} \in R^n \mid \|\mathbf{x} - \mathbf{x}^{(j)}\| < p\}$ and $Q_j = \{\mathbf{x} \in R^n \mid \|\mathbf{x} - \mathbf{x}^{(j)}\| > q\}$; then $\varphi_j(\mathbf{x}) = d^2(\mathbf{x}, \bar{Q}_j)/[d^2(\mathbf{x}, \bar{P}_j) + d^2(\mathbf{x}, \bar{Q}_j)]$ is a continuously differentiable function: $R^n \to R^1$ such that $0 \leqq \varphi_j(\mathbf{x}) \leqq 1$, $\varphi_j(\mathbf{x}) = 1$ for $\mathbf{x} \in \bar{P}_j$ and $\varphi_j(\mathbf{x}) = 0$ for $\mathbf{x} \in \bar{Q}_j$, where $d(\mathbf{x}, S)$ stands for the Euclidean distance between point $\mathbf{x} \in R^n$ and set $S \subset R^n$. Now let

$$\varphi(\mathbf{x}) = \sum_{j=1}^{m} \varphi_j(\mathbf{x})$$

and $\hat{\mathbf{f}}^{(k)} = \varphi\mathbf{f} + (1-\varphi)\mathbf{f}^{(k)}$, for $k = 1, 2, \cdots$; then $\hat{\mathbf{f}}^{(k)}$ is continuously differentiable because $\varphi\mathbf{f} = \mathbf{0}$ on the boundary hyperplanes on which $\mathbf{f}$ is not continuously differentiable. Furthermore, $\hat{\mathbf{f}}^{(k)}(\mathbf{x}) = \mathbf{f}(\mathbf{x})$ for $\mathbf{x} \in P_1 \cup \cdots \cup P_m$ and $\lim_{k\to\infty} \|\hat{\mathbf{f}}^{(k)} - \mathbf{f}\|_C = 0$ since $\|\hat{\mathbf{f}}^{(k)}(\mathbf{x}) - \mathbf{f}(\mathbf{x})\| = |1 - \varphi(\mathbf{x})| \cdot \|\mathbf{f}^{(k)}(\mathbf{x}) - \mathbf{f}(\mathbf{x})\| \leqq \|\mathbf{f}^{(k)}(\mathbf{x}) - \mathbf{f}(\mathbf{x})\|$ for any $\mathbf{x} \in R^n$. It is also easy to see from the compactness of $\bar{C} - P$ that $\hat{\mathbf{f}}^{(k)}(\mathbf{x}) = \mathbf{y}$ has no solution in $C - P$ for all sufficiently large k. Therefore $\{\mathbf{x} \in C \mid \hat{\mathbf{f}}^{(k)}(\mathbf{x}) = \mathbf{y}\} = \Gamma$ for all sufficiently large k and furthermore, $\hat{\mathbf{f}}^{(k)\prime}(\mathbf{x}^{(j)}) = \mathbf{J}_j$ for all j and sufficiently large k. Thus (A.1) and (A.2) lead to the definition (12) in § 3.

Appendix B. In practical applications, one often encounters a resistor network containing so-called "active" elements such as tunnel diodes and transistors. It is demonstrated here that the algorithm proposed in this paper works even if a given network contains such "active" elements. More precisely, a pertinent piecewise-linear model of such an "active" element has points possessing Property U.

(I). *Tunnel Diode.* Consider the diode which is characterized by $h: R^1 \to R^1$ illustrated in Fig. 10. Then it is seen that any point given by

$$u \in (-\infty, x^{(1)}) \cup (x^{(2)}, x^{(3)}) \cup (x^{(3)}, +\infty) \tag{B.1}$$

possesses Property U.

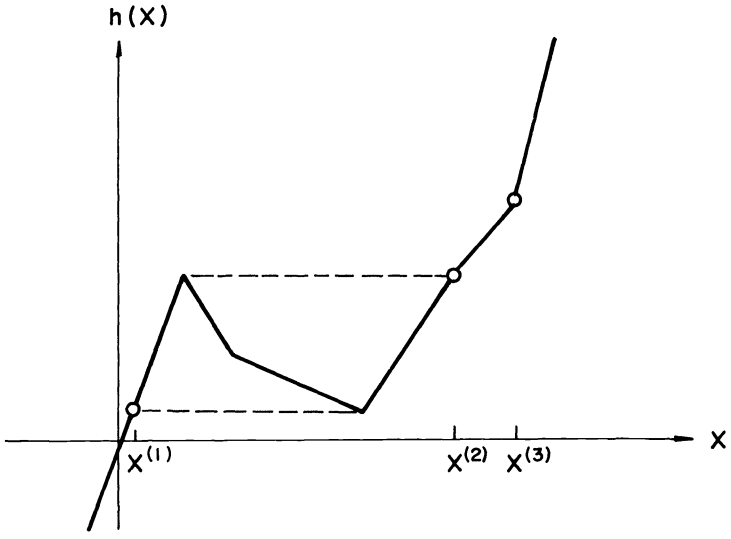


FIG. 10. *Piecewise-linear resistor of tunnel-diode type*

(II). *Transistor of Ebers-Moll Type.* Consider the 2-port resistor characterized by $\mathbf{g}(\mathbf{v})$: $R^2 \to R^2$ which can be written in the form[6]

$$\mathbf{g}(\mathbf{v}) = \left[\begin{array}{c|c} 1 & -\alpha_1 \\ \hline -\alpha_2 & 1 \end{array}\right] \cdot \left[\begin{array}{c} m_1 q(n_1 v_1) \\ \hline m_2 q(n_2 v_2) \end{array}\right], \tag{B.2}$$

where

$$\begin{aligned} &0 \leqq \alpha_1, \alpha_2 < 1, \\ &\alpha_1 < m_1/m_2, n_1/n_2 < 1/\alpha_2, \\ &m_1 n_1, m_2 n_2 > 0 \end{aligned} \tag{B.3}$$

and $q(\cdot)$ is a strictly monotone piecewise-linear mapping.[7] Note that, due to Proposition 5, there exists $k > 0$ such that

$$[q(x) - q(x^0)] \cdot (x - x^0) \geqq k(x - x^0)^2 \tag{B.4}$$

for all $x \in R^1$.

Let $\mathbf{u} = (u_1, u_2)^T$ be a point such that

$$n_1 u_1 = n_2 u_2 = x^0, \tag{B.5}$$

where x^0 is an arbitrary boundary-free point in the domain of $q(\cdot)$. In order to

[6] The class of 2-port resistors discussed here is the same as the transistor model discussed in [15, Thm. 6], except that $q(\cdot)$ has to be strictly monotone and piecewise-linear in this paper.

[7] More generally, $q(\cdot)$ can be of the form as in Fig. 10. This condition is imposed to simplify the proof.

prove that such a $\mathbf{u}$ possesses Property U, the same line of logical development as in the proof of Theorem 6 of [15] is applied.

Define a scalar function $\theta(\mathbf{v})$ by

$$\theta(\mathbf{v}) = \langle \mathbf{g}(\mathbf{v}) - \mathbf{g}(\mathbf{u}), \mathbf{v} - \mathbf{u} \rangle \tag{B.6}$$

and the transformation $\mathbf{v} = (v_1, v_2)^{\mathrm{T}} \to \mathbf{x} = (x_1, x_2)^{\mathrm{T}}$ by

$$x_1 = n_1 v_1, \qquad x_2 = n_2 v_2. \tag{B.7}$$

Note that, due to (B.5) and (B.7),

$$(x_1 - x^0)^2 + (x_2 - x^0)^2 \geqq n^2 \|\mathbf{v} - \mathbf{u}\|^2, \tag{B.8}$$

where $n^2 = \min\{n_1^2, n_2^2\}$. Let

$$\varphi(\mathbf{x}) = [x_1 - x^0 \,|\, x_2 - x^0] A \begin{bmatrix} q(x_1) - q(x^0) \\ \hline q(x_2) - q(x^0) \end{bmatrix}, \tag{B.9}$$

where

$$A = \left[\begin{array}{c|c} a_{11} & -a_{12} \\ \hline -a_{21} & a_{22} \end{array}\right] = \left[\begin{array}{c|c} m_1/n_1 & -\alpha_1 m_2/n_1 \\ \hline -\alpha_2 m_1/n_2 & m_2/n_2 \end{array}\right]. \tag{B.10}$$

Then it is easily seen that $\varphi(\mathbf{x}) = \theta(\mathbf{v})$. Furthermore it follows from (B.3) that

$$a_{11} > a_{12},\ a_{21} > 0,$$
$$a_{22} > a_{21},\ a_{12} > 0,$$

and hence $a = \min\{a_{11} - a_{12}, a_{11} - a_{21}, a_{22} - a_{12}, a_{22} - a_{21}\}$ is a positive number.

In order to observe the behavior of

$$\begin{aligned} \varphi(\mathbf{x}) = {} & [q(x_1) - q(x^0)] \cdot [a_{11}(x_1 - x^0) - a_{21}(x_2 - x^0)] \\ & + [q(x_2) - q(x^0)] \cdot [a_{22}(x_2 - x^0) - a_{12}(x_1 - x^0)], \end{aligned} \tag{B.11}$$

the $x_1 - x_2$ plane is divided into six regions as illustrated in Fig. 11.

In region I, $x_1 - x^0 \geqq x_2 - x^0 \geqq 0$. In the case where $a_{22}(x_2 - x^0) - a_{12}(x_1 - x^0) \geqq 0$, the second term of the right-hand side of (B.11) is nonnegative. Then since $x_1 - x^0 \geqq x_2 - x^0 \geqq 0$, it follows that

$$\varphi(\mathbf{x}) \geqq (a_{11} - a_{21})[q(x_1) - q(x^0)](x_1 - x^0) \geqq ka(x_1 - x^0)^2.$$

On the other hand, if $a_{22}(x_2 - x^0) - a_{12}(x_1 - x^0) \leqq 0$, then since $q(x_1) - q(x^0) \geqq q(x_2) - q(x^0) \geqq 0$, it also follows that

$$\begin{aligned} \varphi(\mathbf{x}) \geqq {} & [q(x_1) - q(x^0)] \cdot [(a_{11} - a_{12})(x_1 - x^0) \\ & + (a_{22} - a_{21})(x_2 - x^0)] \geqq ka(x_1 - x^0)^2. \end{aligned}$$

Since $x_1 - x^0 \geqq x_2 - x^0$ in region I, it follows from (B.8) that

$$\theta(\mathbf{v}) = \varphi(\mathbf{x}) \geqq \tfrac{1}{2} kan^2 \|\mathbf{v} - \mathbf{u}\|^2. \tag{B.12}$$

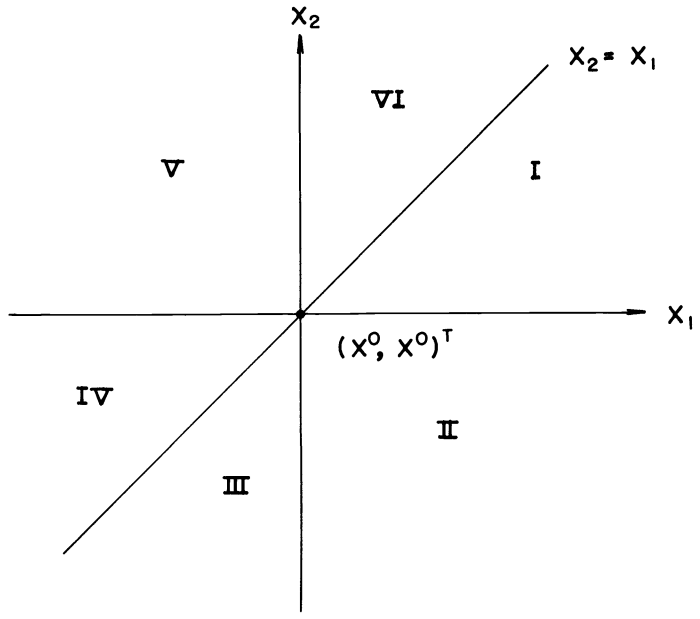


FIG. 11. *Regions considered in an Ebers–Moll type transistor*

In region II, $x_1-x^0 \geqq 0 \geqq x_2-x^0$, and hence both terms of the right-hand side of (B.11) are nonnegative. If $x_1-x^0 \geqq -(x_2-x^0)$, consider the inequalities $\varphi(\mathbf{x}) \geqq [q(x_1)-q(x^0)]a_{11}(x_1-x^0) \geqq ka(x_1-x^0)^2$. On the other hand, if $x_1-x^0 \leqq -(x_2-x^0)$, consider the inequalities $\varphi(\mathbf{x}) \geqq [q(x_2)-q(x^0)]a_{22}(x_2-x^0) \geqq ka(x_2-x^0)^2$. Thus by means of (B.8), (B.12) holds for any $\mathbf{x}$ in region II.

In region III, one can also derive (B.12) by means of a few obvious changes in the argument used for region I. Now, due to the symmetric nature of $\varphi(\mathbf{x})$ (with respect to the subscripts 1, 2) it is clear that the behavior is similar in the other half-plane.

Now it has been proved that $\mathbf{g}$ is uniformly passive on any $\mathbf{u}$ satisfying (B.5). Therefore, unless x^0 is a boundary point of the piecewise-linear mapping $q(\cdot)$, any point $\mathbf{u}$ satisfying (B.5) possesses Property U.

REFERENCES

[1] J. KATZENELSON, *An algorithm for solving nonlinear resistive networks*, Bell System Tech. J., 44 (1965), pp. 1605–1620.

[2] L. O. CHUA, *Efficient computer algorithm for piecewise-linear analysis of resistive nonlinear networks*, IEEE Trans. Circuit Theory, CT-18 (1971), pp. 73–85.

[3] T. OHTSUKI AND N. YOSHIDA, *DC analysis of nonlinear networks based on generalized piecewise-linear characterization*, Ibid., pp. 146–152.

[4] E. S. KUH AND I. N. HAJJ, *Nonlinear circuit theory: resistive networks*, Proc. IEEE, 59 (1971), pp. 340–355.

[5] T. FUJISAWA AND E. S. KUH, *Piecewise-linear theory of nonlinear networks*, SIAM J. Appl. Math., 22 (1972), pp. 307–328.

[6] T. FUJISAWA, E. S. KUH AND T. OHTSUKI, *A sparse matrix method for analysis of piecewise-linear resistive networks*, IEEE Trans. Circuit Theory, CT-19 (1972), pp. 571–584.

[7] K. KAWAKITA AND T. OHTSUKI, *NECTAR 2: a circuit analysis program based on piecewise-linear approach*, Proc. 1975 IEEE Internat. Symp. on Circuits and Systems.

[8] J. M. ORTEGA AND W. C. RHEINBOLDT, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[9] B. GOPINATH AND D. MITRA, *When are transistors passive?*, Bell System Tech. J., 50 (1971), pp. 2835–2847.

[10] J. L. KELLY, *General Topology*, Van Nostrand, New York, 1955.

[11] L. S. PONTRYAGIN, *Foundations of Combinatorial Topology*, English translation, Graylock Press, Rochester, 1952.

[12] F. F. WU, *Existence of an operating points for a nonlinear circuit using the degree of mapping*, IEEE Trans. Circuits and Systems, CAS-21 (1974), pp. 671–677.

[13] E. S. KUH AND R. A. ROHRER, *The state-variable approach to network analysis*, Proc. IEEE, 53 (1965), pp. 672–686.

[14] 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.

[15] I. W. SANDBERG AND A. N. WILLSON, JR., *Existence of solutions for the equations of transistor-resistor-voltage source networks*, Ibid., CT-18 (1971), pp. 619–625.

SIAM J. MATH. ANAL.
Vol. 8, No. 1, February 1977

SOME ALTERNATIVE EXPANSIONS FOR THE DISTRIBUTION FUNCTION OF A NONCENTRAL CHI-SQUARE RANDOM VARIABLE*

RUDY A. GIDEON† AND JOHN GURLAND‡

Abstract. A differential operator is defined and applied to the gamma density function. By formal mathematical manipulations, the resulting function is identified with the distribution function of the noncentral chi-square distribution. Several series expansions of a general nature result, and a table is presented comparing the effectiveness of seven series in evaluating this distribution function.

1. Introduction. The main purpose of this paper is to present a formal operator technique for obtaining series expansions of the distribution function of quadratic forms in independent normal variables, and to compare the effectiveness of these series expansions. Although this formal operator procedure with its algebraic manipulations can be applied in a general way to many distributions, the emphasis will be on the noncentral chi-square distribution. In particular, it will be shown how this unified approach leads to the development of power, chi-square, and various Laguerre series.

A general development and specific parameter choices for these series can be found in [15, Chaps. 28–9]. The operator approach of this paper extends the theory somewhat by allowing a broader choice of the parameters involved in these series expansions.

A paper by R. B. Davies [3] investigates a numerical integration method of evaluating the noncentral chi-square distribution function by an inversion formula for the characteristic function. D. R. Jensen and H. Solomon [11] give an improved Gaussian Wilson–Hilferty-type approximation formula. H. Ruben [19] summarizes many recursive formulas and gives finite series expansions in terms of Chebyshev-Hermite polynomials. Finally, C. G. Khatri [12] uses the approach of Kotz, et al. [13], [14] on the multivariate noncentral case. The approach of this paper may prove to be useful in the multivariate case, since like Khatri's paper, it identifies an infinite series expansion with a specific distribution. This paper complements the work of D. R. Jensen and H. Solomon in that the table lists parameter values of the noncentral chi-square variable in which a Gaussian-type approximation is expected to do well (see the Edgeworth series). Finally, the operator approach of H. Ruben [19] relies on the Hermite polynomials, whereas the approach of this paper is based on the generalized Laguerre polynomials.

Although there are a large number of papers developing theoretical forms for the distribution function (d.f.) of these quadratic forms, only a few, e.g., Gurland [7], Tiku [22], and Kotz, Johnson, and Boyd [13], [14], [15], give an indication of which series are to be preferred in obtaining probabilities as the values of the two parameters in the noncentral χ^2-distribution vary. The formal series development in this paper is followed by a table giving a comparison of relative effectiveness of seven series expansions.

* Received by the editors November 12, 1974, and in revised form September 22, 1975.
† Department of Mathematics, University of Montana, Missoula, Montana 59801.
‡ Statistics Department, University of Wisconsin—Madison, Madison, Wisconsin 53706.

2. Specification of an operator function for quadratic forms. Let $g_{\alpha,\lambda}$ denote a gamma density function of the form

$$g_{\alpha,\lambda}(x)=\frac{\lambda(\lambda x)^{\alpha}e^{-\lambda x}}{\Gamma(\alpha+1)}, \qquad x\geqq 0, \tag{2.1}$$

and let $L_n^{(\alpha)}(x)$ denote the nth generalized Laguerre polynomial. The differential operator D^r on $g_{\alpha,\lambda}(x)$ is defined as

$$D^r[g_{\alpha,\lambda}(x)]=\left(\frac{d}{dx}\right)^r g_{\alpha+r,\lambda}=\frac{\lambda^{\alpha+r+1}x^{\alpha}e^{-\lambda x}\Gamma(r+1)L_r^{(\alpha)}(\lambda x)}{\Gamma(\alpha+r+1)}. \tag{2.2}$$

If w is defined as $w=it/(1-it/\lambda)$, where $i=\sqrt{-1}$, then the characteristic function (c.f.) of D^r applied to $g_{\alpha,\lambda}(x)$ is given by

$$\int_0^{\infty} e^{itx}D^r g_{\alpha,\lambda}(x)\,dx=(-w)^r/(1-it/\lambda)^{\alpha+1}. \tag{2.3}$$

In the following, the operator D and the variable w are to be treated formally as a variable in a power series. A function of D which will generate the series expansions for quadratic forms in independent normal variates can be defined as

$$\frac{\exp\left[\sum_{j=1}^{m}(F_jD)/(1-G_jD)\right]}{\prod_{j=1}^{m}(1-A_jD)^{K_j}}, \tag{2.4}$$

where the F_j, G_j, A_j, K_j, $j=1,2,\cdots,m$, are constants. It will presently be shown that appropriate choices of these constants will yield series expansions for the noncentral chi-square distribution ($\chi'^2_{n,\Lambda}$) with degrees of freedom n and noncentrality parameter Λ. By formally computing the characteristic function $\psi(t)$ of the operator function (2.4) applied to density function (2.1) and using (2.3) with various algebraic manipulations, it can easily be shown that

$$\psi(t)=\frac{(1-it/\lambda)^{-(\alpha+1)}\exp\left[\sum_{r=1}^{m}(-F_rw)/(1+G_rw)\right]}{\prod_{j=1}^{m}(1+A_jw)^{-K_j}}. \tag{2.5}$$

In order to utilize $\psi(t)$, it is necessary to rewrite (2.5) in two distinct forms. The first form will be called the expansion form and it regroups all the terms involving like powers of w as follows:

$$\psi(t)=(1-it/\lambda)^{-(\alpha+1)}\exp\left[\sum_{l=1}^{\infty}(-w)^l d_l\right],$$

where

$$d_l=\lambda^{-l}\left[\sum_{j=1}^{m}\{\lambda F_j(\lambda G_j)^{l-1}+l^{-1}K_j(\lambda A_j)^l\}\right]. \tag{2.6}$$

The second form will be called the identification form as it is used to equate $\psi(t)$ to the charcteristic function of a particular random variable (r.v.). In order to obtain this form, replace w by $it/(1-it/\lambda)$ in (2.5), and after regrouping $\psi(t)$ in powers of

it,

$$\psi(t) = (1 - it/\lambda)^{-(\alpha+1)} \exp\left[\sum_{l=1}^{\infty} (-it)^l h_l\right],$$

(2.7) where

$$h_l = \lambda^{-l}\left[\sum_{j=1}^{m} \{\lambda F_j(\lambda G_j - 1)^{l-1} + l^{-1}K_j((\lambda A_j - 1)^l - (-1)^l)\}\right].$$

In order to generate series expansions, the following recursive formulation is needed (cf. [13] or [15, p. 160]). If

$$\exp\left\{\sum_{k=1}^{\infty} \frac{d_k}{k} t^k\right\} = \sum_{k=0}^{\infty} C_k t^k,$$

(2.8) then

$$C_k = \frac{1}{k}\sum_{j=0}^{k-1} d_{k-j}C_j, \qquad k = 1, 2, \cdots,$$

$$C_0 = 1.$$

3. Alternative expansions. The constants F_j, G_j, K_j, A_j will be specified completely by the r.v. for which we seek an infinite series expansion. However, the constants $\alpha + 1$ and λ can be suitably selected to give one or two moment fit agreement between the r.v. whose d.f. is being expanded and the first term of this series. However, in order to get a three moment fit, a location parameter θ is introduced. By letting $\psi_\theta(t) = e^{\theta it}\psi(t)$, and using the formal identities

$$(3.1) \qquad e^{\theta it} = \exp\left\{\theta \frac{w}{1 + (w/\lambda)}\right\} = \exp\left\{\theta\lambda \sum_{l=1}^{\infty} (-1)^{l-1}\left(\frac{w}{\lambda}\right)^l\right\},$$

$\psi_\theta(t)$ can be used to represent the Laguerrian expansion of the c.f. of a r.v. $Z = X + \theta$, where $\psi(t)$ is the c.f. of the r.v. X and θ is a constant to be selected. If $(1 - (it/\lambda)^{-(\alpha+1)}$ in (2.6), and (2.7) is rewritten as

$$(3.2) \qquad \exp\left\{+(\alpha+1)\sum_{l=1}^{\infty} \frac{1}{l}\left(\frac{it}{\lambda}\right)^l\right\} \quad \text{or} \quad \exp\left\{(\alpha+1)\sum_{l=1}^{\infty} \frac{(-1)^{l-1}}{l}\left(\frac{w}{\lambda}\right)^l\right\},$$

then equations (2.6) or (2.7) can be viewed as exponentiated series in w and it respectively. By combining (2.6) and (3.2) with (3.1) to form $\psi_\theta(t)$, the first term of the Laguerrian expansion can be made to have 1, 2 or 3 moments in common with the r.v. being expanded by setting the coefficients of w^l, $l = 1, 2, 3$, to zero and solving the equation for $l = 1$, $l = 1$ and 2, $l = 1$, 2 and 3 according as a 1, 2, or 3 moment agreement is desired. The detailed solution for $\chi'^2_{n,\Lambda}$ will be presented below.

4. Laguerrian expansions for the d.f. of a noncentral chi-square r.v. with n degrees of freedom and noncentrality parameter Λ; $\chi'^2_{n,\Lambda}$. In this section, the general formulation of the previous sections is used to obtain specific expansions for the r.v. $\chi'^2_{n,\Lambda}$. First, the constants F_j, G_j, A_j, K_j from (2.7) are chosen so that $\psi(t)$

represents the c.f. of $\chi'^2_{n,\Lambda}$. Equation (2.7) in the form utilizing (3.2) is used to equate $\psi(t)$ to the c.f. of $\chi'^2_{n,\Lambda}$.

If Y is a r.v. defined as $\sum_{i=1}^{n} \alpha_i(Z_i+\delta_i)^2$, where α_i are positive constants, δ_i are constants, and Z_i independent $N(0,1)$-r.v.'s, then Y is a noncentral positive definite quadratic form in normal variable. If all the $\alpha_i = 1$ and all $\delta_i = \delta$, then Y becomes $\chi'^2_{n,\Lambda}$, where $\Lambda = n\delta^2$. Although the generalities of the previous work can be applied to Y, it is restricted at this time to the special but important case of $\chi'^2_{n,\Lambda}$. Since the c.f. of $\chi'^2_{n,\Lambda}$ can be written in the form

$$\exp\left\{\sum_{l=1}^{\infty} \frac{(it)^l}{l!} 2^{l-1}(l-1)!n(1+l\delta^2)\right\}, \tag{4.1}$$

$\psi(t)$ can be equated to (4.1) by choosing

$$\begin{aligned} &\lambda A_j = \lambda G_j = 1-2\lambda, \quad m = n, \\ &F_j = -\delta^2 \quad \text{and} \quad K_j = \tfrac{1}{2} \quad \text{for } j = 1, 2, \cdots, n, \end{aligned} \tag{4.2}$$

and letting, for the moment, $\alpha+1 = \sum_{j=1}^{n} K_j = n/2$. Now that $\psi(t)$ has been identified with the c.f. of $\chi'^2_{n,\Lambda}$, $\psi(t)$ can be used in the form (2.6) directly or with the location shift modification (3.1) to obtain Laguerrian expansions.

The general case is taken by expanding $\psi_\theta(t)$, since $\psi(t) = \psi_\theta(t)$ if $\theta = 0$. Thus let $d_l' = d_l - (\theta\lambda/\lambda^l)$, where d_l is defined in (2.6) and $\theta\lambda/\lambda^l$ is the coefficient of $(-w)^l$ in (3.1). In order to use recursive formula (2.8), we must multiply d_l' by l so that the coefficient of $(-w)^l/l$ becomes $ld_l - l\theta/\lambda^{l-1}$. Thus the C_k-coefficients of (2.8) are defined as

$$C_0 = 1, \quad C_k = \frac{1}{k}\sum_{j=0}^{k-1} C_j(k-j)\left\{d_{k-j} - \frac{\theta}{\lambda^{k-j-l}}\right\}$$

(4.3) and

$$\psi_\theta(t) = \left(1-\frac{it}{\lambda}\right)^{-(\alpha+1)}\left\{1+\sum_{k=1}^{\infty} C_k(-w)^k\right\}.$$

Now inversion formula (2.3) can be used with (2.2) and formal term by term integration applied to obtain (letting $x+\theta = x_1$)

$$\begin{aligned} P\{\chi'^2_{n,\Lambda} \leqq x\} \sim &\int_0^{x_1} \frac{\lambda^{\alpha+1} y^\alpha e^{-\lambda y}}{\Gamma(\alpha+1)}\, dy \\ &+ e^{-\lambda x_1}(\lambda x_1)^{\alpha+1} \sum_{k=1}^{\infty} \frac{\lambda^k C_k \Gamma(k)}{\Gamma(\alpha+1+k)} L_{k-1}^{(\alpha+1)}(\lambda x_1). \end{aligned} \tag{4.4}$$

The symbol $\sim$ is meant to imply the formal equating of two mathematical quantities which under certain conditions on the parameters may be identical.

In formula (4.4), the two parameters θ and λ are still undefined, and although $\alpha+1$ has been chosen as $n/2$, other choices are possible, as will soon be demonstrated.

The notation L_i will be used to denote an expression of the form (4.4) in which the first i moments of the r.v. associated with the leading term of the expansion (a gamma r.v.) have been equated to corresponding moments of $\chi'^2_{n,\Lambda}$.

The subscript i will be equal to 0, 1, 2 or 3, and for the cases $i=0$, 1 or 2, the parameter θ is chosen to be zero. For the case $i=0$, $\alpha+1=n/2$, but λ is still free to be selected.

Since $\psi(t)$, by utilizing (3.2) and (4.2) in (2.6), has already been equated to the c.f. of $\chi'^2_{n,\Lambda}$, only θ and λ are free to be chosen. However, $\alpha+1$ can also be included as a free parameter if $\psi(t)$ is multiplied by $(1-it/\lambda)^{-(\alpha+1)}$, and from (3.2) it is seen that $\psi(t)$ is unchanged if $-(\alpha+1)\sum_l((-1)^{l+1}/l)(w/\lambda)^l$ is added as a cancellation in the exponent. $\psi_\theta(t)$ is now in its most general form for the Laguerrian expansions

$$\psi_\theta(t)=\psi(t)\,e^{\theta it}=\left(1-\frac{it}{\lambda}\right)^{-(\alpha+1)} \cdot \exp\left[\sum_{l=1}^{\infty}\frac{(-w)^l}{l}\frac{1}{\lambda^l}\left\{(\alpha+1)-\frac{n}{2}-l\theta\lambda-\frac{n}{2}(1-2\lambda)^{l-1}(2\lambda+2\lambda l\delta^2-1)\right\}\right]. \tag{4.5}$$

The L_i series are now obtained as indicated by solving equations involving the coefficients of $(w)^l$, $l=1,2,3$.

For L_0, $\alpha+1=n/2$, $\theta=0$, $\lambda=1/2$.

For L_1, $\alpha+1=n/2$, $\theta=0$, $\lambda=(\alpha+1)/n(1+\delta^2)$.

For L_2, $\alpha+1=n\lambda(1+\delta^2)$, $\theta=0$, $\lambda=(1+\delta^2)/2(1+2\delta^2)$.

For L_3, $\alpha+1=2n\lambda^2(1+2\delta^2)$, $\theta=2n\lambda(1+2\delta^2)-n(1+\delta^2)$, $\lambda=(1+2\delta^2)/2(1+3\delta^2)$.

The various Laguerrian expansions are of the same form as (4.4), where the constants C_k are derived recursively from d''_l, where d''_l are defined to be the coefficients of $(-w)^l/l$ in the exponent of (4.5).

5. Chi-square, power, and Edgeworth-expansions. In order to ascertain the relative effectiveness of the alternative Laguerrian expansions above, the following chi-square, power, and Edgeworth-expansions are provided and included in a computer study.

For the χ^2-expansion, define $u=(1-(it/\lambda))^{-1}$. The c.f. needed for the term by term integration in the χ^2-case is given by

$$\int_0^\infty e^{itx}g_{\alpha+k,\lambda}(x)\,dx=u^{+(\alpha+k+1)}. \tag{5.1}$$

In order to obtain a χ^2-expansion formula, (2.6) needs to be rewritten in terms of the parameter u. If $w=\lambda(u-1)$ is substituted in equation (2.5) in which the choices of the constants are given by (4.2) and a few algebraic manipulations performed, $\psi(t)$ becomes

$$\psi(t)=\frac{\exp(-(n/2)\delta^2)}{2^{n/2}}\left(\frac{u}{\lambda}\right)^{n/2}\left[\exp\sum_{l=1}^{\infty}\frac{u^l}{l}e_l\right], \tag{5.2}$$

where

$$e_l=\frac{n}{2}\left(\frac{2\lambda-1}{2\lambda}\right)^{l-1}\left(1+\frac{l\delta^2-1}{2\lambda}\right).$$

$\psi(t)$ can now be rewritten using recursive formula (2.8) in which $C_0 = 1$, $C_k = (1/k)\sum_{j=0}^{k-1} e_{k-j}C_j$. After applying term by term inversion (5.1) and integration, the final result is

$$P(\chi'^2_{n,\Lambda} \leqq x) \sim \frac{\exp(-(n/2)\delta^2)}{2^{n/2}}\left[\sum_{k=0}^{\infty} C_k \int_0^x \frac{\lambda(\lambda t)^{\alpha+k} e^{-\lambda t}}{\Gamma(\alpha+k+1)}\,dt\right]. \tag{5.3}$$

In order to develop the power series-expansion, the Fourier transform of the function x^{n-1} over the positive real axis is needed:

$$\int_0^{\infty} e^{itx}\frac{x^{n-1}}{\Gamma(n)}\,dx = (-it)^{-n}. \tag{5.4}$$

As in the χ^2-case, it is necessary to rewrite $\psi(t)$ of (2.5), however, this time in a series involving powers of $(it)^{-1}$. This is accomplished by substituting for w the quantity $it/(1-it/\lambda)$, where again the choices of the constants given by (4.2) are also inserted. After some algebraic manipulations, the following result follows:

$$\psi(t) = \frac{1}{2^{n/2}}\exp\left\{-\frac{n}{2}\delta^2\right\}(-it)^{-(n/2)}\exp\left[\sum_{l=1}^{\infty}\frac{(-it)^{-l}}{l}b_l\right], \tag{5.5}$$

where

$$b_l = \frac{n}{2}\left(-\frac{1}{2}\right)^l(1-l\delta^2).$$

$\psi(t)$ is now in a form to which the recursive formula (2.8) can be applied, where $C_0 = 1$, $C_k = (1/k)\sum_{j=0}^{k-1} b_{k-j}C_j$. After this, the term by term inversion (5.4) is applied and then integration; the result is

$$P(\chi'^2_{n,\Lambda} \leqq x) \sim \frac{1}{2^{n/2}}\exp\left\{-\frac{n}{2}\delta^2\right\}\sum_{k=0}^{\infty} C_k\frac{x^{\alpha+k+1}}{\Gamma(\alpha+k+1)}. \tag{5.6}$$

For a complete comparison, the Hermite–Edgeworth-expansion is included. This expansion (cf. Cramer [2]) is a regrouping of an expansion of the Hermite polynomials for the d.f. of $\chi'^2_{n,\Lambda}$ in which coefficients of like powers of $n^{-k/2}$ are grouped together. A recursive development for this expansion is now presented. Let χ_k be the kth cumulant of $\chi'^2_{n,\Lambda}$. Then from (4.1),

$$\chi_k = 2^{k-1}(k-1)!n(1+k\delta^2). \tag{5.7}$$

If $d = -\chi_1$ and $\rho^2 = \chi_2$, then the c.f. of $(\chi'^2_{n,\Lambda}+d)/\rho$ is

$$\begin{aligned}\psi\left(\frac{t}{\rho}\right)e^{(d/\rho)it} &= E\left(e^{it\frac{\chi'^2_{n,\Lambda}+d}{\rho}}\right)\\ &= e^{-(t^2/2)}\exp\{(it)^2\sum_{k=1}^{\infty} d_k(it)^k\},\end{aligned} \tag{5.8}$$

where

$$d_k = \frac{\chi_{k+2}}{(k+2)!\rho^{k+2}} = \frac{1}{n^{k/2}}\frac{2^{k+1}\{1+\delta^2(k+2)\}}{(k+2)\{2(1+2\delta^2)\}^{(k+2)/2}}.$$

The c.f. of $(\chi'^2_{n,\Lambda}+d)/\rho$ can be rewritten as

$$\psi\left(\frac{t}{\rho}\right)e^{(d/\rho)it}=e^{-t^2/2}+\sum_{k=1}^{\infty}\left\{\sum_{h=1}^{k}\frac{C_{h,k}}{h!}(it)^{k+2h}e^{-t^2/2}\right\}, \tag{5.9}$$

where

$$C_{1,k}=d_k,\quad C_{2,k}=\sum_{m=1}^{k-1}C_{1,m}C_{1,k-m},\ \cdots,\ C_{h,k}=\sum_{m=1}^{k-h+1}C_{1,m}C_{h-1,k-m}.$$

It is easily seen that $C_{h,k}$ contains as a factor the term $n^{-k/2}$. If $H_k(y)$ denotes the usual Hermite polynomial and the well-known relationships of the polynomials are used, then (5.9) can be inverted and integrated term by term to produce

$$P(\chi'^2_{n,\Lambda}\leqq ey-d)\sim\int_{-\infty}^{y}\frac{e^{-u^2/2}}{\sqrt{2\pi}}\,du-\sum_{k=1}^{\infty}\sum_{h=1}^{k}\frac{C_{h,k}}{k!}H_{k+2h-1}(y)\,e^{-y^2/2}. \tag{5.10}$$

The seven series which have been compared on a computer have been developed in formulas (5.10), the Edgeworth; (5.6), the power; (5.3), the χ^2; and (4.4), the various alternative Laguerre series.

6. Comparison of the effectiveness of the series expansions for the d.f. of $\chi'^2_{n,\Lambda}$. All the comments made here for the seven series expansion for the d.f. of $\chi'^2_{n,\Lambda}$ are based on the results of computer evaluations. The results are summarized in Table 1, and a few general comments are given now.

In general, for $\chi'^2_{n,\Lambda}$, the Laguerre series expansions (L_i) are best for Λ small in relation to the parameter n. In contrast, the Edgeworth series (E) is best for Λ large in relation to n and, in particular, is very strong in the central areas of the distributions, that is, probabilities between 0.2 and 0.8. For low degrees of freedom and small noncentrality Λ, the series E, L_2 and L_3 are very poor, but the series L_0 is very good. The series L_1 is also good for low degrees of freedom n and small noncentrality Λ.

Although the chi-square (χ^2) and power (P) series exhibit their best performance for n and Λ both relatively small, they fall substantially short of the performance of L_0. These series usually perform best in the left tails of the distributions. Characteristically, the χ^2 and P-series converge very rapidly after a period of very slow convergence. For example, after 20 terms, the partial sums may be accurate only to one place, but after four to eight additional terms, five place decimal accuracy may be achieved. The Edgeworth series E also has this characteristic, but, in general, if the E-series has not attained four to eight place decimal accuracy after about 12 terms, it will start to diverge.

For n large and small noncentrality Λ, the series L_0, L_1 are the best. The series L_2 and L_3 produce good results in this range also. In choosing between a higher moment series L_2 or L_3 and a lower moment series such as L_0, it should be pointed out that the higher moment series usually give three to five place decimal accuracy with just a few terms (one to five) but may gain additional accuracy slowly. The series L_0 however, usually produces less accuracy up to the first five to ten partial sums, but after that converges rapidly to the true probability.

For n large and Λ large, the strength of the (E, L_1)-series in producing good approximations to the true probability is conspicuous from Table 1. For small n

TABLE 1

Table of comparison of series for the d.f. of $\chi'^2_{n,\Lambda}$

	L_0					L_1				
	Degrees of freedom									
noncentrality	3	8	13	18	23	3	8	13	18	23
2	4, 4, 5	5, 4, 4	3, 4, 4	4, 3, 3	3, 3, 4	5, 5, 5	5, 5, 5	5, 5, 5	5, 5, 5	5, 5, 5
7	4, 4, 5	3, 3, 4	4, 3, 4	3, 3, 3	3, 3, 4	3, 4, 5	4, 4, 5	5, 4, 5	5, 5, 5	5, 5, 5
14	4, 3, 5	4, 2, 5	3, 3, 3	2, 3, 3	2, 3, 3	3, 2, 2	4, 3, 3	3, 3, 4	4, 4, 4	4, 4, 4
21	3, 2, 3	2, 2, 3	2, 2, 3	2, 3, 3	2, 3, 3	3, 2, 2	4, 3, 3	3, 3, 4	4, 4, 4	4, 4, 4
	L_2					L_3				
2	1, 1, 2	2, 2, 2	4, 3, 3	4, 4, 4	4, 4, 5	1, 1, 1	2, 2, 2	4, 3, 3	4, 4, 4	4, 4, 5
7	2, 1, 1	2, 2, 2	4, 3, 2	3, 4, 3	4, 3, 5	2, 2, 1	2, 2, 2	4, 3, 3	4, 4, 4	4, 4, 5
14	1, 1, 1	2, 2, 2	3, 2, 3	3, 3, 4	4, 3, 4	2, 2, 1	3, 3, 4	5, 3, 4	5, 4, 4	4, 4, 4
21	1, 1, 1	2, 2, 2	2, 2, 3	3, 3, 4	3, 3, 4	3, 3, 4	4, 4, 4	4, 4, 4	5, 5, 5	4, 4, 5
	E					χ^2				
2	1, 2, 1	1, 4, 1	1, 4, 3	2, 5, 2	2, 5, 3	4, 3, 4	4, 3, 4	3, 3, 4	3, 2, 2	2, 2, 3
7	1, 5, 1	2, 5, 2	2, 5, 4	2, 5, 3	3, 5, 4	3, 4, 4	3, 3, 3	3, 2, 3	2, 2, 2	2, 2, 3
14	2, 5, 2	3, 5, 4	5, 5, 5	5, 5, 5	4, 5, 4	4, 3, 3	4, 2, 4	2, 2, 2	2, 2, 2	2, 2, 2
21	5, 5, 5	5, 5, 5	5, 5, 5	5, 5, 5	5, 5, 5	3, 2, 3	3, 2, 3	2, 2, 2	2, 3, 3	2, 2, 2
	P									
2	5, 2, 2	5, 2, 1	4, 2, 1	3, 1, 1	1, 1, 1					
7	5, 2, 1	5, 2, 1	4, 1, 1	2, 1, 1	1, 1, 1					
14	5, 2, 1	5, 2, 1	4, 1, 1	2, 1, 1	1, 1, 1					
21	4, 1, 1	4, 1, 1	4, 1, 1	2, 1, 1	1, 1, 1					

Table entries: there are 3 numbers for each degree of freedom and noncentrality parameter; (left, center, right) tail comparisons where 5 = best, $\cdots$, 1 = worst

and large Λ, the E-series is good if Λ is very large, but L_0 is much preferred if Λ is only moderately large. The series L_2, L_3 are very poor in this latter range. A word of caution for the L_3-series in the left-hand tail of the distribution, is that for x_1 values such that $x_1 = x + \theta < 0$, the series leads to negative probability values. Although not indicated in the table, the series L_0 was evaluated with other values of λ keeping $\alpha + 1 = n/2$. It was found that for Λ large relative to n, values of $\lambda = 1/3$ and $\lambda = 1/4$ are quite effective; that is, as Λ increases, a decrease in λ improves the convergence rate of the L_0-expansion.

Prologue to Table 1. *Description of and instructions for using Table* 1. *Codes for entries* (K_1, K_2, K_3) *in Table* 1.

$K_i = 1, 2, 3, 4,$ or 5 for $i = 1, 2, 3$.

K_1 represents left tail, that is, evaluated probabilities are in the interval (0, .2).

K_2 represents center of distribution (.2, .8).

K_3 right tail (.8, 1.0).

The number codes 1 through 5 signify how well the series behave on a five point scale which is described as follows:

5 means *very best* among the series compared. Usually 5 place decimal accuracy is achieved with only a few terms. In those cases, however, where it is difficult to achieve 5 place decimal accuracy for all the series, code number 5 again indicates the best, but more than a few terms of the series are required to achieve 5 place decimal accuracy.

4 indicates *very good*, but not best; that is, 3, 4, or 5 place decimal accuracy can be attained after a few terms in the series, but more terms are necessary than for code number 5.

3 indicates *good*, but to obtain 3, 4, or 5 place decimal accuracy, a larger number of terms is required than for code number 4.

2 indicates *fair*, that is, only 3 or 4 place decimal accuracy can be obtained after a moderate number of terms, but more terms are necessary than for code number 3.

1 indicates *poor*, that is, 3 or 4 place decimal accuracy can not be obtained after a moderate number of terms, or the series diverges before even two place accuracy can be attained.

Categories in Table 1. For each of the seven series the ranges of parameter n and Λ are as follows. The degrees of freedom n are divided into five intervals: $2 \leqq n \leqq 5$, $5 < n \leqq 10$, $10 < n \leqq 15$, $15 < n \leqq 20$, $20 < n \leqq 25$. The approximate centers of these intervals 3, 8, 13, 18, 23 appear as headings for the columns of the table. Thus, for example, the entry (1, 2, 1) in the upper left corner of the table under the Edgeworth-series corresponds to the interval $2 \leqq n \leqq 5$, and is under the column labeled 3.

The values of the noncentrality parameter Λ are divided into four intervals $0 \leqq \Lambda \leqq 5$, $5 < \Lambda \leqq 10$, $10 < \Lambda \leqq 18$, $18 < \Lambda \leqq 25$. The approximate centers of these intervals 2, 7, 14, 21 appear as headings for the rows of the table. Thus the aforementioned entry (1, 2, 1) for the Edgeworth series corresponds to a noncentrality parameter in the interval $0 \leqq \Lambda \leqq 5$.

Comparison between series. Since for each of the series compared there are five categories for degrees of freedom and four categories for noncentrality, there

will be 20 classes in which the various series can be compared. For example, one such class is for $n = 8$ and $\Lambda = 14$. In this class, (K_1, K_2, K_3) appears as (4, 2, 5) for the series L_0, but (3, 5, 4) for the series E. Thus we see that the L_0–series is better in the left tail, but the Edgeworth is much preferred in the center areas, and both are good in the right tail with the L_0-expansions slightly better.

7. Conclusion. Table 1 is intended as an aid in obtaining probabilities for $\chi'^2_{n,\Lambda}$ efficiently. Each entry in the table was the result of the comparison of a large number of evaluations, and enough probabilities were computed in order to discern clearly the pattern of the table. It should be pointed out that the first term of the L_2- and L_3-expansions are the well-known Patnaik and Pearson approximations, respectively, and the first term of the Edgeworth-series is the simple first two moment fit of the standardized Gaussian distribution. The Laguerrian expansions resulting from (4.5) are set up in such a manner that other values of $\alpha + 1$, θ, and λ may be inserted and the resulting expansion evaluated for its efficiency. The table can be used as a starting point for choosing values of $\alpha + 1$, θ, and λ, in order to obtain probabilities efficiently for particular values of n, Λ and x.

REFERENCES

[1] S. H. ABDEL ATY, *Approximate formula for the percentage points and the probability integral of the non-central χ^2*, Biometrika, 41 (1954), pp. 538–540.

[2] H. CRAMER, *Mathematical Methods of Statistics*, Princeton Univ. Press, Princeton, N.J., 1946.

[3] R. B. DAVIES, *Numerical inversion of a characteristic function*, Biometrika, 60 (1973), pp. 415–417.

[4] R. A. FISHER, *The general sampling distribution of the multiple correlation coefficient*, Proc. Roy. Soc. London, 121A (1928), pp. 654–673..

[5] E. FIX, *Tables of non-central χ^2*, Univ. of Calif. Publications in Statistics, Univ. of Calif. Press, Berkeley, 1 (1949), no. 2, pp. 15–19.

[6] W. FREIDBERGER AND R. H. JONES, *Computation of the frequency function of a quadratic form in random normal deviates*, J. Assoc. Comp. Mach., 7 (1960), pp. 245–250.

[7] J. GURLAND, *Distribution of quadratic forms and ratios of quadratic forms*, Ann. Math. Statist., 24 (1953), pp. 416–427.

[8] ———, *Distribution of definite and of indefinite quadratic forms*, Ibid., 26 (1955), pp. 122–127.

[9] ———, *Quadratic forms in normally distributed random variables*, Sankhyā, 17 (1956), pp. 37–50.

[10] J. P. IMHOF, *Computing the distribution of quadratic forms in normal variables*, Biometrika, 48 (1961), pp. 418–426.

[11] D. R. JENSEN AND H. SOLOMON, *A Gaussian approximation to the distribution of a definite quadratic form*, J. Amer. Stat. Assoc., 67 (1972), pp. 898–902.

[12] G. C. KHATRI, *Distribution of a quadratic form in normal vectors*, paper presented at the Calgary Conf. on Statistical Distributions, Calgary, Alberta, Canada, 1974.

[13] S. KOTZ, N. L. JOHNSON AND D. W. BOYD, *Series representations of distributions of quadratic forms in normal variables. I: Central case*, Ann. Math. Statist., 38 (1967), pp. 823–837.

[14] ———, *Series representations of distributions of quadratic forms in normal variables II: Non-central case*, Ibid., 38 (1967), pp. 838–848.

[15] S. KOTZ AND N. L. JOHNSON, *Continuous Univariate Distributions*, vol. 2, John Wiley, New York, 1970.

[16] P. B. PATNAIK, *The non-central χ^2 and F distributions and their applications*, Biometrika, 36 (1949), pp. 202–232.

[17] E. S. PEARSON, *Note on an approximation to the distribution of non-central χ^2*, Ibid., 46 (1959), p. 364.

[18] H. RUBEN, *A new result on the distribution of quadratic forms*, Ann. Math. Statist., 34 (1963), pp. 1582–1583.

[19] ———, *Non-central chi-square and gamma revisited*, Comm. Statist., 3 (1974), pp. 607–633.
[20] M. SANKARAN, *Approximations to the non-central χ^2 distribution*, Biometrika, 50 (1963), pp. 199–204.
[21] G. A. SEBER, *The non-central χ^2 and Beta distributions*, Ibid., 50 (1963), pp. 542–544.
[22] M. L. TIKU, *Laguerre series forms of non-central χ^2 and F distributions*, Ibid., 52 (1965), pp. 415–427.

SIAM J. MATH. ANAL.
Vol. 8, No. 1, February 1977

UNIQUENESS OF SOLUTIONS TO SINGULAR BOUNDARY VALUE PROBLEMS*

C. C. TRAVIS† AND E. C. YOUNG‡

Abstract. The paper is concerned with the uniqueness of solutions to non-well-posed singular boundary value problems for partial differential equations. A method of proof is employed which is completely independent of the type of the partial differential equation being considered, and which allows the consideration of a class of nonstandard boundary conditions.

1. Introduction. The classical work of Bourgin and Duffin [1] concerning the uniqueness of solutions of the Dirichlet and Neumann problems for the wave equation $u_{tt}-u_{xx}=0$ in a rectangle with sides parallel to the coordinate axis has been extended in recent years to various normal and singular hyperbolic and ultrahyperbolic partial differential equations. See, for instance, [2]–[6]. All of these papers employ essentially the same method of proof, namely, an energy integral argument together with the assumption that a complete set of eigenfunctions exists for an associated eigenvalue problem. More recently, by using an entirely different approach, Travis [7] considered the singular hyperbolic boundary value problem studied in [4] and obtained more general results under weaker assumptions. The object of this paper is to employ the method of [7] to investigate the question of uniqueness of solutions for the singular equation

$$\sum_{i=1}^{n}\left(u_{x_ix_i}+\frac{\alpha_i}{x_i}u_{x_i}\right)-\sum_{j,k=1}^{m}(a_{jk}u_{y_j})_{y_k}+cu=0, \tag{1.1}$$

under a more general class of boundary conditions than studied by Young [6]. We shall consider (1.1) in the domain $D=H\times G$, where $H=\{(x_1,\cdots,x_n)\in E^n|0<x_i<a_i;\ 1\leqq i\leqq n\}$ and G is a bounded regular domain in E^m. The α_i $(1\leqq i\leqq n)$ are real parameters, $-\infty<\alpha_i<\infty$, and the coefficients a_{jk}, c are functions of the variables $y_1, y_2,\cdots, y_m$ only. On the boundary $H\times\partial G$, we prescribe the boundary condition

$$\partial u/\partial n+\sigma u=0, \tag{1.2}$$

where $\partial u/\partial n$ denotes the transversal derivative

$$\frac{\partial u}{\partial n}=\sum_{i,j=1}^{m}a_{ij},u_{y_i}\nu_j,$$

$(\nu_1,\cdots,\nu_m)$ being the outward unit normal vector on ∂G and σ being a continuous function of $y_i,\cdots,y_m$ on ∂G. Additional boundary conditions on $\partial H\times G$ will be given later.

The study of the singular equation (1.1) raises some interesting questions which are not present in the nonsingular case. In particular, there is the question as

* Received by the editors May 20, 1975, and in revised form September 15, 1975.

† Department of Mathematics, University of Tennessee, Knoxville, Tennessee 37916.

‡ Department of Mathematics, Florida State University, Tallahassee, Florida 32306.

to what is the proper boundary condition to be imposed on the solutions along the singular planes $x_i = 0$ $(1 \leqq i \leqq n)$. Previous work in this connection (see [4], [6]) requires the solutions to vanish on these planes and to be class C^2 in D and C^1 in $\bar{D}$. In this paper we shall replace these assumptions with much weaker conditions that allow the solution to be infinite on the singular planes. This relaxation enables us to impose less stringent differentiability conditions on the solutions. For brevity, we shall let $x = (x_1, \cdots, x_n)$ denote a point in H and $y = (y_1, \cdots, y_m)$ a point in G.

In obtaining our uniqueness results, we shall have need of the associated eigenvalue problem

$$
\begin{aligned}
&-\sum_{i,j=1}^{m} (a_{ij} v_{y_i})_{y_j} + cv = \lambda v \quad \text{in } G, \\
&\frac{\partial v}{\partial n} + \sigma v = 0 \qquad\qquad\qquad \text{on } \partial G.
\end{aligned} \tag{1.3}
$$

In contrast to previous works cited above, which assume that the equation in (1.3) is elliptic and self-adjoint and that $c \geqq 0$ in G, $\sigma \geqq 0$ on ∂G, here we shall assume only that the eigenvalue problem (1.3) has a complete set of eigenfunctions. We make no assumption about the symmetry or definite positivity of the matrix (a_{ij}), nor do we assume that all the eigenvalues of (1.3) are of the same sign.

2. Preliminary lemmas and conditions. In this section, we shall present certain lemmas and conditions that will be needed in the statement and proof of our uniqueness results. For convenience, we shall write $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ and $\int_0^a f\, dx = \int_0^{a_1} \cdots \int_0^{a_n} f\, dx_1 \cdots dx_n$.

LEMMA 2.1. *Let $\alpha_i \leqq -1$ for $i = 1, \cdots, n$. The eigenvalues and eigenfunctions of the singular eigenvalue problem*

$$
\begin{aligned}
&\sum_{i=1}^{n} (x^\alpha X_{x_i})_{x_i} + x^\alpha \mu X = 0 \quad \text{in } H, \\
&\int_0^a x^\alpha X^2\, dx < \infty, \\
&X(x) = 0 \quad \text{on } x_i = a_i, \qquad 1 \leqq i \leqq n,
\end{aligned} \tag{2.1}
$$

are given by $\mu = \mu_1 + \cdots + \mu_n$, where μ_i $(1 \leqq i \leqq n)$ is a root of the equation

$$
J_{|1-\alpha_i|/2}(\sqrt{\mu_i} a_i) = 0, \qquad 1 \leqq i \leqq n, \tag{2.2}
$$

and

$$
X(x) = \prod_{i=1}^{n} x_i^{(1-\alpha_i)/2} J_{(1-\alpha_i)/2}(\sqrt{\mu_i} x_i).
$$

Here $J_p(t)$ denotes the Bessel function of the first kind of order p.

LEMMA 2.2. *Let* $-1<\alpha_i$ *for* $i=1,\cdots,n$. *The eigenvalues and eigenfunctions of the singular eigenvalue problem*

$$\sum_{i=1}^{n}(x^{\alpha}X_{x_i})_{x_i}+x^{\alpha}\mu X=0 \quad \text{in } H,$$

$$\lim_{x_i\to 0} x_i^{\alpha_i}X_{x_i}=0, \qquad 1\leqq i\leqq n, \tag{2.3}$$

$$X(x)=0 \quad \text{on } x_i=a_i, \qquad 1\leqq i\leqq n,$$

are given by $\mu=\mu_1+\cdots+\mu_n$, *where* μ_i $(1\leqq i\leqq n)$ *satisfies* (2.2) *and*

$$X(x)=\prod_{i=1}^{n} x_i^{(1-\alpha_i)/2}J_{(\alpha_i-1)/2}(\sqrt{\mu_i}\,x_i).$$

Remarks. If the condition $X=0$ on $x_i=a_i$ $(1\leqq i\leqq n)$ in (2.1) and (2.3) is replaced by $X_{x_i}=0$ on $x_i=a_i$ $(1\leqq i\leqq n)$, then both Lemmas 2.1 and 2.2 remain valid provided the μ_i $(1\leqq i\leqq n)$ are roots of the equation

$$J_{|1+\alpha_i|/2}(\sqrt{\mu_i}\,a_i)=0, \qquad 1\leqq i\leqq n. \tag{2.4}$$

Both Lemmas 2.1 and 2.2, and these remarks can be easily established by the usual separation of variables argument.

To facilitate the statement of our results, we shall refer to the conditions that are used in the uniqueness theorems as

Condition A. Given any eigenvalue λ_k of the eigenvalue problem (1.3), and given any sequence of possibly complex constants μ_i $(1\leqq i\leqq n)$ satisfying $\mu_1+\cdots+\mu_n=\lambda_k$, there exists, a μ_r $(1\leqq r\leqq n)$ such that

$$J_{|1-\alpha_r|/2}(\sqrt{\mu_r}\,a_r)\neq 0; \tag{2.5}$$

Condition B. Under the same hypotheses as in Condition A, there exists a μ_r $(1\leqq r\leqq n)$ such that

$$J_{|1+\alpha_r|/2}(\sqrt{\mu_r}\,a_r)\neq 0. \tag{2.6}$$

3. Uniqueness results. We are now in a position to state and prove our uniqueness results.

THEOREM 3.1. *Let* $\alpha_i\leqq -1$ *for* $i=1,\cdots,n$, *and suppose that* $u\in C^2(H\times G)$ *is a solution of* (1.1) *satisfying* (1.2) *and the conditions*

$$\int_0^a\int_G x^{\alpha}|u(x,y)|^2\,dy\,dx<\infty,$$

$$u(x,y)=0 \quad \text{on } x_i=a_i, \qquad 1\leqq i\leqq n, \tag{3.1}$$

for all $y\in G$. *Then* $u\equiv 0$ *if and only if Condition* A *holds.*

Proof. Suppose Condition A is not satisfied. Then corresponding to an eigenvalue λ_k of the eigenvalue problem (1.3), there exist constants $\mu_1,\cdots,\mu_m$

satisfying $\mu_1 + \cdots + \mu_n = \lambda_k$ such that $J_{(1-\alpha_i)/2}(\sqrt{\mu_i}\, a_i) = 0$ $(i = 1, \cdots, n)$. Let v_k be an eigenfunction of the problem (1.3) corresponding to the eigenvalue λ_k. Then

$$u(x, y) = \prod_{i=1}^{n} x_i^{(1-\alpha_i)/2} J_{(1-\alpha_i)/2}(\sqrt{\mu_i}\, x_i) v_k(y)$$

is a nontrivial solution of the problem (3.1), as is readily verified.

Conversely, suppose Condition A holds. Let λ_k be an eigenvalue of the problem (1.3) with the corresponding normalized eigenfunctions $v_k(y)$. Let $\mu_1, \cdots, \mu_n$ be constants satisfying $\mu_1 + \cdots + \mu_m = \lambda_k$ for which (2.5) holds for some r, $1 \leqq r \leqq n$. Define

$$h_k(x) = \int_G u(x, y) v_k(y)\, dy,$$

where u is a solution of the problem (3.1). Then h_k satisfies the equation

$$\sum_{i=1}^{n} (x^\alpha h_{x_i})_{x_i} + x^\alpha \lambda_k h = 0$$

and the condition $h(x) = 0$ on $x_i = a_i$ for $i = 1, \cdots, n$. Moreover, by the Cauchy-Schwarz inequality and the fact that v_k is a normalized eigenfunction, it follows that

$$\int_0^a x^\alpha |h|^2\, dx \leqq \int_0^a \int_G x^\alpha |u(x, y)|^2\, dx\, dy < \infty.$$

Therefore h_k is a solution of the problem (2.1) corresponding to $\mu = \lambda_k$. However, in view of (2.5), $\lambda_k = \mu_1 + \cdots + \mu_n$ is not an eigenvalue of (2.1). Hence h_k must be a trivial solution, that is,

$$\int_G u(x, y) v_k(y)\, dy \equiv 0, \qquad k = 1, 2, \cdots.$$

By the completeness of the set of eigenfunctions v_k, it follows that $u \equiv 0$, as we wish to prove.

COROLLARY 3.1. *Let (a_{ij}) be symmetric and positive definite in $\bar{G}$ and let $c \leqq 0$ in G and $\sigma \geqq 0$ on ∂G. If $\alpha_i \leqq -1$, then every solution of the generalized* GASPT *equation*

$$\sum_{i=1}^{n} \left(u_{x_i x_i} + \frac{\alpha_i}{x_i} u_{x_i} \right) + \sum_{j,k=1}^{m} (a_{jk} u_{y_i})_{y_k} + cu = 0 \quad \text{in } H \times G \tag{3.2}$$

satisfying (1.2) *and vanishing on $x_i = a_i$ $(i = 1, \cdots, n)$ such that*

$$\int_0^a \int_G x |u(x, y)|^2\, dx\, dy < \infty$$

is identically zero.

Proof. From the assumptions on (a_{ij}), c and σ, it follows that the eigenvalues

λ_k of the associated eigenvalue problem

$$\sum_{j,k=1}^{m} (a_{jk}v_{y_j})_{y_k} + cv = \lambda v \quad \text{in } G$$

$$\frac{\partial v}{\partial n} + \sigma v = 0 \qquad \text{on } \partial G,$$

are negative. Thus, for any sequence of constants $\mu_1, \cdots, \mu_n$ such that $\mu_1 + \cdots + \mu_n = \lambda_k$, there is a μ_r $(1 \leqq r \leqq n)$ such that $\mu_r < 0$, and hence (2.5) holds. The conclusion now follows from Theorem 3.1.

THEOREM 3.2. *Let* $-1 < \alpha_i$ *for* $i = 1, \cdots, n$. *Then every solution* $u \in C^2(H \times G)$ *of* (1.1), (1.2) *satisfying the conditions*

$$\begin{aligned} &\lim_{x_i \to 0} x^{\alpha_i} u_{x_i} = 0, \qquad i = 1, \cdots, n, \\ &u(x, y) = 0 \quad \text{on } x_i = a_i, \qquad i = 1, \cdots, n, \end{aligned} \tag{3.3}$$

for $y \in G$ *is identically zero if and only if Condition* A *is satisfied.*

The proof of this theorem, which makes use of Lemma 2.2, is similar to that of Theorem 3.1 and is therefore omitted.

COROLLARY 3.2. *If* $-1 < \alpha_i$, *for* $i = 1, \cdots, n$, *then every solution* $u \in C^2(H \times G)$ *of* (3.2) *satisfying* (1.2) *and* (3.3) *is identically zero.*

If we replace the boundary condition $u = 0$ on $x_i = a_i$ $(1 \leqq i \leqq n)$ in Theorems 3.1 and 3.2 by $u_{x_i} = 0$ on $x_i = a_i$ $(1 \leqq i \leqq n)$, we obtain the following results:

THEOREM 3.3. *Let* $\alpha_i \leqq -1$ *for* $i = 1, \cdots, n$. *Then every solution* $u \in C^2(H \times G)$ *of* (1.1) *satisfying* (1.2) *and the conditions*

$$\begin{aligned} &\int_0^a \int_G x^{\alpha} |u(x, y)|^2 \, dx \, dy < \infty, \\ &u_{x_i}(x, y) = 0 \quad \text{on } x_i = a_i, \qquad i = 1, \cdots, n, \end{aligned}$$

for all $y \in G$, *vanishes identically if and only if Condition* B *holds.*

THEOREM 3.4. *Let* $-1 < \alpha_i$ *for* $i = 1, \cdots, n$. *Then every solution* $u \in C^2(H \times G)$ *of* (1.1) *satisfying* (1.2) *and the conditions*

$$\begin{aligned} &\lim_{x_i \to 0} x_i^{\alpha_i} u_{x_i} = 0, \qquad i = 1, \cdots, n, \\ &u_{x_i}(x, y) = 0 \quad \text{on } x_i = a_i, \qquad i = 1, \cdots, n, \end{aligned}$$

for all $y \in G$ *vanishes identically if and only if Condition* B *holds.*

The proofs of these theorems follow the same line of argument as those of Theorems 3.1 and 3.2, using the remarks made following Lemma 2.2. Obviously, corresponding theorems can also be stated for the generalized GASPT equation (3.2).

4. Other boundary value problems. In this section we consider uniqueness conditions for (1.1) subject to more classical boundary conditions that include the

ones treated in [6]. Because of our different approach, we are able to obtain more general results under weaker assumptions.

THEOREM 4.1. *If* $\alpha_i < 1$, *for* $i = 1, \cdots, n$, *then every solution* $u \in C^2(H \times G)$ *of* (1.1) *satisfying* (1.2) *and vanishing on* $x_i = 0$, $x_i = a_i$ $(1 \leqq i \leqq n)$ *is identically zero if and only if Condition* A *holds.*

The theorem is proved in the same manner as Theorem 3.1.

Remarks. We remark that in [6] a result corresponding to Theorem 4.1 was obtained under the stronger assumption that $u \in C^2(H \times G) \cap C^1(\bar{H} \times \bar{G})$; that result was valid only for $\alpha_i \leqq 0$ $(1 \leqq i \leqq n)$. It is interesting to note that Theorem 4.1 remains valid for all values of the parameters α_i if we impose the boundary conditions $u = 0$ on $x_i = a_i$ $(1 \leqq i \leqq n)$ and $|u(0, y)| < \infty$ for every $y \in G$. A corresponding result involving the generalized GASPT equation (3.2) can also be stated. We conclude this paper by stating a uniqueness result corresponding to the Neumann boundary conditions

$$u_{x_i} = 0 \quad \text{on } x_i = 0 \quad \text{and} \quad x_i = a_i, \qquad 1 \leqq i \leqq n. \tag{4.1}$$

This result, too, is independent of the value of the parameter α_i.

THEOREM 4.2. *For any value of* α_i $(1 \leqq i \leqq n)$, *every solution* $u \in C^2(H \times G)$ *of* (1.1) *satisfying* (1.2) *and* (4.1) *is identically zero, or* $u = \text{const.}$ *if* $c \equiv 0$ *and* $\sigma \equiv 0$, *if and only if Condition* B *is satisfied.*

We observe that in the case when $c \equiv 0$ and $\sigma \equiv 0$, the eigenvalue problem (1.3) has a zero eigenvalue with the corresponding eigenfunction $v_0 = \text{const.}$ Thus the result $h_k(x) = \int_G u(x, y) v_k(y)\, dy = 0$ $(k = 1, 2, \cdots)$ in the proof implies only that u is a nonzero constant. Theorem 4.2 extends the result in [8] to all values of the parameter α_i, instead of $0 \leqq \alpha_i$ only.

5. Concluding remarks. When $m = n = 1$ and $\alpha = 0$, Theorems 3.4 and 4.1 reduce to the classical results concerning the uniqueness of solutions of the Neumann and Dirichlet problems ($\sigma = 0$ and $\sigma = \infty$ in (1.2), respectively) for the one-dimensional wave equation $u_{xx} - u_{yy} = 0$ in the rectangular region $\{x | 0 < x < a\} \times \{y | 0 < y < b\}$. The eigenvalues of the problem (1.3), in this one-dimensional case, are given by $\lambda_m = (m\pi)^2 / b^2$, where $m = 0, 1, 2, \cdots$, for $\sigma = 0$ and $\lambda_m = m$, where $m = 1, 2, \cdots$, for $\sigma = \infty$. Conditions B and A of Theorems 3.4 and 4.1 then reduce to the condition $\sin(m\pi a/b) \neq 0$, which leads to the classical result that the ratio a/b is irrational.

It is clear that the method employed in this paper is completely independent of the type of the partial differential equation being considered. Thus it provides an alternative method of establishing uniqueness theorems that is more general than the usual energy integral approach.

REFERENCES

[1] D. G. BOURGIN AND R. DUFFIN, *The Dirichlet problem for the vibrating string equation*, Bull. Amer. Math. Soc., 45 (1939), pp. 851–858.

[2] D. R. DUNNINGER AND E. C. ZACHMANOGLOU, *The condition for uniqueness of solutions of the Dirichlet problem for the wave equation in coordinate rectangles*, J. Math. Anal. Appl., 20 (1967), pp. 17–21.

[3] ———, *The condition for uniqueness of the Dirichlet problem for hyperbolic equations in cylindrical domains*, J. Math. Mech., 18 (1969), pp. 763–766.

[4] E. C. YOUNG, *Uniqueness of solutions of the Dirichlet and Neumann problems for hyperbolic equations*, Trans. Amer. Math. Soc., 160 (1971), pp. 402–408.

[5] J. B. DIAZ AND E. C. YOUNG, *Uniqueness of solutions of certain boundary value problems for ultrahyperbolic equations*, Proc. Amer. Math. Soc., 29 (1971), pp. 569–574.

[6] E. C. YOUNG, *Uniqueness of solutions of the Dirichlet problem for singular ultrahyperbolic equations*, Proc. Amer. Math. Soc., 36 (1972), pp. 130–136.

[7] C. C. TRAVIS, *On the uniqueness of solutions to hyperbolic boundary value problems*, Trans. Amer. Math. Soc., 216 (1976), pp. 327–336.

[8] E. C. YOUNG, *Uniqueness of solutions of the Neumann problem for singular ultrahyperbolic equations*, Portugal. Math., to appear.

SIAM J. MATH. ANAL.
Vol. 8, No. 1, February 1977

THE GEOMETRY OF INDEFINITE J-SPACES AND STRONG STABILITY CRITERIA OF CANONICAL DIFFERENTIAL EQUATIONS WITH PERIODIC COEFFICIENTS*

KUO-LIANG CHIOU†

Abstract. In this paper we shall use the general results in J-unitary and related operators to study equation $dX(t)/dt = iJH(t)X(t)$ and various special cases. We obtain several stability criteria which generalize some cases treated by Lyapunov [9], Borg [2], Kreĭn [7], Gohberg and Kreĭn [6], Brockett [3] and Daleckiĭ and Kreĭn [8, Thm. 5.3].

1. Introduction. In this paper we consider stability criteria for the following linear system with periodic coefficient of period $T(>0)$:

$$\frac{dX(t)}{dt} = iJH(t)X(t), \qquad 0<t<\infty. \tag{1.1}$$

We assume that $H^*(t) = H(t) = H(t+T)$, $J = J^*$, $J^2 = I$ and that $H(t)$ is a function with values in the ring R of bounded operators acting on a Hilbert space S. Here A^* denotes the adjoint operator of A. For the sake of simplicity, we will always assume that the above properties hold for all differential equations in this paper. We shall define stability and strong stability in the sense of Daleckiĭ and Kreĭn [8, p. 220] as follows:

DEFINITION.

(i) The system (1.1) is called *stable* if all its solutions are bounded as $t \to \infty$.

(ii) The system (1.1) is called *strongly stable* if it is stable and if there exists a $\varepsilon > 0$ such that for any real symmetric bounded operator satisfying the condition

$$\tilde{H}(t+T) = \tilde{H}(t) \quad \text{and} \quad \int_0^T \|\tilde{H}(t) - H(t)\|\, dt < \varepsilon,$$

all solutions of the system

$$\frac{dX(t)}{dt} = iJ\tilde{H}(t)X(t)$$

are bounded as $t \to \infty$. Here $\|\cdot\|$ denotes the induced operator norm.

In this paper, we shall use the general results in J-unitary and related operators to study (1.1) and various special cases to obtain several stability criteria which generalize results of Lyapunov [9], Borg [2], Kreĭn [7], Gohberg and Kreĭn [6], Brockett [3], and Daleckiĭ and Kreĭn [8, Thm. 5.3]. Related questions are discussed in [5], [6], [8] and [12].

2. The geometry of indefinite J-spaces and differential equations. Let us recall some facts from the theory of linear differential equations with real

* Received by the editors May 15, 1975, and in revised form October 29, 1975.

† Division of Engineering and Applied Physics, Harvard University, Cambridge, Massachusetts. Now at Department of Mathematics, Wayne State University, Detroit, Michigan 48202. This work was supported by the U.S. Office of Naval Research under the Joint Services Electronics Program by Contract N00014-75-C-0648 at the Division of Engineering and Applied Physics, Harvard University, Cambridge, Massachusetts.

parameter λ having the form

$$\frac{dX(t)}{dt} = i\lambda JH(t)X(t), \qquad 0 \leqq t < \infty, \tag{2.1}$$

where $H(t)$ and J have been defined above. Denote by $U(t;\lambda)$ the operator satisfying (2.1) with initial condition $U(0;\lambda) = I$. Since $(d/dt)[U^*(t;\lambda) \cdot JU(t;\lambda)] = 0$ and $U(0;\lambda) = I$, we see that $U^*(t;\lambda)JU(t;\lambda) = J$. Such operators are called J-unitary operators. On the other hand, if we take for any positive integer n,

$$iJ = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}, \qquad I_n = \text{the identity matrix with dimension } n,$$

then $iJH(t) \in \mathrm{sp}(n)$ (Hamiltonian Lie algebra) and $U(t;\lambda) \in \mathrm{Sp}(n)$ (Symplectic group). From the uniqueness of the solutions of (2.1) we obtain $U(t+nT;\lambda) = U(t;\lambda)U^n(T;\lambda)$ with $0 \leqq t \leqq T$ and n a positive integer. This suggests that we study the bounded solutions of (2.1) by considering the spectrum of the monodromy operator of (2.1), $U(T;\lambda)$. We shall define stability and strong stability of the J-unitary operator U in the sense of Daleckiĭ and Kreĭn [8] as follows.

DEFINITION.

(i) A J-unitary operator U is said to be *stable* if

$$\|U^n\| \leqq c, \qquad n = 1, 2, \cdots, \quad c = \text{constant}.$$

(ii) A J-unitary operator U is said to be *strongly stable* if there exists $\varepsilon > 0$ such that all J-unitary operators U_1 satisfying $\|U_1 - U\| < \varepsilon$ are stable.

If we use the equation $U(t;\lambda)x(0) = x(t)$ and the principle of uniform boundedness, we see that in order for any solution $x(t)$ of (2.1) to be stable (strongly stable) on $0 \leqq t < \infty$, it is necessary and sufficient that $U(T;\lambda)$ is stable (strongly stable).

Let us call $L^{[\perp]} = \{x \in S \mid (Jx, y) = 0 \ \forall y \in L\}$, the J-orthogonal complement of L, and $\sigma(A) = \{\lambda \mid (A - \lambda I)$ does not have a continuous inverse$\}$, the spectrum of A. We also say that L is J-nonnegative if $(Jx, x) \geqq 0 \ \forall x \in L$ and J-nonpositive if $(Jx, x) \leqq 0 \ \forall x \in L$.

DEFINITION. A real number $\lambda = \lambda_0$ is called a *point of strong stability* of (2.1) if for $\lambda = \lambda_0$ all of the solutions of (2.1) are strongly stable.

DEFINITION. By the *spectrum* of the boundary value problem

$$\frac{dX(t)}{dt} = i\lambda JH(t)X(t), \qquad X(0) + X(T) = 0, \tag{2.2}$$

we mean the closed set Ω consisting of those λ for which the operator $U(T;\lambda) + I$ does not have a continuous inverse. We let λ_{-1} and λ_1 denote, respectively, the maximal negative and minimal positive points of the spectrum of problem (2.2).

If we let

$$iJ = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \quad \text{and} \quad H(t) = \begin{pmatrix} B(t) & 0 \\ 0 & A(t) \end{pmatrix}, \tag{2.3}$$

then (2.1) becomes

$$\frac{dX(t)}{dt}=\lambda\begin{pmatrix}0 & A(t)\\ -B(t) & 0\end{pmatrix}X(t), \qquad 0\leqq t<\infty, \tag{2.4}$$

with $A^*(t)=A(t)=A(t+T)$, and $B^*(t)=B(t)=B(t+T)$.

The main result of this section is the following.

THEOREM 2.1. *Suppose that*

$$\text{(i)} \qquad \int_0^T H(s)\,ds \gg 0, \qquad H(t)=\begin{pmatrix}B(t) & 0\\ 0 & A(t)\end{pmatrix}$$

and

(ii) *either* $A(t)\geqq 0$, $0\leqq t\leqq T$, *or* $B(t)\geqq 0$, $0\leqq t\leqq T$. *Then all the solutions of* (2.4) *are strongly stable whenever* $0<\lambda<\lambda_1$.

Remark. It is relevant to note that Daleckiĭ and Kreĭn [8, Thm. 4.4, p. 224] have proved the following theorem.

THEOREM (Daleckiĭ and Kreĭn). *Suppose that*

$$H(t)\geqq 0, \qquad 0\leqq t\leqq T, \quad \int_0^T H(s)\,ds \gg 0.$$

Then all solutions of (2.1) *are strongly stable whenever*

$$\lambda_{-1}<\lambda<\lambda_1, \qquad \lambda\neq 0.$$

Therefore Theorem 2.1 is a generalization of the above theorem in the case of equation (2.4). Before proving Theorem 2.1 let us recall and prove some theorems.

THEOREM 2.2 [8, p. 221]. *In order for a J-unitary operator U to be strongly stable it is necessary and sufficient that S decomposes into a direct sum of two J-orthogonal subspaces* L_1 *and* L_2;

$$S=L_1[+]L_2,$$

where the subspaces L_1 *and* L_2 *are invariant with respect to U,* $\sigma(U|L_1)\cap\sigma(U|L_2)=\phi$, *and* L_1 *is J-nonnegative and* L_2 *is J-nonpositive. This decomposition is unique.*

THEOREM 2.3 [8, p. 221]. *If the operator*

$$\int_0^T H(s)\,ds$$

is uniformly positive (denoted by $\int_0^T H(s)\,ds\gg 0$), *then there exists a* $\delta>0$ *such that all points* λ, $-\delta<\lambda<\delta$, $\lambda\neq 0$, *are points of strong stability of* (2.1).

LEMMA 1 [8, p. 222]. *Suppose that* $U(T,\lambda)$ *is the monodromy operator of* (2.4) *for* $0<\lambda<\lambda_1$. *Then*

(i) *the spectrum of* $U(T;\lambda)$ *is on the unit circle, and*

(ii) *if* $V(\lambda)=-i[U(T;\lambda)-I][U(T;\lambda)+I]^{-1}$, *then*

$$V(\lambda)=JN(\lambda),$$

where

$$N(\lambda)=2\int_0^{\lambda}[U^*(T;\mu)+I]^{-1}R(\mu)[U(T;\lambda)+I]^{-1}\,d\mu$$

and

$$R(\lambda)=\int_0^T U^*(s;\lambda)H(s)U(s;\lambda)\,ds.$$

LEMMA 2. *Under the assumption of Theorem* 2.1, *if there exists a* λ_0, $0<\lambda_0<\lambda_1$, *which is a point of strong stability of* (2.4), *then* $R(\lambda_0)\geqq 0$ (*nonnegative definite*) *where* $R(\lambda)$ *is defined above.*

Proof. For any $X_0\in S$

$$X_0R(\lambda_0)X_0=\int_0^T X_0^*U^*(s;\lambda_0)H(s)U(s;\lambda_0)X_0\,ds$$

$$=\int_0^T X^*(s)H(s)X(s)\,ds,$$

where $X(t)$ is the solution of (2.4) with initial condition $X(0)=X_0$. Let

$$X^*(t)=(Y^*(t),Z^*(t)).$$

Then from (2.4) we obtain

$$X_0^*R(\lambda_0)X_0=\int_0^T (Y^*(s),Z^*(s))\begin{pmatrix}B(s)&0\\0&A(s)\end{pmatrix}\begin{pmatrix}Y(s)\\Z(s)\end{pmatrix}ds$$

$$=\frac{1}{\lambda}(Z_0^*Y_0^*-Z^*(T)Y^*(T))+2\int_0^T Z^*(s)A(s)Z(s)\,ds \tag{2.5}$$

$$=\frac{1}{\lambda}(Z^*(T)Y(T)-Z_0^*Y_0^*)+2\int_0^T Y^*(s)B(s)Y(s)\,ds. \tag{2.6}$$

Daleckiĭ and Kreĭn [8, p. 50] have shown that $U(T;\lambda_0)$ can be represented in the form

$$U(T;\lambda_0)=\begin{pmatrix}\exp(iH_1)&0\\0&\exp(iH_1^*)\end{pmatrix}$$

for some H_1 because $U(T;\lambda_0)$ is strongly stable.

Since

$$U(T;\lambda_0)\begin{pmatrix}y_0\\Z_0\end{pmatrix}=\begin{pmatrix}Y(T)\\Z(T)\end{pmatrix},$$

we obtain

$$Y(T)=\exp(iH_1)Y(0)$$

and

$$Z(T)=\exp(iH_1^*)Z(0),$$

which implies that

$$Z^*(T)Y(T) = Z_0^* Y_0.$$

Therefore, (2.5) (or (2.6)) now becomes

$$X_0^* R(\lambda_0) X_0 = 2\int_0^T Z^*(s)A(s)Z(s)\,ds \left(= 2\int_0^T Y^*(s)B(s)Y(s)\,ds\right)$$
$$\geqq 0$$

because of the assumption (ii) of Theorem 2.1. Thus Lemma 2 is established.

Proof of Theorem 2.1. From Theorem 2.3 we know that there exists $\delta > 0$ such that all points λ, $0 < \lambda < \delta$, are points of strong stability of (2.1). Now we want to show that $\delta \geqq \lambda_1$; i.e., all points λ, $0 < \lambda < \lambda_1$, are points of strong stability of (2.1). Suppose to the contrary, that there exists λ_0, $0 < \lambda_0 < \lambda_1$, such that all points λ, $0 < \lambda < \lambda_0$, are points of strong stability and $\lambda = \lambda_0$ is not a point of strong stability. Since $U(T; \mu)$ is close to I for μ sufficiently small, $R(\mu) \gg 0$ and then $N(\mu) \gg 0$. From Lemma 2, we obtain that $N(\lambda) \gg 0$ for $0 < \lambda < \lambda_0$. Since $N(\lambda)$ is continuous with respect to λ and by definition of $N(\lambda)$, there exists a small $\varepsilon > 0$ such that $N(\lambda_0 + \varepsilon) \gg 0$, i.e., $N(\lambda) \gg 0$, $0 < \lambda < \lambda_0 + \varepsilon$. It will then follow from Lemma 1 and the properties of normally decomposable operators (see Daleckiĭ and Kreĭn [8, § I.8.2]) that $V(\lambda)$ is normally J-decomposable, $0 < \lambda < \lambda_0 + \varepsilon$. Since $U(T; \lambda)$ can be expressed in the form of a linear fractional transformation of $V(\lambda)$ it too is normally J-decomposable. In addition, this operator is J-unitary. By virtue of Daleckiĭ and Kreĭn's theorem (see Daleckiĭ and Kreĭn [8, Thm. I.8.3]), this operator will be strongly stable. This contradicts $\lambda = \lambda_0$ is not a point of strong stability. Therefore, Theorem 2.1 is now established.

3. Stability criteria. In this section we consider stability criteria of the following system:

$$\frac{d^2X(t)}{dt^2} + p(t)X(t) = 0, \qquad 0 \leqq t < \infty, \tag{3.1}$$

with $p^*(t) = p(t) = p(t+T)$. Let

$$Y(t) = \begin{pmatrix} X(t) \\ \dfrac{dx(t)}{dt} \end{pmatrix} \in S \oplus S,$$

where S is the Hilbert space. We may write (3.1) in the form of (1.1) as follows:

$$\frac{dY(t)}{dt} = \begin{pmatrix} 0 & I \\ -P(t) & 0 \end{pmatrix} Y(t), \qquad 0 \leqq t < \infty. \tag{3.2}$$

By letting $iJ = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$ and $H(t) = \begin{pmatrix} P(t) & 0 \\ 0 & I \end{pmatrix}$, we see that (3.2) has the same form as (1.1). (For details see Daleckĭ and Kreĭn [8, p. 225–226].)

Daleckiĭ and Kreĭn [8, Thm. 5.3] have obtained a stability criterion in the infinite-dimensional case which is a generalization of Lyapunov's criterion [9]. In their paper, $p(t)$ is positive over $[0, T]$ and the average of $p(t)$ over $[0, T]$ is

uniform positive. Here for the case $\beta = 1$ (see the following theorem) we only consider that the average of $p(t)$ over $[0, T]$ is uniformly positive. Thus we have a criterion which generalizes Borg's criteria [2] and Daleckiĭ and Kreĭn's criteria [8, Thm. 5.3] in some cases. We also consider the case $\beta > 1$. (For the case $\beta = 2$, c.f. Borg's criteria [1] for the 2-dimensional case.)

THEOREM 3.1. *All solutions of* (3.1) *are strongly stable whenever there exists* $\beta \geqq 1$ *such that*

$$\text{(i)} \qquad \int_0^T p(s)\,ds \gg 0,$$

(3.3)

$$\text{(ii)} \qquad \int_0^T \|p(s)\|^\beta\,ds < c,$$

where $\|\cdot\|$ *is the induced operator norm and*

$$\text{if } \beta = 1, \quad c = \frac{4}{T},$$

$$\text{if } \beta > 1, \quad c = \left(\frac{4}{T}\right)^\beta \left(\frac{\alpha+1}{T}\right)^{\beta-1} \quad \text{and} \quad \alpha = \frac{\beta}{\beta-1}.$$

Proof. If we show that the number λ_1 from Theorem 2.1 satisfies the estimate

$$\frac{1}{\lambda_1} < \frac{1}{c}\int_0^T \|p(s)\|^\beta\,ds,$$

then (3.3) will, accordingly, imply that $\lambda_1 > 1 = \lambda$, and the validity of Theorem 3.1 will follow from Theorem 2.1.

Consider the class $C(S)$ of all continuous functions $X(t)$ $(0 \leqq t \leqq T)$ with values in S. We define a norm

$$\|X\| = \max_{0 \leqq t \leqq T} \{\|X(t)\|\}.$$

Thus $\langle C(S), \|\cdot\| \rangle$ is a Banach space. Define the integral operator K from $C(S)$ into $C(S)$ by

$$(KX)(t) = \int_0^T g(t-s)p(s)X(s)\,ds,$$

where

$$g(t-s) = \begin{cases} \dfrac{T}{4} - \dfrac{t-s}{2}, & 0 < s < t, \\[2ex] \dfrac{T}{4} + \dfrac{t-s}{2}, & t < s < T. \end{cases}$$

Thus by Schwarz's inequality,

$$\|(KX)(t)\| \leqq \int_0^T |g(t-s)|\,\|p(s)\|\,\|X(s)\|\,ds$$
$$\leqq c^{-1/\beta}\left(\int_0^T \|p(s)\|^\beta \|X(s)\|^\beta\,ds\right)^{1/\beta}.$$

Therefore the induced operator K has the following estimate:

$$\|K\| = \sup_{\|x\|=1}\{\|KX\|\} \leqq c^{-1/\beta}\left(\int_0^T \|p(s)\|^\beta\,ds\right)^{1/\beta} \tag{3.4}$$
$$< 1.$$

Since $\lambda_1 \in \Omega$ and is a positive real number, there exists a nonzero $X(t)$ such that

$$(KX)(t) = \frac{1}{\lambda_1^2}X(t),$$

which is equivalent to $d^2X(t)/dt^2 = \lambda^2\begin{pmatrix} 0 & I \\ -p(t) & 0 \end{pmatrix}X(t)$. Thus

$$\|K\| = \sup_{\|x\|\neq 0}\left\{\frac{\|(KX)\|}{\|x\|}\right\} > \frac{1}{\lambda_1^2}. \tag{3.5}$$

From (3.4) and (3.5) we obtain that $\lambda_1 > 1$. Thus Theorem 3.1 is established.

COROLLARY 3.1. *All solutions of* (3.1) *are strongly stable whenever*

(i) $$\int_0^T p(s)\,ds \gg 0,$$

(ii) $$\sup_{0\leqq t\leqq T}\|p(t)\| < \frac{8}{T^2}.$$

Proof. From the proof of Theorem 3.1 we know

$$\|(KX)(t)\| \leqq \int_0^T |g(t-s)|\,\|p(s)\|\,\|X(s)\|\,ds$$
$$\leqq \frac{T^2}{8}\|X(t)\| \sup_{0\leqq t\leqq T}\|p(t)\|$$
$$< \|X(t)\|.$$

Then we may follow the proof of Theorem 3.1 to obtain the desired conclusion.

Remark. For the scalar case (1-dimensional case) of (3.1), Zubovskii [10] proved that the solutions of

$$\frac{d^2x(t)}{dt^2} + a(t)x(t) = 0, \qquad a(t+T) = a(t), \qquad 0 < t < \infty,$$

where $a(t)$ is a piecewise continuous function, are all bounded provided $n^2\pi^2T^{-2} \leqq a(t) \leqq (n+1)^2\pi^2T^{-2}$ for all t and some $n = 0, 1, \cdots$. For the case of $n = 0$ we may compare the above Zubovskii result with Corollary 3.1.

In the following theorem we shall deal with (2.4) and obtain a stability criterion which is a generalization of Brockett's criterion [3] and some criteria in [11]. Brockett considers the case where $B(t) > 0$ over $[0, T]$. Here we assume the average of $B(t)$ over $[0, T]$ is uniformly positive.

THEOREM 3.2. *All solutions of* (2.4) *with* $\lambda = 1$ *are strongly stable whenever*

(i) *either* $A(t) > 0$, $0 \leqq t \leqq T$, *or* $B(t) \geqq 0$, $0 \leqq t \leqq T$,

(ii) $$\int_0^T A(s)\,ds \gg 0, \qquad \int_0^T B(s)\,ds \gg 0$$

and

(iii) $$\int_0^T \|A(s)\|\,ds \int_0^T \|B(s)\|\,ds < 4.$$

Proof. We note that (2.4) with $\lambda = 1$ is equivalent to

$$\frac{d}{dt}\left(A^{-1}(t)\frac{dy(t)}{dt}\right) + B(t)y(t) = 0.$$

We will follow the proof of Theorem 3.1. Define an operator K from $C(S)$ into $C(S)$ by

$$(KY)(t) = \frac{1}{4}\int_t^T A(s)\int_0^s B(z)Y(z)\,dz\,ds - \frac{1}{4}\int_t^T A(s)\int_s^T B(z)Y(z)\,dz\,ds$$
$$-\frac{1}{4}\int_0^t A(s)\int_0^s B(z)Y(z)\,dz\,ds + \frac{1}{4}\int_0^T A(s)\int_s^T B(z)Y(z)\,dz\,ds.$$

Thus the induced operator K has the estimate

$$\|K\| < \frac{1}{4}\int_0^T \|A(s)\|\,ds \int_0^T \|B(s)\|\,ds < 1.$$

Now we follow the proof of Theorem 3.1 to obtain $\lambda_1 > 1$. This proves Theorem 3.2.

Acknowledgment. I would like to thank R. W. Brockett for his encouragement, helpful suggestions and general interest in this work. I also would like to thank P. Crough for his helpful suggestions.

REFERENCES

[1] G. BORG, *Uber die Stabilitat gewisser Klassen von linear Differential-gleichunger*, Ark. Mat. Astronom. Fys., 31 (1944), p. 31.

[2] ———, *On a Lyapunov criterion of stability*, Amer. J. Math., 71 (1949), pp. 67–70.

[3] R. W. BROCKETT, *Stability of periodic linear systems and the geometry of Lie groups*, Dynamical Systems, L. Cesari, J. Hale, J. LaSalle, eds., Academic Press, New York, 1975, to appear.

[4] L. CESARI, *Asymptotic Behavior and Stability Problems in Ordinary Differential Equations*, Springer-Verlag, Berlin, 1959.

[5] I. M. GEL'FAND AND V. B. LIDSKII, *On the structure of the regions of stability of linear canonical systems of differential equations with periodic coefficients*, Uspehi Mat. Nauk, 10 (1955), pp. 3–40; English transl., Amer. Math. Soc. Transl. (2), 9 (1958), pp. 143–181.

[6] I. C. GOHBERG AND M. G. KREĬN, *Theory and Applications of Volterra Operators in Hilbert Space*, vol. 24, Trans. of Math. Monographs, American Mathematical Society, Providence, R.I., 1970.

[7] M. G. KREĬN, *Generalization of certain investigations of A. M. Lyapunov on linear differentia equations with periodic coefficients*, Dokl. Akad. Nauk SSSR, 73 (1950), pp. 445–448. (In Russian.)

[8] JU. L. DALECKIĬ AND M. G. KREĬN, *Stability of Solutions of Differential Equations in Banach Space*, vol. 43, Transl. of Math. Monographs, American Mathematical Society, Providence, R.I., 1974.

[9] A. LYAPUNOV, *Problème général de la stabilité du mouvement*, Comm. Soc. Math. Kharkov, 3 (1893), pp. 265–272.

[10] N. E. ZUBOVSKII, *Conditions for finiteness of integrals of equations* $y''+P(t)y=0$, Mat. Sb., 16 (1892), pp. 582–591.

[11] V. M. STARZINSKII, *A survey of works on the conditions of stability of the trivial solution of a system of linear differential equations with periodic coefficients*, Amer. Math. Soc. Transl. (2), 1 (1950), pp. 189–237.

[12] W. A. COPPEL AND A. HOWE, *On the stability of linear canonical systems with periodic coefficients*, J. Austral. Math. Soc., 5 (1965), pp. 161–195.

SIAM J. MATH. ANAL.
Vol. 8, No. 1, February 1977

CONNECTION FORMULAS FOR SECOND-ORDER DIFFERENTIAL EQUATIONS WITH MULTIPLE TURNING POINTS*

F. W. J. OLVER†

Abstract. A study is made of the differential equation

$$d^2w/dx^2 = \{u^2 f(u,x) + g(u,x)\}w, \qquad a_1 < x < a_2,$$

in which u is a positive parameter. For each value of u, $\partial^2 f(u,x)/\partial x^2$ and $g(u,x)$ are assumed to be continuous in the finite or infinite open interval (a_1, a_2). The function $f(u,x)$ is real and its only zero within (a_1, a_2) is a single zero of multiplicity $m-2$, where m ($\geqq 2$) is an arbitrary integer. For large values of u, asymptotic approximations for the solutions are constructed in terms of Bessel functions of order $1/m$, subject to certain restrictions on the behavior of $f(u,x)$ and $g(u,x)$ as $u \to \infty$. These restrictions are satisfied, for example, if $f(u,x)$ is independent of u and $g(u,x) = O(u^{\varpi})$, where $\varpi < \min(4/m, 1)$. Each approximation is uniformly valid throughout (a_1, a_2) and is accompanied by a strict and realistic bound for the error term.

In the case in which a_1 and a_2 are singularities of the differential equation, the uniform approximations are applied to solve the problem of connecting the known asymptotic solutions in terms of elementary functions (the Liouville–Green approximations) valid in a neighborhood of a_2, with the corresponding asymptotic solutions valid in a neighborhood of a_1.

1. Introduction and summary

1.1. Introduction. Consider the following two questions concerning a differential equation of the form

$$d^2w/dx^2 = f(x)w, \tag{1.01}$$

in which the independent variable x ranges over a finite or infinite interval (a_1, a_2),[1] and $f(x)$ is a real function that has singularities at a_1 and a_2, and is continuous within (a_1, a_2):

What is the nature of the solutions as x approaches the endpoints?

Given a solution with a certain behavior at one endpoint, how does this solution behave as x approaches the other endpoint?

A satisfactory answer to the first question is furnished by the theory of the Liouville–Green approximation given in [23, Chap. 6]. The general result is as follows. Assume that:

(i) *In the neighborhoods of a_1 and a_2, $f(x)$ is nonzero, $f''(x)$ is continuous, and the function*

$$F(x) \equiv \int \frac{1}{|f(x)|^{1/4}} \frac{d^2}{dx^2}\left(\frac{1}{|f(x)|^{1/4}}\right) dx$$

is of bounded variation.

---

* Received by the editors June 12, 1975, and in revised form October 17, 1975. 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.

† Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742, and National Bureau of Standards, Washington, D.C. 20234.

[1] Throughout, we adhere to the convention that (a, b) denotes the open interval $a < x < b$, and $[a, b]$ denotes the corresponding closed interval $a \leqq x \leqq b$; similarly for the partly closed intervals $(a, b]$ and $[a, b)$.

(ii)

$$\left|\int f^{1/2}(x)\,dx\right| \to \infty \quad \text{as } x \to a_1+ \quad \text{or} \quad a_2-.$$

Then equation (1.01) has solutions $w_1(x)$, $w_2(x)$, $w_3(x)$ and $w_4(x)$, such that

$$\begin{aligned} w_1(x) &\sim f^{-1/4}(x) \exp\left\{\int f^{1/2}(x)\,dx\right\}, \\ w_2(x) &\sim f^{-1/4}(x) \exp\left\{-\int f^{1/2}(x)\,dx\right\}, \end{aligned} \qquad x \to a_1+, \tag{1.02}$$

$$\begin{aligned} w_3(x) &\sim f^{-1/4}(x) \exp\left\{\int f^{1/2}(x)\,dx\right\}, \\ w_4(x) &\sim f^{-1/4}(x) \exp\left\{-\int f^{1/2}(x)\,dx\right\}, \end{aligned} \qquad x \to a_2-. \tag{1.03}$$

The following comments on this result are pertinent to the present investigation. First, the given conditions admit almost all kinds of singularities at a_1 and a_2, including most regular singularities and all irregular singularities of finite or infinite rank; see [23, pp. 197–202]. Secondly, the choice of integration constant in the indefinite integral of $f^{1/2}(x)$ appearing in (1.02) and (1.03) is unimportant, since it only affects the solutions by a constant factor. Thirdly, even when this constant of integration is fixed, the solutions may or may not be unique, but this is immaterial at the moment.

To answer the second question posed in the opening paragraph, it suffices to know the coefficients in two of the linear relations holding between any three of the four solutions, for example,

$$w_1(x) = A_1 w_3(x) + B_1 w_4(x), \qquad w_2(x) = A_2 w_3(x) + B_2 w_4(x).$$

These linear relations are called the *connection formulas* for the given differential system. Exact expressions for the coefficients A_1, B_1, A_2 and B_2 can be found only in rare cases, and in general approximate methods have to be employed. In certain physical problems, for example, barrier penetration and the approximate harmonic oscillator [10], [6], [3], $f(x)$ is twice continuously differentiable throughout (a_1, a_2) and contains a real parameter—which we shall denote by u^2—as a factor. Asymptotic approximations for the coefficients in the connection formulas are needed for large values of u. Of great importance in constructing these approximations are the number and multiplicities of the zeros of $f(x)$ within (a_1, a_2). A zero of $f(x)$ of multiplicity m is called a *turning point* (or *transition point*) *of* (1.01) *of multiplicity m.*

Satisfactory approximations, complete with strict error bounds, have been constructed for the solutions of the problem outlined in the preceding paragraph in cases when (a_1, a_2) contains at most two turning points, both of which are simple; see [3] and [23, Chap. 13]. The purpose of the present paper is to supply a similar theory when (a_1, a_2) contains a single turning point of arbitrary (integer) multiplicity. A subsequent paper [26] extends the analysis to the general case in which (a_1, a_2) contains an arbitrary number of turning points of arbitrary multiplicities.

1.2. Summary. The paper is arranged as follows.

In § 2 we consider the simplest form of differential equation having a multiple turning point, obtained by setting $f(x)$ equal to a constant factor times a power of x. Depending on the sign of the constant factor and whether the power is even or odd, three essentially distinct cases arise, which we designate I, II and III. With m denoting a positive integer such that $m \geqq 2$, our classification is given by:

Case I: $f(x) = \frac{1}{4}m^2x^{m-2}$, m even.

Case II: $f(x) = -\frac{1}{4}m^2x^{m-2}$, m even.

Case III: $f(x) = \frac{1}{4}m^2x^{m-2}$, m odd.

Thus in each case the multiplicity of the turning point at the origin is $m-2$. The normalizing factor $\pm\frac{1}{4}m^2$ has been introduced to simplify subsequent notation. In each case, the differential equation is solvable exactly in terms of Bessel functions (or modified Bessel functions) of order $1/m$. Also, in § 2 the correct choice of standard solutions is discussed, and relevant properties of the solutions are supplied, including exact connection formulas and illustrative graphs.[2]

In § 3 we study the equation

$$d^2w/dx^2 = \{u^2f(u, x) + g(u, x)\}w, \qquad a_1 < x < a_2, \tag{1.04}$$

in which u is a positive parameter, and $f(u, x)$ is nonvanishing within (a_1, a_2), except for a zero x_0, say, of multiplicity $m-2$. By means of an appropriate Liouville transformation, equation (1.04) is transformed into one of the forms

$$d^2W/d\zeta^2 = \{\pm\tfrac{1}{4}m^2u^2\zeta^{m-2} + \phi(u, \zeta)\}W,$$

in which the function $\phi(u, \zeta)$ is expressible in terms of $f(u, x)$ and $g(u, x)$. On neglecting $\phi(u, \zeta)$, the transformed equation is solvable in terms of the standard solutions of § 2. Strict error bounds are constructed for the approximations found in this way. The analysis in this section is very similar to that of [25], which treats the more general case in which the zero of $f(u, x)$ may have any multiplicity, integer or fractional. It is not possible to quote needed results directly from this reference however, because solutions are given directly in [25] for only one of the intervals $[x_0, a_2)$ or $(a_1, x_0]$. Moreover, solutions that comprise a numerically satisfactory pair[3] in the neighborhoods of x_0 and a_2 (or a_1), do not comprise a numerically satisfactory pair in the neighborhoods of a_1 and a_2.

In § 4 the results of §§ 2 and 3 are applied to derive the desired connection formulas for (1.04), complete with strict bounds for the error terms in the coefficients.

In § 5 the asymptotic nature of the error bounds of § 4 is explored in the commonly occurring case in which $f(u, x)$ is independent of u and $g(u, x) = O(u^{\varpi})$ as $u \to \infty$, uniformly with respect to x in compact intervals within (a_1, a_2), ϖ denoting a real constant. With suitable supplementary conditions, it is proved that each error bound vanishes as $u \to \infty$, provided that $\varpi < \min(4/m, 1)$. This establishes the asymptotic validity of the approximate connection formulas in these circumstances.

---

[2] A detailed investigation of solutions of the equation $d^2w/dz^2 = z^m w$ for integer values of m and complex values of z has been carried out by Swanson and Headley [28], and Headley and Barwell [7]. Solutions that are appropriate for complex z, however, are not necessarily the most suitable for real z; in consequence, these references do not supply all the results needed in the present paper.

[3] In the sense of Miller [19]; see also [23, pp. 154–155].

Lastly, in § 6 the work of other writers on the problem is described and related to the present analysis.

2. Standard solutions of the basic equations

2.1. Case I. The basic differential equation is given by

$$d^2w/dt^2 = \tfrac{1}{4}m^2t^{m-2}w, \tag{2.01}$$

where m is an even positive integer.[4] All solutions of this equation are infinitely differentiable when $t \in (-\infty, \infty)$. In defining standard solutions, we are guided by asymptotic behavior as $t \to \pm\infty$. From the theory of the Liouville–Green approximation given in [23, pp. 197–202], we know that there exists a unique solution that is asymptotic to $t^{(2-m)/4}\exp(-t^{m/2})$ as $t \to +\infty$. Since this solution is recessive compared with all linearly independent solutions in these circumstances, we adopt it as one of the standard solutions, and denote it by $U_m(t)$. Thus

$$U_m(t) \sim t^{(2-m)/4}\exp(-t^{m/2}), \qquad t \to +\infty, \tag{2.02}$$

the relative error in this relation being $O(t^{-m/2})$. It is also known from the theory of the Liouville–Green approximation that the last relation is differentiable; thus

$$U_m'(t) \sim -\tfrac{1}{2}mt^{(m-2)/4}\exp(-t^{m/2}), \qquad t \to +\infty, \tag{2.03}$$

again with relative error $O(t^{-m/2})$. Next, from (2.02), (2.03) and the differential equation (2.01), it follows that as t increases from $-\infty$ to $+\infty$, $U_m(t)$ is positive and decreasing, and $U_m'(t)$ is negative and increasing.[5] To determine the asymptotic behavior of $U_m(t)$ as $t \to -\infty$, as well as to derive other properties, we express $U_m(t)$ in terms of modified Bessel functions.

From the properties of Bessel functions collected in [21, Chap. 9], it is known that the general solution of (2.01) is $|t|^{1/2}\mathscr{Z}_{1/m}(|t|^{m/2})$, where $\mathscr{Z}_{1/m}$ denotes any modified cylinder function of order $1/m$. Examination of the asymptotic form of the modified Bessel functions for large positive argument yields the identification

$$U_m(t) = (2t/\pi)^{1/2}K_{1/m}(t^{m/2}), \qquad t > 0. \tag{2.04}$$

Letting $t \to 0$ in this relation and its differentiated form, we find that

$$\begin{aligned} U_m(0) &= \pi^{-1/2}2^{(2-m)/(2m)}\Gamma(1/m), \\ U_m'(0) &= \pi^{-1/2}2^{-(2+m)/(2m)}\Gamma(-1/m). \end{aligned} \tag{2.05}$$

These relations enable $U_m(t)$ to be identified in terms of modified Bessel functions when t is negative; thus

$$U_m(t) = (2|t|/\pi)^{1/2}\{\pi \csc(\pi/m)I_{1/m}(|t|^{m/2}) + K_{1/m}(|t|^{m/2})\}, \qquad t < 0. \tag{2.06}$$

---

[4] Although (2.01) has no turning point when $m = 2$, no complications arise on including this case in the subsequent analysis. In fact, because the solutions of (2.01) are elementary functions when $m = 2$, a useful check on the results becomes available. Similarly in Case II below.

[5] Let t_1 be the (algebraically) largest zero, if any, of $U_m(t)$. From (2.02) and graphical considerations, there exists $t_2 \in (t_1, \infty)$, such that $U_m'(t_2) = 0$. Referring to (2.03), however, we see that this is impossible, since from (2.01) and the fact that m is even it follows that $U_m''(t) \geqq 0$ when $t \in (t_1, \infty)$. Hence for all t, $U_m(t)$ is positive, $U_m''(t)$ is nonnegative, and $U_m'(t)$ is negative.

Thence we derive

$$U_m(t) \sim \csc(\pi/m)|t|^{(2-m)/4} \exp(|t|^{m/2}), \qquad t \to -\infty, \tag{2.07}$$

$$U_m'(t) \sim -\tfrac{1}{2}m \csc(\pi/m)|t|^{(m-2)/4} \exp(|t|^{m/2}), \qquad t \to -\infty, \tag{2.08}$$

with relative errors $O(t^{-m/2})$ in both cases.

As the second standard solution of (2.01) we employ $U_m(-t)$. As t increases from $-\infty$ to $+\infty$, this function and its first derivative increase strictly from 0 to $+\infty$. Clearly $U_m(t)$ and $U_m(-t)$ are linearly independent; they also comprise a numerically satisfactory pair in the sense of Miller [19]. Their Wronskian is easily calculated by letting $t \to +\infty$ or $t \to 0$. Either way yields

$$\mathcal{W}\{U_m(t), U_m(-t)\} = m \csc(\pi/m). \tag{2.09}$$

The connection formulas for equation (2.01) are in effect given by the combination of (2.02) and (2.07).

Lastly, in the special cases $m = 2$ and 4, we have

$$U_2(t) = e^{-t}, \qquad U_4(t) = 2^{1/2}U(0, 2t), \tag{2.10}$$

where $U(a, t)$ is the parabolic cylinder function in the notation of Miller [20], or [21, Chap. 19].

2.2. Case II. The basic differential equation in this case is given by

$$d^2w/dt^2 = -\tfrac{1}{4}m^2t^{m-2}w, \tag{2.11}$$

where m is again an even positive integer. The Liouville–Green approximation shows that as $t \to +\infty$, all solutions of this equation have the form

$$Xt^{(2-m)/4}\{\cos(t^{m/2} + \chi) + O(t^{-m/2})\},$$

where the amplitude X and phase χ are arbitrary constants. Similarly, as $t \to -\infty$, the solutions reduce to

$$\hat{X}|t|^{(2-m)/4}\{\cos(|t|^{m/2} + \hat{\chi}) + O(t^{-m/2})\},$$

where $\hat{X}$ and $\hat{\chi}$ are further constants. Following Miller [19], we seek a pair of solutions that differ in phase by $\frac{1}{2}\pi$ as $t \to +\infty$ *and also as* $t \to -\infty$. The general solution of (2.11) is $|t|^{1/2}\mathscr{C}_{1/m}(|t|^{m/2})$, where $\mathscr{C}_{1/m}$ denotes any cylinder function of order $1/m$. By considering the asymptotic forms of the $J_{1/m}$ and $Y_{1/m}$ functions for large positive argument and connecting the solutions at $t = 0$, we find that an appropriate pair of solutions is $W_m(t)$ and $W_m(-t)$, where[6]

$$W_m(t) = -\left(\frac{\pi t}{2}\right)^{1/2}\left\{\tan\left(\frac{\pi}{2m}\right)J_{1/m}(t^{m/2}) + Y_{1/m}(t^{m/2})\right\}, \qquad t > 0, \tag{2.12}$$

$$W_m(t) = \left(\frac{\pi|t|}{2}\right)^{1/2}\left\{\cot\left(\frac{\pi}{2m}\right)J_{1/m}(|t|^{m/2}) - Y_{1/m}(|t|^{m/2})\right\}, \qquad t < 0. \tag{2.13}$$

[6] The analysis also shows that when $m \geqq 4$, the *only* solutions meeting the stipulated phase conditions are constant multiples of $W_m(t)$ and $W_m(-t)$. The normalizing factor has been chosen to make $W_m(t)$ and $W_m'(t)$ agree with $U_m(t)$ and $U_m'(t)$, respectively, at $t = 0$; compare (2.05) and (2.18) below. This has some notational advantages when we come to Case III.

Relevant properties of these solutions are as follows:

$$W_m(t) = \sec\left(\frac{\pi}{2m}\right) t^{(2-m)/4}\{\cos(t^{m/2} + \tfrac{1}{4}\pi) + O(t^{-m/2})\}, \qquad t \to +\infty, \tag{2.14}$$

$$W_m'(t) = -\frac{m}{2}\sec\left(\frac{\pi}{2m}\right) t^{(m-2)/4}\{\sin(t^{m/2} + \tfrac{1}{4}\pi) + O(t^{-m/2})\}, \qquad t \to +\infty, \tag{2.15}$$

$$W_m(t) = \csc\left(\frac{\pi}{2m}\right) |t|^{(2-m)/4}\{\cos(|t|^{m/2} - \tfrac{1}{4}\pi) + O(t^{-m/2})\}, \qquad t \to -\infty, \tag{2.16}$$

$$W_m'(t) = \frac{m}{2}\csc\left(\frac{\pi}{2m}\right) |t|^{(m-2)/4}\{\sin(|t|^{m/2} - \tfrac{1}{4}\pi) + O(t^{-m/2})\}, \qquad t \to -\infty, \tag{2.17}$$

$$\begin{aligned} W_m(0) &= \pi^{-1/2} 2^{(2-m)/(2m)} \Gamma(1/m), \\ W_m'(0) &= \pi^{-1/2} 2^{-(2+m)/(2m)} \Gamma(-1/m), \end{aligned} \tag{2.18}$$

$$\mathcal{W}\{W_m(t), W_m(-t)\} = m \csc(\pi/m). \tag{2.19}$$

Special cases are:

$$W_2(t) = 2^{1/2}\cos(t + \tfrac{1}{4}\pi), \qquad W_4(t) = 2^{3/4} W(0, 2t), \tag{2.20}$$

where $W(a, t)$ is the modified parabolic cylinder function defined by Miller, as in [20], or [21, Chap. 19].[7]

In order to have a convenient way of assessing the magnitudes of the functions $W_m(t)$ and $W_m(-t)$ and their derivatives in the error analysis in subsequent sections, we introduce auxiliary weight, modulus and phase functions as was done, for example, for the Airy functions in [23, Chap. 11]. Let $t = q_m$ denote the smallest nonnegative root of the equation

$$\cos\left(\frac{\pi}{2m}\right) W_m(t) = \sin\left(\frac{\pi}{2m}\right) W_m(-t);$$

compare (2.14) and (2.16). Then the weight function $E_m(t)$ is defined to be

$$E_m(t) = \begin{cases} \left(\cot\dfrac{\pi}{2m}\right)^{1/2}, & t \geqq q_m; \\ \left\{\dfrac{W_m(-t)}{W_m(t)}\right\}^{1/2}, & -q_m \leqq t \leqq q_m; \\ \left(\tan\dfrac{\pi}{2m}\right)^{1/2}, & t \leqq -q_m. \end{cases} \tag{2.21}$$

Using (2.19), we see that

$$\frac{d}{dt}\left\{\tan\left(\frac{\pi}{2m}\right)\frac{W_m(-t)}{W_m(t)}\right\} = \frac{m}{2}\sec^2\left(\frac{\pi}{2m}\right)\frac{1}{W_m^2(t)} > 0.$$

[7] It is on account of the second of (2.10) and the second of (2.20) that the symbols U and W have been chosen for the solutions of (2.01) and (2.11), respectively.

From this relation and the fact that $W_m(0)>0$, it follows that the functions $W_m(t)$ and $W_m(-t)$ have no zeros in the interval $-q_m \leqq t \leqq q_m$, and hence that $E_m(t)$ *is continuous and nondecreasing in the interval* $(-\infty, \infty)$. It is also clear that $E_m(0)=1$, and if $E_m^{-1}(t)$ denotes $1/E_m(t)$, then

$$E_m(-t)=E_m^{-1}(t). \tag{2.22}$$

Modulus functions $M_m(t)$, $N_m(t)$, and phase functions $\theta_m(t)$, $\omega_m(t)$, are defined by

$$W_m(t)=E_m^{-1}(t)M_m(t)\sin\theta_m(t), \qquad W_m(-t)=E_m(t)M_m(t)\cos\theta_m(t), \tag{2.23}$$

$$W'_m(t)=E_m^{-1}(t)N_m(t)\sin\omega_m(t), \qquad W'_m(-t)=E_m(t)N_m(t)\cos\omega_m(t). \tag{2.24}$$

Each is continuous in $(-\infty, \infty)$. From (2.22) it is seen that $M_m(t)$ and $N_m(t)$ are even in t, and

$$\theta_m(t)+\theta_m(-t)=\tfrac{1}{2}\pi, \qquad \omega_m(t)+\omega_m(-t)=\tfrac{1}{2}\pi. \tag{2.25}$$

Using (2.21), we arrive at the following expressions, valid when $t \geqq q_m$:

$$\begin{aligned} M_m(t)&=\left\{\cot\left(\frac{\pi}{2m}\right)W_m^2(t)+\tan\left(\frac{\pi}{2m}\right)W_m^2(-t)\right\}^{1/2}, \\ N_m(t)&=\left\{\cot\left(\frac{\pi}{2m}\right)W_m'^2(t)+\tan\left(\frac{\pi}{2m}\right)W_m'^2(-t)\right\}^{1/2}, \end{aligned} \tag{2.26}$$

$$\begin{aligned} \theta_m(t)&=\tan^{-1}\left\{\cot\left(\frac{\pi}{2m}\right)\frac{W_m(t)}{W_m(-t)}\right\}, \\ \omega_m(t)&=\tan^{-1}\left\{\cot\left(\frac{\pi}{2m}\right)\frac{W'_m(t)}{W'_m(-t)}\right\}. \end{aligned} \tag{2.27}$$

Similarly, when $-q_m \leqq t \leqq q_m$, we have

$$\begin{aligned} M_m(t)&=\{2W_m(t)W_m(-t)\}^{1/2}, \\ N_m(t)&=\left\{\frac{W_m'^2(t)W_m^2(-t)+W_m'^2(-t)W_m^2(t)}{W_m(t)W_m(-t)}\right\}^{1/2}, \end{aligned} \tag{2.28}$$

$$\theta_m(t)=\tfrac{1}{4}\pi, \qquad \omega_m(t)=\tan^{-1}\left\{\frac{W'_m(t)W_m(-t)}{W'_m(-t)W_m(t)}\right\}. \tag{2.29}$$

Lastly, when $t\to\pm\infty$, we obtain from (2.14) to (2.17) the asymptotic forms

$$\begin{aligned} M_m(t)&\sim\{2\csc(\pi/m)\}^{1/2}|t|^{(2-m)/4}, \\ N_m(t)&\sim m\{\tfrac{1}{2}\csc(\pi/m)\}^{1/2}|t|^{(m-2)/4}. \end{aligned} \tag{2.30}$$

2.3. Case III. The basic differential equation in this case is the same as in Case I, that is,

$$d^2w/dt^2=\tfrac{1}{4}m^2t^{m-2}w, \tag{2.31}$$

except that now $m=3, 5, 7, \cdots$. For positive t, the solutions have exponential-type behavior. As in Case I, the first standard solution, which we denote by $V_m(t)$,

is recessive as $t \to +\infty$ and fixed by the condition

$$V_m(t) = t^{(2-m)/4} \exp(-t^{m/2})\{1 + O(t^{-m/2})\}, \qquad t \to +\infty. \tag{2.32}$$

Because this relation has the same form as (2.02), the formulas for $U_m(t)$ given in § 2.1 carry over directly to $V_m(t)$ when $t \geqq 0$. Thus

$$V_m'(t) \sim -\tfrac{1}{2}mt^{(m-2)/4} \exp(-t^{m/2}), \qquad t \to +\infty, \tag{2.33}$$

with relative error $O(t^{-m/2})$,

$$V_m(t) = (2t/\pi)^{1/2} K_{1/m}(t^{m/2}), \qquad t > 0, \tag{2.34}$$

and

$$V_m(0) = \pi^{-1/2} 2^{(2-m)/(2m)} \Gamma(1/m), \qquad V_m'(0) = \pi^{-1/2} 2^{-(2+m)/(2m)} \Gamma(-1/m). \tag{2.35}$$

When t is negative, the differential equation (2.31) has the same form as in Case II, hence the solutions are oscillatory. Moreover, since the initial conditions (2.35) agree with (2.18), the formulas for $W_m(t)$ given in § 2.2 carry over to $V_m(t)$ when $t \leqq 0$. Thus

$$V_m(t) = \left(\frac{\pi|t|}{2}\right)^{1/2} \left\{\cot\left(\frac{\pi}{2m}\right) J_{1/m}(|t|^{m/2}) - Y_{1/m}(|t|^{m/2})\right\}, \qquad t < 0, \tag{2.36}$$

$$V_m(t) = \csc\left(\frac{\pi}{2m}\right) |t|^{(2-m)/4} \{\cos(|t|^{m/2} - \tfrac{1}{4}\pi) + O(t^{-m/2})\}, \qquad t \to -\infty, \tag{2.37}$$

$$V_m'(t) = \frac{m}{2} \csc\left(\frac{\pi}{2m}\right) |t|^{(m-2)/4} \{\sin(|t|^{m/2} - \tfrac{1}{4}\pi) + O(t^{-m/2})\}, \qquad t \to -\infty. \tag{2.38}$$

Again, following Miller [19], we choose the second standard solution of (2.31), $\bar{V}_m(t)$, say, in such a way that $V_m(t)$ and $\bar{V}_m(t)$ are $\frac{1}{2}\pi$ out of phase as $x \to -\infty$. A convenient normalization is given by

$$\bar{V}_m(t) = \sec\left(\frac{\pi}{2m}\right) |t|^{(2-m)/4} \{\cos(|t|^{m/2} + \tfrac{1}{4}\pi) + O(t^{-m/2})\}, \qquad t \to -\infty, \tag{2.39}$$

since this has the same form as $W_m(-t)$ when $t \to -\infty$; compare (2.14). Then from § 2.2 we immediately derive

$$\bar{V}_m'(t) = \frac{m}{2} \sec\left(\frac{\pi}{2m}\right) |t|^{(m-2)/4} \{\sin(|t|^{m/2} + \tfrac{1}{4}\pi) + O(t^{-m/2})\}, \qquad t \to -\infty, \tag{2.40}$$

$$\bar{V}_m(t) = -\left(\frac{\pi|t|}{2}\right)^{1/2} \left\{\tan\left(\frac{\pi}{2m}\right) J_{1/m}(|t|^{m/2}) + Y_{1/m}(|t|^{m/2})\right\}, \qquad t < 0, \tag{2.41}$$

and

$$\begin{aligned} \bar{V}_m(0) &= \pi^{-1/2} 2^{(2-m)/(2m)} \Gamma(1/m), \\ \bar{V}_m'(0) &= -\pi^{-1/2} 2^{-(2+m)/(2m)} \Gamma(-1/m). \end{aligned} \tag{2.42}$$

For positive t, we have, as in Case I,

$$\bar{V}_m(t) = (2t/\pi)^{1/2}\{\pi \csc(\pi/m) I_{1/m}(t^{m/2}) + K_{1/m}(t^{m/2})\}, \qquad t > 0, \tag{2.43}$$

$$\bar{V}_m(t) \sim \csc(\pi/m) t^{(2-m)/4} \exp(t^{m/2}), \qquad t \to +\infty, \tag{2.44}$$

$$\bar{V}'_m(t) \sim \tfrac{1}{2} m \csc(\pi/m) t^{(m-2)/4} \exp(t^{m/2}), \qquad t \to +\infty, \tag{2.45}$$

the relative error in the last two relations being $O(t^{-m/2})$.

The Wronskian relation in the present case is given by

$$\mathcal{W}\{V_m(t), \bar{V}_m(t)\} = m \csc(\pi/m). \tag{2.46}$$

In the special case $m = 3$, the solutions are expressible as Airy functions, as follows:

$$V_3(t) = 2^{5/6} 3^{1/6} \pi^{1/2} \operatorname{Ai}\{(\tfrac{3}{2})^{2/3} t\}, \qquad \bar{V}_3(t) = 2^{5/6} 3^{-1/3} \pi^{1/2} \operatorname{Bi}\{(\tfrac{3}{2})^{2/3} t\}. \tag{2.47}$$

Next, when $t \in [0, \infty)$, $V_m(t)$ is positive and decreasing, $V'_m(t)$ is negative and increasing, and both $\bar{V}_m(t)$ and $\bar{V}'_m(t)$ are positive and increasing. In the case of $V_m(t)$ and $V'_m(t)$, this result is proved by similar analysis to that sketched in § 2.1 for $U_m(t)$ and $U'_m(t)$. For $\bar{V}_m(t)$ and $\bar{V}'_m(t)$, the result follows from the Maclaurin series for $\bar{V}_m(t)$.

Auxiliary functions are defined as follows. First, consider the equation

$$\cos\left(\frac{\pi}{2m}\right) \bar{V}_m(t) = \sin\left(\frac{\pi}{2m}\right) V_m(t).$$

There are no positive roots, since the left-hand side exceeds the right-hand side at $t = 0$, and the former is increasing and the latter decreasing. Let $t = -q_m$ be the negative root of smallest absolute value. Then the weight function we shall use is defined by

$$E_m(t) = \left\{\frac{\bar{V}_m(t)}{V_m(t)}\right\}^{1/2}, \qquad t \geqq -q_m;$$

$$E_m(t) = \left(\tan\frac{\pi}{2m}\right)^{1/2}, \qquad t \leqq -q_m. \tag{2.48}$$

From (2.46) we have

$$\frac{d}{dt}\frac{\bar{V}_m(t)}{V_m(t)} = m \csc\left(\frac{\pi}{m}\right)\frac{1}{V_m^2(t)} > 0.$$

Hence as t decreases continuously from $t = 0$, the value $-q_m$ is reached ahead of the first zero of $\bar{V}_m(t)$, which in turn is reached ahead of the first zero of $V_m(t)$. Thus $E_m(t)$ *is continuous and nondecreasing in the interval* $(-\infty, \infty)$.

With $E_m^{-1}(t)$ again denoting $1/E_m(t)$, modulus functions $M_m(t)$, $N_m(t)$ and phase functions $\theta_m(t)$, $\omega_m(t)$, are defined by

$$V_m(t) = E_m^{-1}(t) M_m(t) \sin\theta_m(t), \qquad \bar{V}_m(t) = E_m(t) M_m(t) \cos\theta_m(t), \tag{2.49}$$

$$V'_m(t) = E_m^{-1}(t) N_m(t) \sin\omega_m(t), \qquad \bar{V}'_m(t) = E_m(t) N_m(t) \cos\omega_m(t). \tag{2.50}$$

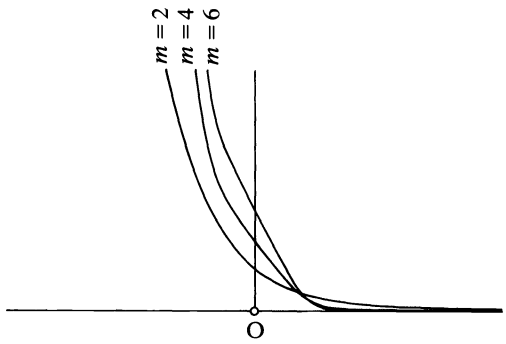


FIG. 2.1. *Case* I. $U_m(t)$, $m = 2, 4, 6$

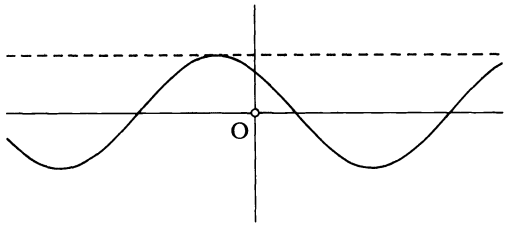


FIG. 2.2. *Case* II. $W_2(t)$ ——; $E_2^{-1}(t)M_2(t)$ – – –

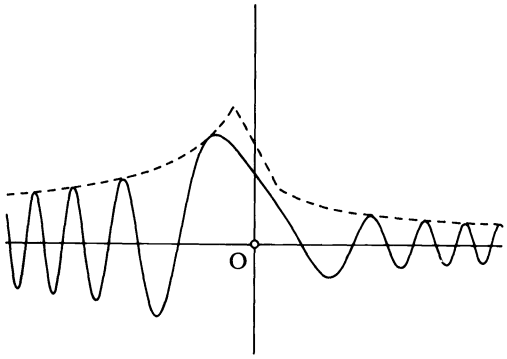


FIG. 2.3. *Case* II. $W_4(t)$ —— ; $E_4^{-1}(t)M_4(t)$ – – –

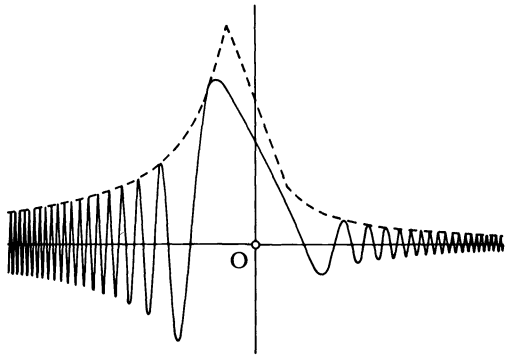


FIG. 2.4. *Case* II. $W_6(t)$ —— ; $E_6^{-1}(t)M_6(t)$ – – –

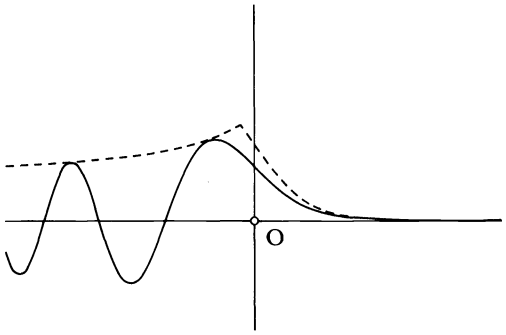


FIG. 2.5. *Case* III. $V_3(t)$ —— ; $E_3^{-1}(t)M_3(t)$ - - -

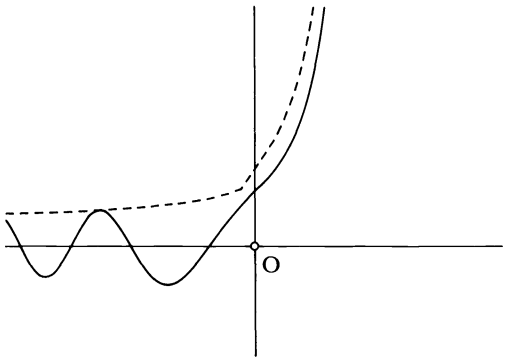


FIG. 2.6. *Case* III. $\bar{V}_3(t)$ —— ; $E_3(t)M_3(t)$ - - -

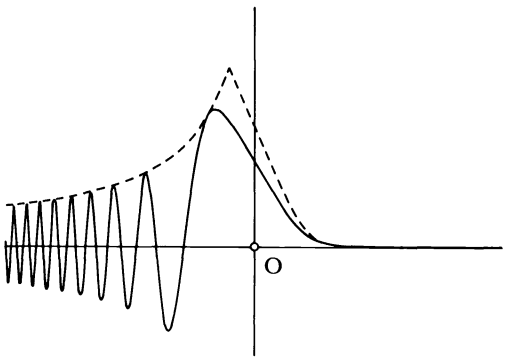


FIG. 2.7. *Case* III. $V_5(t)$ —— ; $E_5^{-1}(t)M_5(t)$ - - -

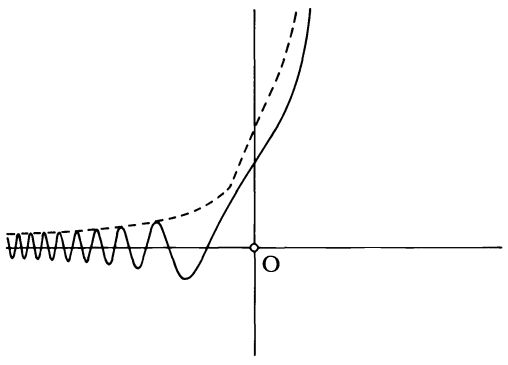


FIG. 2.8. *Case* III. $\bar{V}_5(t)$ —— ; $E_5(t)M_5(t)$ - - -

Each is continuous. Using (2.48), we see that when $t \geqq -q_m$,

$$(2.51)\quad M_m(t) = \{2V_m(t)\bar{V}_m(t)\}^{1/2}, \qquad N_m(t) = \left\{\frac{V_m'^2(t)\bar{V}_m^2(t) + \bar{V}_m'^2(t)V_m^2(t)}{V_m(t)\bar{V}_m(t)}\right\}^{1/2},$$

$$(2.52)\quad \theta_m(t) = \tfrac{1}{4}\pi, \qquad \omega_m(t) = \tan^{-1}\left\{\frac{V_m'(t)\bar{V}_m(t)}{\bar{V}_m'(t)V_m(t)}\right\}.$$

Alternatively, when $t \leqq -q_m$,

$$(2.53)\quad \begin{aligned} M_m(t) &= \left\{\tan\left(\frac{\pi}{2m}\right)V_m^2(t) + \cot\left(\frac{\pi}{2m}\right)\bar{V}_m^2(t)\right\}^{1/2}, \\ N_m(t) &= \left\{\tan\left(\frac{\pi}{2m}\right)V_m'^2(t) + \cot\left(\frac{\pi}{2m}\right)\bar{V}_m'^2(t)\right\}^{1/2}, \end{aligned}$$

$$(2.54)\quad \theta_m(t) = \tan^{-1}\left\{\tan\left(\frac{\pi}{2m}\right)\frac{V_m(t)}{\bar{V}_m(t)}\right\}, \qquad \omega_m(t) = \tan^{-1}\left\{\tan\left(\frac{\pi}{2m}\right)\frac{V_m'(t)}{\bar{V}_m'(t)}\right\}.$$

Lastly, as $t \to \pm\infty$, we find that

$$(2.55)\quad M_m(t) \sim \{2\csc(\pi/m)\}^{1/2}|t|^{(2-m)/4}, \qquad N_m(t) \sim m\{\tfrac{1}{2}\csc(\pi/m)\}^{1/2}|t|^{(m-2)/4}.$$

2.4. Graphs. Typical graphs for the standard solutions in Cases I, II and III are given in the accompanying diagrams.

The continuous curves in Figs. 2.1–2.4 represent the solutions $U_m(t)$, in Case I, and $W_m(t)$, in Case II, for $m = 2$ (no turning point), $m = 4$ (double turning point) and $m = 6$ (quadruple turning point). The graphs of $W_m(t)$ in Figs. 2.2–2.4 are accompanied by broken curves representing the auxiliary function $E_m^{-1}(t)M_m(t)$. This function is used as a majorant for $|W_m(t)|$ in constructing error bounds in subsequent sections; compare the first of equations (2.23). The discontinuities in the direction of the tangent to the broken curves occur at the points $t = \pm q_m$, defined in § 2.2. Numerical values to three decimal places are found to be

$$q_2 = 0.000, \quad q_4 = 0.431, \quad q_6 = 0.596.$$

In a similar manner, the solutions $V_m(t)$ and $\bar{V}_m(t)$ in Case III are represented by the continuous curves in Figs. 2.5–2.8 for $m = 3$ (simple turning point) and $m = 5$ (triple turning point). The graphs of $V_m(t)$ and $\bar{V}_m(t)$ are accompanied by broken curves representing $E_m^{-1}(t)M_m(t)$, and $E_m(t)M_m(t)$, respectively; compare equations (2.49). The discontinuities in the direction of the tangent to the broken curves occur at $t = -q_m$, defined in § 2.3. Numerical values to three decimal places are given by

$$q_3 = 0.279, \qquad q_5 = 0.528.$$

All diagrams are drawn to the same scale, that is, from -5 to 5 horizontally and from -2.75 to 6 vertically, except that the vertical scales in Figs. 2.1 and 2.2 have been shortened to 0 to 6, and -2.75 to 2.75, respectively.

3. Approximations for the solutions

3.1. Preliminary conditions. In this section we investigate solutions of the differential equation

$$d^2w/dx^2 = \{u^2 f(u, x) + g(u, x)\}w, \tag{3.01}$$

when x ranges over a finite or infinite interval (a_1, a_2) and u is a positive parameter. For each value of u, we suppose that $g(u, x)$ is a real or complex function that is continuous (or sectionally continuous) in (a_1, a_2), and $f(u, x)/(x - x_0)^{m-2}$ is a real function that is nonvanishing and twice continuously differentiable in (a_1, a_2). Here x_0 is a given interior point of (a_1, a_2) (which may depend on u), and m is an integer not less than 2. Thus the only possible zero of $f(u, x)$ in (a_1, a_2) is a zero of multiplicity $m-2$ at $x = x_0$. As in § 2, there are three distinct cases I, II and III to consider, depending on the sign of $f(u, x)/(x - x_0)^{m-2}$ and whether m is even or odd.

Following similar investigations in [24] and [25], we introduce a *balancing function* $\Omega_m(t)$, an *auxiliary variable* $\zeta \equiv \zeta(u, x)$, an *auxiliary function* $\hat{f}(u, x)$ and an *error-control function* $H_m(u, x)$. These are defined as follows.

First, $\Omega_m(t)$ is any conveniently chosen continuous[8] even function of the real variable t that is positive, except possibly at $t = 0$, and has the properties

$$\Omega_m(t) = O(t^{(m-2)/2}), \qquad t \to \pm\infty. \tag{3.02}$$

For example, a suitable choice is given by

$$\Omega_m(t) = 1 + |t|^{(m-2)/2}. \tag{3.03}$$

Secondly,

$$\begin{aligned} \zeta &= -\left\{\int_x^{x_0} |f(u, y)|^{1/2}\, dy\right\}^{2/m}, \qquad a_1 < x \leqq x_0; \\ \zeta &= \left\{\int_{x_0}^{x} |f(u, y)|^{1/2}\, dy\right\}^{2/m}, \qquad x_0 \leqq x < a_2. \end{aligned} \tag{3.04}$$

Clearly ζ is a continuous and increasing function of x.

Thirdly,

$$\hat{f}(u, x) = 4|f(u, x)|/(m^2|\zeta|^{m-2}). \tag{3.05}$$

Fourthly,

$$H_m(u, x) = \int \left\{ \frac{1}{\hat{f}^{1/4}(u, x)} \frac{d^2}{dx^2}\left(\frac{1}{\hat{f}^{1/4}(u, x)}\right) - \frac{g(u, x)}{\hat{f}^{1/2}(u, x)} \right\} \frac{dx}{\Omega_m(u^{2/m}\zeta)}, \tag{3.06}$$

the arbitrary constant of integration being immaterial.

For example, in the case of the equation

$$d^2w/dx^2 = \tfrac{1}{4}u^2 m^2 x^{m-2} w$$

[8] Actually "continuous" may be replaced here by "sectionally continuous." Also, $\Omega_m(t)$ may depend on u, but this does not need emphasizing in the notation.

we may take $f(u, x) = \frac{1}{4}m^2x^{m-2}$, $x = x_0$ and $g(u, x) = 0$. Then $\zeta = x$, $\hat{f}(u, x) = 1$, and $H_m(u, x) = 0$.

3.2. Case I.

THEOREM I. *Assume the conditions and notation of* § 3.1, *and also that m is even and the sign of* $f(u, x)/(x - x_0)^{m-2}$ *is positive. Then in* (a_1, a_2), *equation* (3.01) *has twice continuously differentiable solutions* $w_1(u, x)$ *and* $w_2(u, x)$ *such that*

$$w_1(u, x) = \hat{f}^{-1/4}(u, x)\{U_m(-u^{2/m}\zeta) + \varepsilon_1(u, x)\}, \tag{3.07}$$

$$w_2(u, x) = \hat{f}^{-1/4}(u, x)\{U_m(u^{2/m}\zeta) + \varepsilon_2(u, x)\}, \tag{3.08}$$

where

$$\frac{|\varepsilon_1(u, x)|}{U_m(-u^{2/m}\zeta)}, \frac{|\partial\varepsilon_1(u, x)/\partial x|}{\mu_m u^{2/m}\hat{f}^{1/2}(u, x)|U_m'(-u^{2/m}\zeta)|} \leqq \exp\left\{\frac{\lambda_m}{u^{2/m}}\mathcal{V}_{a_1,x}(H_m)\right\} - 1, \tag{3.09}$$

$$\frac{|\varepsilon_2(u, x)|}{U_m(u^{2/m}\zeta)}, \frac{|\partial\varepsilon_2(u, x)/\partial x|}{\mu_m u^{2/m}\hat{f}^{1/2}(u, x)|U_m'(u^{2/m}\zeta)|} \leqq \exp\left\{\frac{\lambda_m}{u^{2/m}}\mathcal{V}_{x,a_2}(H_m)\right\} - 1, \tag{3.10}$$

and

$$\lambda_m = \sup_{t\in(-\infty,\infty)}\left\{\frac{1}{m}\sin\left(\frac{\pi}{m}\right)\Omega_m(t)U_m(t)U_m(-t)\right\}, \tag{3.11}$$

$$\mu_m = \sup_{t\in(-\infty,\infty)}\left\{m\csc\left(\frac{\pi}{m}\right)\frac{1}{|U_m'(t)|U_m(-t)}\right\}. \tag{3.12}$$

In this theorem and subsequent analysis, $\mathcal{V}$ denotes the variational operator defined and discussed in [23, pp. 27–29]; thus, for example,

$$\mathcal{V}_{a_1,x}(H_m) = \int_{a_1}^{x}\left|\frac{\partial H_m(u, y)}{\partial y}\right| dy.$$

We note that as a consequence of the behavior of $U_m(t)$ and $U_m'(t)$ as $t \to \pm\infty$, given in § 2.1, and the conditions on $\Omega_m(t)$ imposed in § 3.1, the suprema in (3.11) and (3.12) are finite.

To prove Theorem I, we begin by transforming from x to ζ as independent variable in the differential equation. Let $\zeta = \zeta_1, \zeta_2$ correspond to the endpoints $x = a_1, a_2$, respectively. Either ζ_1 or ζ_2 or both may be at infinity. From (3.04) and the fact that in the present case, $f(u, x)$ is nonnegative, we see that

$$\tfrac{1}{4}m^2\zeta^{m-2} = \dot{x}^2 f(u, x),$$

where the dot signifies differentiation with respect to ζ. In place of w, we adopt a new dependent variable, given by

$$W = \dot{x}^{-1/2}w. \tag{3.13}$$

The transformation from x and w to ζ and W is then a Liouville transformation [23, pp. 190–193]; accordingly the new differential equation is given by

$$d^2W/d\zeta^2 = \{\tfrac{1}{4}m^2u^2\zeta^{m-2} + \phi(u,\zeta)\}W, \tag{3.14}$$

where

$$\phi(u,\zeta) = \dot{x}^{1/2}\{d^2(\dot{x}^{-1/2})/d\zeta^2\} + \dot{x}^2 g(u,x).$$

From the definition (3.05) we see that

$$\hat{f}(u,x) = 1/\dot{x}^2 \tag{3.15}$$

and hence that

$$\phi(u,\zeta) = \frac{1}{\hat{f}^{1/4}}\frac{d^2}{d\zeta^2}(\hat{f}^{1/4}) + \frac{g}{\hat{f}} = -\frac{1}{\hat{f}^{3/4}}\frac{d^2}{dx^2}\left(\frac{1}{\hat{f}^{1/4}}\right) + \frac{g}{\hat{f}}.$$

LEMMA. *With the conditions in the opening paragraph of* § 3.1, $\zeta/(x-x_0)$ *is a positive, twice continuously differentiable function of* x *when* $a_1 < x < a_2$, *and* $\phi(u,\zeta)$ *is sectionally continuous in the corresponding interval* $\zeta_1 < \zeta < \zeta_2$.

The proof of this result in the case $m = 3$ can be found, for example, on p. 399 of [23]. The extension to other values of m is straightforward, and it is unnecessary to record details.

We return to the proof of Theorem I. Substituting $w_2(u,x)$ for w on the right-hand side of (3.13) and using (3.08) and (3.15), we obtain

$$W = U_m(u^{2/m}\zeta) + \varepsilon_2(u,x).$$

On combining (2.01), with $w = U_m$ and $t = u^{2/m}\zeta$, and (3.14), we arrive at the following inhomogeneous differential equation for the error term:

$$(d^2\varepsilon_2/d\zeta^2) - \tfrac{1}{4}m^2u^2\zeta^{m-2}\varepsilon_2 = \phi(u,\zeta)\{U_m(u^{2/m}\zeta) + \varepsilon_2\}.$$

Using the Wronskian (2.09), we construct an equivalent Volterra integral equation:

$$\eta_2(u,\zeta) = \frac{1}{mu^{2/m}}\sin\left(\frac{\pi}{m}\right)\int_{\zeta}^{\zeta_2} K(\zeta,v)\phi(u,v)\{U_m(u^{2/m}v) + \eta_2(u,v)\}\,dv, \tag{3.16}$$

where $\eta_2(u,\zeta)$ has been written for $\varepsilon_2(u,x)$ and

$$K(\zeta,v) = U_m(u^{2/m}\zeta)U_m(-u^{2/m}v) - U_m(-u^{2/m}\zeta)U_m(u^{2/m}v). \tag{3.17}$$

From the monotonicity properties of the function $U_m(t)$ stated in § 2.1, it follows that the kernel is bounded by

$$0 \leqq K(\zeta,v) < U_m(u^{2/m}\zeta)U_m(-u^{2/m}v), \qquad \zeta \leqq v.$$

Next, if we differentiate (3.17) with respect to ζ and use (2.09), bearing in mind that U_m' is always negative, we find that

$$\left|\frac{\partial K(\zeta,v)}{\partial\zeta}\right| \leqq u^{2/m}m\csc\left(\frac{\pi}{m}\right)\frac{U_m(-u^{2/m}v)}{U_m(-u^{2/m}\zeta)}, \qquad \zeta \leqq v,$$

and hence

$$\left|\frac{\partial K(\zeta, v)}{\partial \zeta}\right| \leqq \mu_m u^{2/m} |U_m'(u^{2/m}\zeta)| U_m(-u^{2/m}v), \qquad \zeta \leqq v,$$

where μ_m is defined by (3.12).

Having bounded the kernel and its ζ-derivative, we can solve the integral equation (3.16) by the standard procedure of successive approximations, for example, by applying Theorem 10.2 of [23, Chap. 6]. Then returning from ζ to x as variable, we arrive at the desired inequalities (3.10). The proof of (3.09) is similar, or we may merely replace x by $-x$ in the result for $w_2(u, x)$ and $\varepsilon_2(u, x)$.

3.3. Case II.

THEOREM II. *Assume the conditions and notation of* § 3.1, *and also that m is even and the sign of* $f(u, x)/(x-x_0)^{m-2}$ *is negative. Then in* (a_1, a_2), *equation* (3.01) *has twice continuously differentiable solutions* $w_1(u, x)$ *and* $w_2(u, x)$ *such that*

$$w_1(u, x) = \hat{f}^{-1/4}(u, x)\{W_m(-u^{2/m}\zeta) + \varepsilon_1(u, x)\}, \tag{3.18}$$

$$w_2(u, x) = \hat{f}^{-1/4}(u, x)\{W_m(u^{2/m}\zeta) + \varepsilon_2(u, x)\}, \tag{3.19}$$

where

$$\frac{|\varepsilon_1(u, x)|}{M_m(u^{2/m}\zeta)}, \frac{|\partial\varepsilon_1(u, x)/\partial x|}{u^{2/m}\hat{f}^{1/2}(u, x)N_m(u^{2/m}\zeta)} \leqq \frac{\sigma_m}{\rho_m} E_m(u^{2/m}\zeta)\left[\exp\left\{\frac{\rho_m}{u^{2/m}}\mathcal{V}_{a_1,x}(H_m)\right\} - 1\right], \tag{3.20}$$

$$\frac{|\varepsilon_2(u, x)|}{M_m(u^{2/m}\zeta)}, \frac{|\partial\varepsilon_2(u, x)/\partial x|}{u^{2/m}\hat{f}^{1/2}(u, x)N_m(u^{2/m}\zeta)} \leqq \frac{\sigma_m}{\rho_m} E_m^{-1}(u^{2/m}\zeta)\left[\exp\left\{\frac{\rho_m}{u^{2/m}}\mathcal{V}_{x,a_2}(H_m)\right\} - 1\right], \tag{3.21}$$

and

$$\rho_m = \sup_{t\in(-\infty,\infty)}\left\{\frac{1}{m}\sin\left(\frac{\pi}{m}\right)\Omega_m(t)M_m^2(t)\right\}, \tag{3.22}$$

$$\sigma_m = \sup_{t\in(-\infty,\infty)}\left\{\frac{1}{m}\sin\left(\frac{\pi}{m}\right)\Omega_m(t)|W_m(t)|E_m(t)M_m(t)\right\}. \tag{3.23}$$

In this theorem the functions $W_m(t)$, $E_m(t)$, $M_m(t)$ and $N_m(t)$ are defined in § 2.2. The proof is similar to that of Theorem I, and it is unnecessary to give details. Again, in consequence of the asymptotic properties of $W_m(t)$ and $M_m(t)$ for large $|t|$ given in § 2.2, the positive constants ρ_m and σ_m are finite. It is also clear from (2.23) that $\sigma_m \leqq \rho_m$; in consequence, σ_m may be replaced by ρ_m in (3.20) and (3.21), thereby simplifying the results.

3.4. Case III.

THEOREM III. *Assume the conditions and notation of* § 3.1, *and also that m is odd and the sign of* $f(u, x)/(x-x_0)^{m-2}$ *is positive.*[9] *Then in* (a_1, a_2), *equation* (3.01) *has twice continuously differentiable solutions* $w_1(u, x)$ *and* $w_2(u, x)$ *such that*

$$w_1(u, x) = \hat{f}^{-1/4}(u, x)\{\bar{V}_m(u^{2/m}\zeta) + \varepsilon_1(u, x)\}, \tag{3.24}$$

$$w_2(u, x) = \hat{f}^{-1/4}(u, x)\{V_m(u^{2/m}\zeta) + \varepsilon_2(u, x)\}, \tag{3.25}$$

where

$$\frac{|\varepsilon_1(u, x)|}{M_m(u^{2/m}\zeta)}, \quad \frac{|\partial\varepsilon_1(u, x)/\partial x|}{u^{2/m}\hat{f}^{1/2}(u, x)N_m(u^{2/m}\zeta)} \leqq \frac{\sigma_{m,1}}{\rho_m} E_m(u^{2/m}\zeta)\left[\exp\left\{\frac{\rho_m}{u^{2/m}}\mathcal{V}_{a_1,x}(H_m)\right\} - 1\right], \tag{3.26}$$

$$\frac{|\varepsilon_2(u, x)|}{M_m(u^{2/m}\zeta)}, \quad \frac{|\partial\varepsilon_2(u, x)/\partial x|}{u^{2/m}\hat{f}^{1/2}(u, x)N_m(u^{2/m}\zeta)} \leqq \frac{\sigma_{m,2}}{\rho_m} E_m^{-1}(u^{2/m}\zeta)\left[\exp\left\{\frac{\rho_m}{u^{2/m}}\mathcal{V}_{x,a_2}(H_m)\right\} - 1\right], \tag{3.27}$$

and

$$\rho_m = \sup_{t\in(-\infty,\infty)}\left\{\frac{1}{m}\sin\left(\frac{\pi}{m}\right)\Omega_m(t)M_m^2(t)\right\}, \tag{3.28}$$

$$\sigma_{m,1} = \sup_{t\in(-\infty,\infty)}\left\{\frac{1}{m}\sin\left(\frac{\pi}{m}\right)\Omega_m(t)|\bar{V}_m(t)|E_m^{-1}(t)M_m(t)\right\}, \tag{3.29}$$

$$\sigma_{m,2} = \sup_{t\in(-\infty,\infty)}\left\{\frac{1}{m}\sin\left(\frac{\pi}{m}\right)\Omega_m(t)|V_m(t)|E_m(t)M_m(t)\right\}. \tag{3.30}$$

In this theorem, the functions $\bar{V}_m(t)$, $V_m(t)$, $E_m(t)$, $M_m(t)$ and $N_m(t)$ are defined in § 2.3. The proof is similar to that of Theorems I and II. Again, the positive constants ρ_m, $\sigma_{m,1}$ and $\sigma_{m,2}$ are finite, and since $\sigma_{m,1}$ and $\sigma_{m,2}$ are both bounded by ρ_m, the ratios $\sigma_{m,1}/\rho_m$ and $\sigma_{m,2}/\rho_m$ in (3.26) and (3.27) may be replaced by unity.

3.5. Remarks on Theorems I, II and III. (i) The present results agree with earlier ones in cases in which there are no turning points or one turning point. Thus Theorems I and II reduce respectively to Theorems 2.1 and 2.2 of [23, Chap. 6] when $m = 2$ and $\Omega_2(t) = 1$, and Theorem III reduces to Theorem 3.1 of [23, Chap. 11] when $m = 3$ and $\Omega_3(t) = |t|^{1/2}$.

(ii) The theorems are essentially inequalities for the solutions of the differential equation, and not asymptotic results *per se*. Subject to the given conditions they are valid for any positive value of the parameter u, not necessarily large. But, of course, the results are of interest only when the error terms $\varepsilon_1(u, x)$ and $\varepsilon_2(u, x)$ are small compared with the corresponding approximants $U_m(\mp u^{2/m}\zeta)$,

[9] Cases in which the sign of $f(u, x)/(x-x_0)^{m-2}$ is negative can be accommodated by replacing x by $-x$ throughout.

$W_m(\mp u^{2/m}\zeta)$, $\bar{V}_m(u^{2/m}\zeta)$ or $V_m(u^{2/m}\zeta)$. When this happens, the graphs in Figs. 2.1–2.8) mirror the behavior of the actual solutions, or more precisely, the products of each solution and the function $\hat{f}^{1/4}(u, x)$. In § 5 it will be shown that with suitable dependence of $f(u, x)$ and $g(u, x)$ on u, the error terms are uniformly vanishingly small compared with the approximants as $u \to \infty$.

4. The connection formulas

4.1. Preliminary conditions. In this section we construct connection formulas for a turning point of arbitrary multiplicity which are analogous to the Gans–Jeffreys formulas for the case of a single turning point given, for example, in [23, pp. 491–494].

As in § 3, we consider the differential equation

$$d^2w/dx^2 = \{u^2 f(u, x) + g(u, x)\}w, \qquad u > 0, \quad a_1 < x < a_2, \tag{4.01}$$

in which the given interval (a_1, a_2) is finite or infinite and contains just one zero x_0, say, of $f(u, x)$, the multiplicity of this zero being $m-2$. In addition, we assume that within (a_1, a_2):

(i) $f(u, x)/(x-x_0)^{m-2}$ *is real and twice continuously differentiable.*

(ii) $g(u, x)$ *is continuous.*

(iii) *As* $x \to a_1+$ *or* a_2-, *the integral* $\int_{x_0}^{x} |f(u, y)|^{1/2}\, dy$ *diverges and* $\mathcal{V}(F)$ *converges, where* $F(u, x)$ *is the error-control function for the Liouville–Green approximation, that is,*

$$F(u, x) \equiv \int \left\{ \frac{1}{|f|^{1/4}} \frac{d^2}{dx^2}\left(\frac{1}{|f|^{1/4}}\right) - \frac{g}{|f|^{1/2}} \right\} dx. \tag{4.02}$$

We first show that these conditions ensure that the error-control function $H_m(u, x)$ defined by (3.06) is of bounded variation in (a_1, a_2), provided that the balancing function $\Omega_m(t)$ is chosen to satisfy the extra condition

$$\Omega_m(t)/|t|^{(m-2)/2} \to \text{const.}, \qquad t \to \pm\infty. \tag{4.03}$$

Substituting in (3.06) by means of (3.05) and using the differential relation $|f|^{1/2}\, dx = \frac{1}{2}m|\zeta|^{(m-2)/2}\, d\zeta$, we find that

$$H_m(u, x) = \tag{4.04}$$

$$\frac{m}{2} \int \left\{ \frac{1}{|f|^{1/4}} \frac{d^2}{dx^2}\left(\frac{1}{|f|^{1/4}}\right) - \frac{g}{|f|^{1/2}} \right\} \frac{|\zeta|^{(m-2)/2} dx}{\Omega_m(u^{2/m}\zeta)} - \frac{m^2-4}{16} \int \frac{d\zeta}{\zeta^2 \Omega_m(u^{2/m}\zeta)}.$$

Condition (iii) above implies that $\zeta \to -\infty$ as $x \to a_1+$, and $\zeta \to +\infty$ as $x \to a_2-$. Aided also by (4.03), we see that each integral on the right-hand side of (4.04) converges absolutely, and therefore that $\mathcal{V}(H_m)$ converges, as asserted.

4.2. Case I. Here m ($\geqq 2$) is even and the sign of $f(u, x)/(x-x_0)^{m-2}$ is positive. With the given conditions it is known from the theory of the Liouville–Green approximation [23, pp. 197–200] that for each value of u there exist

unique solutions $\hat{w}_1(u, x)$ and $\hat{w}_2(u, x)$, say, of (4.01) having the properties:

$$\hat{w}_1(u, x) \sim f^{-1/4}(u, x) \exp\left(-u \int_x^{x_0} f^{1/2}(u, y)\, dy\right), \qquad x \to a_1+, \tag{4.05}$$

$$\hat{w}_2(u, x) \sim f^{-1/4}(u, x) \exp\left(-u \int_{x_0}^{x} f^{1/2}(u, y)\, dy\right), \qquad x \to a_2-. \tag{4.06}$$

Our object is to determine the asymptotic form of $\hat{w}_1(u, x)$ as $x \to a_2-$ and the asymptotic form of $\hat{w}_2(u, x)$ as $x \to a_1+$.

All the conditions of Theorem I of § 3.2 are satisfied with the present assumptions, in consequence there exists a solution $w_2(u, x)$ given by (3.08) and (3.10). Furthermore, the right-hand side of (3.10) is finite for each u. Letting $x \to a_2-$, we see that

$$\varepsilon_2(u, x)/U_m(u^{2/m}\zeta) \to 0,$$

and hence from (3.08) that

$$w_2(u, x) \sim \hat{f}^{-1/4}(u, x) U_m(u^{2/m}\zeta).$$

Since $\zeta \to +\infty$, we may replace $U_m(u^{2/m}\zeta)$ by its asymptotic form, obtainable from (2.02). Then using (3.04) and (3.05), we derive

$$w_2(u, x) \sim (\tfrac{1}{2}m)^{1/2} u^{(2-m)/(2m)} f^{-1/4}(u, x) \exp\left(-u \int_{x_0}^{x} f^{1/2}(u, y)\, dy\right).$$

On comparing this relation with (4.06), we identify

$$\hat{w}_2(u, x) = (\tfrac{1}{2}m)^{-1/2} u^{(m-2)/(2m)} w_2(u, x). \tag{4.07}$$

We now let $x \to a_1+$, that is, $\zeta \to -\infty$. From (3.08), (3.10) and (2.07), we obtain

$$w_2(u, x) \sim (\tfrac{1}{2}m)^{1/2} \csc(\pi/m)\ (1+k_2) u^{(2-m)/(2m)} f^{-1/4}(u, x) \exp\left(u \int_x^{x_0} f^{1/2}(u, y)\, dy\right),$$

where k_2 is a constant bounded by

$$|k_2| \leqq \exp\{\lambda_m u^{-2/m} \mathscr{V}_{a_1, a_2}(H_m)\} - 1, \tag{4.08}$$

provided that the right-hand side of this inequality does not exceed unity.[10] Here λ_m is defined by (3.11). Combination of this result with (4.07) immediately yields

$$\hat{w}_2(u, x) \sim (1+k_2) \csc\left(\frac{\pi}{m}\right) f^{-1/4}(u, x) \exp\left(u \int_x^{x_0} f^{1/2}(u, y)\, dy\right), \qquad x \to a_1+. \tag{4.09}$$

Relations (4.06) and (4.09) comprise one of the wanted connection formulas. By symmetry, the other formula is given by (4.05) and

$$\hat{w}_1(u, x) \sim (1+k_1) \csc\left(\frac{\pi}{m}\right) f^{-1/4}(u, x) \exp\left(u \int_{x_0}^{x} f^{1/2}(u, y)\, dy\right), \qquad x \to a_2-, \tag{4.10}$$

[10] Actually it is not obvious that k_2 *is* a constant, that is, independent of x, because Theorem I merely states that $|\varepsilon_2(u, x)|/U_m(u^{2/m}\zeta)$ is bounded as $x \to a_1+$. The assertion is justifiable, however, of Theorem 3.1 of [23, Chap. 6].

where k_1 is a constant subject to the same bound (4.08) as k_2.

4.3. Case II. Here m ($\geqq 2$) is even and the sign of $f(u,x)/(x-x_0)^{m-2}$ is negative. With the conditions of § 4.1, for each value of u there are unique solutions $\hat{w}_1(u,x)$ and $\hat{w}_2(u,x)$ of (4.01), such that

$$\hat{w}_1(u,x) = |f(u,x)|^{-1/4}\left\{\cos\left(u\int_x^{x_0}|f(u,y)|^{1/2}\,dy+\tfrac{1}{4}\pi\right)+o(1)\right\}, \tag{4.11}$$

$$x\to a_1+,$$

$$\hat{w}_2(u,x) = |f(u,x)|^{-1/4}\left\{\cos\left(u\int_{x_0}^{x}|f(u,y)|^{1/2}\,dy+\tfrac{1}{4}\pi\right)+o(1)\right\}, \tag{4.12}$$

$$x\to a_2-.$$

Using analysis similar to that of the preceding subsection and § 7.2 of [23, Chap. 13], we find that

$$\hat{w}_1(u,x) = (1+\gamma_1)\cot\left(\frac{\pi}{2m}\right)|f(u,x)|^{-1/4} \tag{4.13}$$

$$\cdot\left\{\cos\left(u\int_{x_0}^{x}|f(u,y)|^{1/2}\,dy-\tfrac{1}{4}\pi+\delta_1\right)+o(1)\right\}, \qquad x\to a_2-,$$

$$\hat{w}_2(u,x) = (1+\gamma_2)\cot\left(\frac{\pi}{2m}\right)|f(u,x)|^{-1/4} \tag{4.14}$$

$$\cdot\left\{\cos\left(u\int_{x}^{x_0}|f(u,y)|^{1/2}\,dy-\tfrac{1}{4}\pi+\delta_2\right)+o(1)\right\}, \qquad x\to a_1+,$$

where γ_1, γ_2, δ_1 and δ_2 are constants bounded by

$$|\gamma_1|, |\gamma_2|, \frac{2|\delta_1|}{\pi}, \frac{2|\delta_2|}{\pi} \leqq \frac{\sigma_m}{\rho_m}\left[\exp\left\{\frac{\rho_m}{u^{2/m}}\mathscr{V}_{a_1,a_2}(H_m)\right\}-1\right], \tag{4.15}$$

and ρ_m and σ_m are defined by (3.22) and (3.23). These results are the required connection formulas, and are valid whenever the right-hand side of (4.15) does not exceed unity.

4.4. Case III. In this case, m ($\geqq 3$) is odd and the sign of $f(u,x)/(x-x_0)^m$ is positive.[11] Again, with the conditions of § 4.1, for each u there are unique solutions $\hat{w}_1(u,x)$ and $\hat{w}_2(u,x)$, such that

$$\hat{w}_1(u,x) = |f(u,x)|^{-1/4}\left\{\cos\left(u\int_x^{x_0}|f(u,y)|^{1/2}\,dy+\tfrac{1}{4}\pi\right)+o(1)\right\}, \tag{4.16}$$

$$x\to a_1+,$$

$$\hat{w}_2(u,x) \sim f^{-1/4}(u,x)\exp\left(-u\int_{x_0}^{x} f^{1/2}(u,y)\,dy\right), \qquad x\to a_2-. \tag{4.17}$$

[11] As in § 3.4, cases in which $f(u,x)/(x-x_0)^{m-2}$ is negative can be accommodated by reversal of the sign of x.

Using analysis similar to that in Cases I and II, we find that the required connection formulas are given by

$$\hat{w}_1(u,x) \sim \tfrac{1}{2}(1+k)\csc\left(\frac{\pi}{2m}\right) f^{-1/4}(u,x)\exp\left(u\int_{x_0}^{x} f^{1/2}(u,y)\,dy\right), \tag{4.18}$$

$$x \to a_2-,$$

$$\hat{w}_2(u,x) = (1+\gamma)\csc\left(\frac{\pi}{2m}\right)|f(u,x)|^{-1/4} \tag{4.19}$$

$$\cdot\left\{\cos\left(u\int_{x}^{x_0}|f(u,y)|^{1/2}\,dy - \tfrac{1}{4}\pi + \delta\right) + o(1)\right\}, \qquad x \to a_1+,$$

where k, γ and δ are constants bounded by

$$|k| \leqq 2^{1/2}\frac{\sigma_{m,1}}{\rho_m}\left[\exp\left\{\frac{\rho_m}{u^{2/m}}\,\mathcal{V}_{a_1,a_2}(H_m)\right\}-1\right], \tag{4.20}$$

$$|\gamma|,\ \frac{2|\delta|}{\pi} \leqq \frac{\sigma_{m,2}}{\rho_m}\left[\exp\left\{\frac{\rho_m}{u^{2/m}}\,\mathcal{V}_{a_1,a_2}(H_m)\right\}-1\right], \tag{4.21}$$

and ρ_m, $\sigma_{m,1}$ and $\sigma_{m,2}$ are defined by (3.28), (3.29) and (3.30). The only additional condition needed is that in the case of (4.19), the right-hand side of (4.21) must not exceed unity. It is easily verified that these results reduce to those given in [23, pp. 491–494] in the case $m = 3$.

4.5. Remark. In each of Cases I, II and III, we have constructed connection formulas for two solutions of the given differential equation, complete with bounds on the errors in the approximate coefficients. Provided that the error terms are small compared with the coefficients—and this forms the subject of the next section—each pair of solutions of the differential equation is numerically satisfactory, because the approximating functions were chosen in § 2 to be numerically satisfactory solutions of the corresponding basic equation. In consequence, connection formulas for any other solution can be deduced whenever the problem is not ill-posed. This aspect is discussed more fully in [26].

5. Asymptotic estimates of the error terms

5.1. Preliminary remarks. The error terms $\varepsilon_1(u,x)$ and $\varepsilon_2(u,x)$ appearing in the solutions given by Theorems I, II and III of § 3 depend on the parameter u, as do the error terms k, k_1, k_2, γ, γ_1, γ_2, δ, δ_1 and δ_2 in the connection formulas derived in § 4. The behavior of these error terms as $u \to \infty$ obviously depends on the manner in which u enters the functions $f(u,x)$ and $g(u,x)$ in the given differential equation (3.01).

5.2. A typical case. Let us assume that the endpoints a_1 and a_2 of the x-interval are independent of u, and that the conditions of § 3.1 are satisfied. We also assume that:

(i) *$f(u,x) \equiv f(x)$ is independent of u, and the variation of the function $\int f^{-1/4}(x)\{f^{-1/4}(x)\}''\,dx$ converges as $x \to a_1+$ or a_2-.*

(ii) *There is a real constant* ϖ *such that* $u^{-\varpi}|g(u,x)|$ *is bounded when* u *is arbitrarily large and* x *ranges over any fixed compact interval within* (a_1, a_2). *Moreover, the variation of* $\int u^{-\varpi} g(u,x) f^{-1/2}(x)\,dx$ *converges uniformly with respect to* u *as* $x \to a_1+$ *or* a_2-.

The main problem is to find an asymptotic estimate of $\mathcal{V}_{a_1,a_2}(H_m)$ as $u \to \infty$, where $H_m(u,x)$ is defined by (3.06); thus

$$\mathcal{V}_{a_1,a_2}(H_m) = \int_{a_1}^{a_2} \left| \frac{1}{\hat{f}^{1/4}(x)} \frac{d^2}{dx^2}\left(\frac{1}{\hat{f}^{1/4}(x)}\right) - \frac{g(u,x)}{\hat{f}^{1/2}(x)} \right| \frac{dx}{\Omega_m(u^{2/m}\zeta)}. \tag{5.01}$$

Consider first the interval $[x_1, x_2]$, where x_1 and x_2 are any fixed points such that $a_1 < x_1 < x_0 < x_2 < a_2$. From (3.05), we have

$$\hat{f}(x) \equiv \hat{f}(u,x) = \frac{4}{m^2} \frac{|f(x)|}{|x-x_0|^{m-2}} \left(\frac{x-x_0}{\zeta}\right)^{m-2}.$$

By hypothesis, $|f(x)|/|x-x_0|^{m-2}$ is nonvanishing and twice continuously differentiable, and from the Lemma of § 3.2, and the corresponding results for Cases II and III, we know that the same is true of $(x-x_0)/\zeta$. Hence $\hat{f}^{-1/4}(\hat{f}^{-1/4})''$ is continuous in (a_1, a_2) and therefore bounded in absolute value in $[x_1, x_2]$. Next, using Condition (ii) above, we see that $u^{-\varpi}\hat{f}^{-1/1}(x)|g(u,x)|$ is bounded in $[x_1, x_2]$. Combination of these two results shows that

$$\left| \frac{1}{\hat{f}^{1/4}(x)} \frac{d^2}{dx^2}\left(\frac{1}{\hat{f}^{1/4}(x)}\right) - \frac{g(u,x)}{\hat{f}^{1/2}(x)} \right| \leqq A u^{\varpi_0}, \qquad u \geq 1, \quad x \in [x_1, x_2], \tag{5.02}$$

where A is a constant and

$$\varpi_0 = \max(\varpi, 0). \tag{5.03}$$

We now adopt the choice (3.03) for the balancing function $\Omega_m(t)$. Since $(x-x_0)/\zeta$ is bounded in $[x_1, x_2]$, it follows that

$$\frac{1}{\Omega_m(u^{2/m}\zeta)} \leqq \frac{A}{1+u^{(m-2)/m}|x-x_0|^{(m-2)/2}},$$

where the symbol A is now being used generically. On combining this inequality with (5.02) and referring to (5.01), we see that

$$\mathcal{V}_{x_1,x_2}(H_m) \leqq A u^{\varpi_0} \int_{x_1}^{x_2} \frac{dx}{1+u^{(m-2)/m}|x-x_0|^{(m-2)/2}}.$$

With $\xi = u^{2/m}(x-x_0)$ as new integration variable, we have

$$\int_{x_0}^{x_2} \frac{dx}{1+u^{(m-2)/m}|x-x_0|^{(m-2)/2}} = \frac{1}{u^{2/m}} \int_0^{u^{2/m}(x_2-x_0)} \frac{d\xi}{1+\xi^{(m-2)/2}}.$$

For large u, this is $O(u^{-(m-2)/m})$, $O(u^{-2/m} \ln u)$ or $O(u^{-2/m})$, according as $2 \leqq m < 4$, $m = 4$ or $m > 4$. The same estimates hold for the corresponding integral

over the interval $[x_1, x_0]$, hence $\mathcal{V}_{x_1,x_2}(H_m)$ is estimated by

$$\begin{aligned} &O(u^{\varpi_0-1+(2/m)}), \qquad m=2,3;\\ &O(u^{\varpi_0-(1/2)}\ln u), \qquad m=4;\\ &O(u^{\varpi_0-(2/m)}), \qquad m>4. \end{aligned} \tag{5.04}$$

Next, consider the contribution from the interval $[x_2, a_2)$. Instead of (5.01), we shall use the expression (4.04), with the same choice (3.03) for $\Omega_m(t)$. Since ζ is nonzero in the present circumstances, we have

$$\frac{|\zeta|^{(m-2)/2}}{\Omega_m(u^{2/m}\zeta)} = \frac{\zeta^{(m-2)/2}}{1+u^{(m-2)/m}\zeta^{(m-2)/2}} < \frac{1}{u^{(m-2)/m}}.$$

Combining this bound with Conditions (i) and (ii) in turn, we derive

$$\int_{x_2}^{a_2} \left| \frac{1}{f^{1/4}(x)} \frac{d^2}{dx^2}\left\{\frac{1}{f^{1/4}(x)}\right\} \frac{\zeta^{(m-2)/2}}{\Omega_m(u^{2/m}\zeta)}\, dx \right| = O\left(\frac{1}{u^{(m-2)/m}}\right)$$

and

$$\int_{x_2}^{a_2} \left| \frac{g(u,x)}{f^{1/2}(x)} \frac{\zeta^{(m-2)/2}}{\Omega_m(u^{2/m}\zeta)}\, dx \right| = O\left(\frac{u^{\varpi}}{u^{(m-2)/m}}\right).$$

For the remaining contribution in (4.04), let ζ_2 again denote the value of ζ corresponding to $x = a_2$, and $z_2(>0)$ the value of ζ corresponding to $x = x_2$. Then we have

$$\begin{aligned} \int_{z_2}^{\zeta_2} \frac{d\zeta}{\zeta^2\Omega_m(u^{2/m}\zeta)} &= \int_{z_2}^{\zeta_2} \frac{d\zeta}{\zeta^2\{1+u^{(m-2)/m}\zeta^{(m-2)/2}\}}\\ &< \frac{1}{u^{(m-2)/m}} \int_{z_2}^{\infty} \frac{d\zeta}{\zeta^{(m+2)/2}} = O\left(\frac{1}{u^{(m-2)/m}}\right), \end{aligned}$$

since the last integral converges when $m \geqq 2$. Substituting in (4.04) by means of these three relations, we see that

$$\mathcal{V}_{x_2,a_2}(H_m) = O(u^{(m\varpi_0-m+2)/m}),$$

which is absorbable in the estimates (5.04) for $\mathcal{V}_{x_1,x_2}(H_m)$. Similarly $\mathcal{V}_{a_1,x_1}(H_m)$ is absorbable in (5.04). Accordingly $\mathcal{V}_{a_1,a_2}(H_m)$, also, is estimated by (5.04).

Uniform estimates for the error terms $\varepsilon_1(u,x)$ and $\varepsilon_2(u,x)$ appearing in Theorems I, II and III are easily available now by substituting (5.04) for the variations of H_m appearing in these theorems. For example, in Case I both $\varepsilon_1(u,x)/U_m(-u^{2/m}\zeta)$ and $\varepsilon_2(u,x)/U_m(u^{2/m}\zeta)$ are estimated by

$$\begin{aligned} &O(u^{\varpi_0-1}), \qquad m=2,3;\\ &O(u^{\varpi_0-1}\ln u), \qquad m=4;\\ &O(u^{\varpi_0-(4/m)}), \qquad m>4; \end{aligned} \tag{5.05}$$

uniformly for $x \in (a_1, a_2)$. Furthermore, with the additional condition that $\int_{x_0}^{x} |f(y)|^{1/2}\, dy$ diverges as $x \to a_1+$ or a_2- each of the error terms k, k_1, k_2, γ, γ_1, γ_2, δ, δ_1 and δ_2 appearing in the connection formulas of § 4 is estimated by (5.05).

5.3. Specializations. If the conditions stated in the first paragraph of § 5.2 are satisfied and $\varpi \leqq 0$, then from (5.03) we have $\varpi = 0$. Accordingly, the estimates (5.05) become

$$\begin{aligned} &O(u^{-1}), && m = 2, 3; \\ &O(u^{-1} \ln u), && m = 4; \\ &O(u^{-4/m}), && m > 4. \end{aligned} \tag{5.06}$$

In particular, these estimates apply when $g(u, x)$ is independent of u, provided that $\int |gf^{-1/2}|\, dx$ converges as $x \to a_1+$ or a_2-.

On the other hand, if $\varpi > 0$, then $\varpi_0 = \varpi$. The estimates (5.05) for the error terms are valid for any ϖ, but they are useful only when they are $o(1)$ as $u \to \infty$. Thus we require $\varpi < 1$ when $m = 2$, 3 or 4, and $\varpi < /4m$ when $m > 4$. When these conditions are fulfilled Theorems I, II and III, and the corresponding connection formulas of § 4 supply useful information, but not otherwise. To put this another way, when $\varpi \geqq \min(1, 4/m)$, the term $g(u, x)$ in (3.01) cannot be treated as a perturbation compared with $u^2 f(u, x)$; at least part of $g(u, x)$ must be taken into account in constructing the basic equation, in consequence Bessel functions no longer furnish adequate approximations to the solutions.

6. Previous results and conclusions

6.1. Simple turning points. In the case of a simple turning point, two types of method are available for the construction of connection formulas. The first consists of approximating, in some manner, the solutions of the differential equation by Airy functions or, equivalently, Bessel functions of order one-third, and then substituting for the approximants by their asymptotic approximations in terms of elementary functions for large positive and negative arguments. In the case $m = 3$, the method used in the present paper is of this type.

The second method is applicable only when the coefficients $f(u, x)$ and $g(u, x)$ in the given differential equation (3.01) are analytic functions of the complex variable x. We first map the so-called *principal curves* (or *anti-Stokes lines*)

$$\operatorname{Re}\left\{\int_{x_0}^{x} f^{1/2}(u, y)\, dy\right\} = 0 \tag{6.01}$$

in the complex x-plane, where x_0 denotes the turning point. Since $f(u, x)$ has a simple zero at x_0, three of these curves emerge from this point, dividing the x-plane into three regions, which we refer to as the *principal subdomains* corresponding to x_0. Associated with each principal subdomain there is a solution of the differential equation which is recessive as $x \to \infty$ within the subdomain, and dominant as $x \to \infty$ in the other two subdomains. Except in the neighborhood of the turning point, each of the three solutions can be approximated uniformly in terms of elementary functions by the Liouville–Green theory. The linear identity holding between the three solutions is, in effect, the required connection formula. Because two of the solutions are dominant in any chosen principal subdomain, they have to cancel each other in the connection formula as $x \to \infty$ in this subdomain. This condition yields exactly the right number of equations to determine the coefficients in the connection formula. An attractive feature of this

method is that approximations in terms of elementary functions are continued around the turning point without having to construct approximations in terms of nonelementary functions valid in domains that include the turning point. Another advantage is that a simpler error-control function is used in constructing error bounds (although the variation of this function has to be calculated along curves in the complex plane instead of the real axis).

Wasow [29] appropriately has called the first type of method *central connection*, and the second type *lateral connection*. Detailed accounts of rigorous methods of both types are included in [23, Chap. 13]. These references, and also [11], [9], [6, Chap. 1], [3, Chap. 1], [30], [17] and [1, pp. 292–295], sketch the history and modifications of the methods.

6.2. Multiple turning points. As we have seen in the present paper, the method of central connection can still be used for double and higher turning points. A true analogue of the rigorous lateral connection procedure for a simple turning point has yet to be found, however. All published methods, including those mentioned below, that continue an approximate solution around a multiple turning point in the complex plane are based, directly or indirectly, on approximate representations in terms of Bessel functions valid at the turning point. To distinguish these complex-variable methods from direct central connection procedures along the real axis, we shall call them *pseudo-lateral* methods.

Another way in which the connection formula problem for simple turning points differs from that for multiple turning points is that in the former case, the uniform asymptotic approximations for the solutions of the differential equation can be extended in straightforward manner into uniform asymptotic expansions in descending powers of the large parameter u; see [23, Chap. 11]. Although asymptotic expansions in descending powers of u can be constructed for the solutions in regions containing a multiple turning point [18], [27], [16], [31], in marked contrast to the case of a simple turning point, these expansions are uniformly valid only for bounded values of the independent variable; compare [24, § 11]. Moreover, in the case of [18], [27] and [16] the expansions are in terms of functions of two or more variables.

The earliest paper deriving connection formulas for multiple turning points appears to be that of Goldstein [4]. He treats equation (3.01) with $g(u, x) = 0$ and $f(u, x) \equiv f(x)$ independent of x. The zero x_0 of $f(x)$ may have any multiplicity. Goldstein's method is an extension of the procedure successfully employed by Jeffreys for a simple turning point [8]. Thus $f(x)$ is replaced by the first nonvanishing term $f^{(m-2)}(x_0)(x-x_0)^{m-2}/(m-2)!$ in its Taylor series expansion at x_0, where $m-2$ is the multiplicity of the zero. The differential equation is then solvable exactly in terms of Bessel functions (or modified Bessel functions), as we saw in § 2. The connection formulas are found by replacing the Bessel functions by their asymptotic approximations for large positive and negative arguments. Goldstein's connection formulas agree with those obtained by rigorous analysis in the present paper.[12]

[12] Misprints on pp. 84 and 86 of [4] are corrected in [5].

A second paper by Goldstein [5] treats the case in which $f(u, x)$ and $u^{-1}g(u, x)$ are independent of x, and $f(u, x)$ has a double zero at x_0. As we perceived in § 5.3, Bessel functions are no longer adequate approximants to the solutions in these circumstances. By using parabolic cylinder functions instead, Goldstein is able to derive connection formulas by analysis similar to that of [8] and [4]. A somewhat more general version of the same problem was treated subsequently by Langer [12]. Langer's analysis is rigorous, and he supplies uniform asymptotic approximations for the solutions in the complex x-plane. It may be noted, incidentally, that recent results of the present writer for the case of two coalescing simple turning points [24] are applicable to this second problem of Goldstein is able to derive connection formulas by analysis similar to that of [8].

Returning to Goldstein's first problem, we note that almost all subsequent writers have espoused the pseudo-lateral approach, deriving rules for the discontinuous changes in the coefficients in the Liouville–Green approximations as the solutions pass from one principal subdomain in the complex plane to the next. The first systematic treatment of this kind for a turning point of arbitrary multiplicity appears to be the formal analysis of Heading on pp. 89–93 and pp. 110–115 of [6]. Similar analysis was given subsequently by Fröman and Fröman [3, pp. 69–74], and also Evgrafov and Fedoryuk [2, pp. 34–35].[13]

Apart from [12], rigorous analyses of the pseudo-lateral methods begin with papers of Lee [13] and Leung [14], [15], which establish the connection formulas for adjacent principal subdomains for a double turning point. Extensions to turning points of any multiplicity have been made by Nishimoto [22], Sibuya [27] and Leung [16].

6.3. Conclusions. The present paper is essentially a rigorous formulation, with extensions, of the results of Goldstein's first paper [4]. For connection formula problems along the real axis, this approach has the following advantages over the complex-variable approach adopted in the references mentioned in § 6.2:

(i) The coefficients $f(u, x)$ and $g(u, x)$ in the given differential equation (3.01) need not be analytic functions of the complex variable x. It suffices that $\partial^2 f(u, x)/\partial x^2$ and $g(u, x)$ are continuous functions of the real variable x (and even this condition could be relaxed to some extent).

(ii) There is no need to investigate the topology in the complex plane of the principal curves (6.01).

(iii) A solution can be continued along the real axis through the turning point by means of a *single* connection formula. With the complex-variable approach, there are m principal subdomains associated with a turning point of multiplicity

[13] Instead of drawing on known properties of Bessel functions in discussing the equation $d^2w/dt^2 = t^{m-2}w$, Fröman and Fröman, and Evgrafov and Fedoryuk, derive the necessary connection formulas *ab initio* by elementary analysis. However, this analysis hinges upon the Fuchs–Frobenius theory of the differential equation as $t \to 0$, and can be regarded simply as a way of establishing the connection formulas for the Bessel functions. Thus the connection procedure of these authors is properly classified as central and not lateral. Incidentally, although the treatment of multiple turning points in [2] and [3] is formal, the analysis of simple turning points in both references is rigorous. It may also be noted that [2] includes a systematic study of the topology of the principal curves (6.01) in the complex plane when any number of turning points are present.

$m-2$, and a total of almost $[\frac{1}{2}m]$ successive steps from one principal subdomain to the next is needed to achieve the connection along the real axis.

(iv) Explicit and realistic bounds are available for the error terms.

Acknowledgment. Figures 2.1–2.8 were plotted mechanically from values of the solutions obtained by numerical integration of the differential equations (2.01), (2.11) and (2.31). The programs for the calculations and plotting were constructed by Mr. R. E. Kaylor, and the author is pleased to acknowledge this assistance.

REFERENCES

[1] R. B. DINGLE, *Asymptotic Expansions: Their Derivation and Interpretation*, Academic Press, New York, 1973.

[2] M. A. EVGRAFOV AND M. V. FEDORYUK, *Asymptotic behaviour as $\lambda \to \infty$ of the solution of the equation $w''(z)-p(z,\lambda)w(z)=0$ in the complex z-plane*, Russian Math. Surveys, 21 (1966), pp. 1–48.

[3] N. FRÖMAN AND P. O. FRÖMAN, *JWKB Approximation—Contributions to the Theory*, North-Holland, Amsterdam, 1965.

[4] S. GOLDSTEIN, *A note on certain approximate solutions of linear differential equations of the second order, with an application to the Mathieu equation*, Proc. London Math. Soc. [2], 28 (1928), pp. 81–90.

[5] ———, *A note on certain approximate solutions of linear differential equations of the second order* (2), Proc. London Math. Soc. [2], 33 (1932), pp. 246–252.

[6] J. HEADING, *An Introduction to Phase-integral Methods*, John Wiley, New York, 1962.

[7] V. B. HEADLEY AND V. K. BARWELL, *On the distribution of the zeros of generalized Airy functions*, Math. Comp., 29 (1975), pp. 863–877.

[8] H. JEFFREYS, *On certain approximate solutions of linear differential equations of the second order*, Proc. London Math. Soc., [2], 23 (1924), pp. 428–436.

[9] ———, *On the use of asymptotic approximations of Green's type when the coefficient has zeros*, Proc. Cambridge Philos. Soc., 52 (1956), pp. 61–66.

[10] N. D. KAZARINOFF AND R. K. RITT, *Scalar diffraction theory and turning-point problems*, Arch. Rational Mech. Anal., 5 (1960), pp. 177–186.

[11] R. E. LANGER, *The asymptotic solutions of ordinary linear differential equations of the second order, with special reference to the Stokes phenomenon*, Bull. Amer. Math. Soc., 40 (1934), pp. 545–582.

[12] ———, *The asymptotic solutions of certain linear ordinary differential equations of the second order*, Trans. Amer. Math. Soc., 36 (1934), pp. 90–106.

[13] 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.

[14] A. W-K. LEUNG, *Connection formulas for asymptotic solutions of second order turning points in unbounded domains*, this Journal, 4 (1973), pp. 89–103.

[15] ———, *Errata: Connection formulas for asymptotic solutions of second order turning points in unbounded domains*, this Journal, 6 (1975), p. 600.

[16] ———, *Lateral connections for asymptotic solutions around higher order turning points*, J. Math. Anal. Appl., 50 (1975), pp. 560–578.

[17] J. A. M. MCHUGH, *An historical survey of ordinary linear differential equations with a large parameter and turning points*, Arch. History Exact Sci., 7 (1971), pp. 277–324.

[18] R. W. MCKELVEY, *The solutions of second order linear ordinary differential equations about a turning point of order two*, Trans. Amer. Math. Soc., 79 (1955), pp. 103–123.

[19] J. C. P. MILLER, *On the choice of standard solutions for a homogeneous linear differential equation of the second order*, Quart. J. Mech. Appl. Math., 3 (1950), pp. 225–235.

[20] ———, *Tables of Weber Parabolic Cylinder Functions*, H. M. Stationery Office, London, 1955.

[21] 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, D.C., 1964.

[22] T. NISHIMOTO, *On an extension theorem and its application for turning point problems of large order*, Kōdai Math. Sem. Rep., 25 (1973), pp. 458–489.

[23] F. W. J. OLVER, *Asymptotics and Special Functions*, Academic Press, New York, 1974.

[24] ———, *Second-order linear differential equations with two turning points*, Philos. Trans. Roy. Soc. London Ser. A, 278 (1975), pp. 137–174.

[25] ———, *Second-order differential equations with fractional transition points*, Trans. Amer. Math. Soc., to appear.

[26] ———, *Connection formulas for second-order differential equations having an arbitrary number of turning points of arbitrary multiplicities*, this Journal, to appear.

[27] Y. SIBUYA, *Uniform simplification in a full neighborhood of a transition point*, Mem. Amer. Math. Soc., no. 149, (1974).

[28] C. A. SWANSON AND V. B. HEADLEY, *An extension of Airy's equation*, SIAM J. Appl. Math., 15 (1967), pp. 1400–1412.

[29] W. WASOW, *Connection problems for asymptotic series*, Bull. Amer. Math. Soc., 74 (1968), pp. 831–853.

[30] ———, *Simple turning-point problems in unbounded domains*, this Journal, 1 (1970), pp. 153–170.

[31] B. WILLNER AND L. A. RUBENFELD, *Uniform asymptotic solutions for a linear ordinary differential equation with one μ-th order turning point: analytic theory*, Comm. Pure Appl. Math., 29 (1976), pp. 343–367.

SIAM J. MATH. ANAL.
Vol. 8, No. 1, February 1977

WEAK SOLUTIONS OF THE TIME-DEPENDENT CONTINUOUS-ENERGY EQUATION OF NEUTRON DIFFUSION*

JOHN R. CANNON† AND PAUL NELSON‡

Abstract. Existence, uniqueness, continuous dependence upon the data, and nonnegativity of the weak solution is demonstrated for the solution of an initial-boundary value problem for the continuous energy diffusion approximation to the neutron transport equation.

1. Introduction. Pao [1] has recently studied solutions of classical type for the time-dependent continuous-energy version of the neutron diffusion equation. The continuous-energy diffusion approximation to the neutron transport equation is frequently termed the "P_1 equation" [2]. In the present paper we study weak solutions to the time-dependent P_1 equation. Hlaváček [3] has also been motivated by neutron diffusion theory to consider weak solutions of certain operator equations; however his results seem to require in an essential way the appearance of the second time derivative in the underlying equation, and therefore are presumably not applicable to the "parabolic" equation considered in the present paper.

In the next section we introduce the initial-boundary value problem and its weak formulation. Section 3 is devoted to the study of weak solutions of an auxiliary parabolic problem containing a parameter. Application of the results of § 3 is made in § 4 to obtain existence and uniqueness. Properties of the solution are discussed in § 5.

2. The problem. We write the time-dependent continuous-energy neutron diffusion equation in the form

$$\begin{aligned}\frac{\partial\varphi}{\partial t}-v(E)\nabla\cdot(D\nabla\varphi)+v(E)\Sigma\varphi&\\ =\int_0^\infty v(E)k(x,t,E,E')\varphi(x,t,E')\,dE'&\\ +v(E)q(x,t,E),\qquad (x,t,E)\in\tilde{Q}_T=Q_T\times(0,\infty).&\end{aligned}\tag{2.1}$$

Here t is time, $x=(x_1,\cdots,x_n)$ is in the normal domain Ω, E is energy, ∇ denotes the gradient operator in the spatial variable x, $Q_T=\Omega\times\{t:0\leqq t\leqq T\}$ for some fixed $T>0$, the given functions $v(E)$, $D(x,t,E)$, $\Sigma(x,t,E)$, $q(x,t,E)$ and $k(x,t,E,E')$ represent respectively neutron speed, diffusion coefficient, total scattering cross-section, external source and the expected density of neutrons of energy E resulting from a neutron of energy E' undergoing a collision at (x,t), and the neutron flux $\varphi(x,t,E)$ is to be determined. All omitted arguments in (2.1) are to be understood as (x,t,E). Note that the function k includes consideration of all types of collisional events, including fission, scattering and absorption. The initial

* Received by the editors July 24, 1975, and in revised form May 25, 1976.
† Department of Mathematics, University of Texas at Austin, Austin, Texas 78712.
‡ Department of Mathematics, Texas Tech University, Lubbock, Texas 79409.

and boundary conditions for (2.1) will be taken respectively as

$$\varphi(x, 0, E) = \varphi_0(x, E)(\geqq 0), \qquad (x, E) \in \tilde{\Omega} = \Omega \times (0, \infty), \tag{2.2}$$

and

$$\varphi(x, t, E) = 0, \qquad (x, t, E) \in \tilde{S}_T = S_T \times (0, \infty), \tag{2.3}$$

where $S_T = \partial\Omega \times \{t : 0 < t \leqq T\}$.

With regard to the given functions in (2.1)–(2.3) we make the following assumptions.

(H1) $v(E)$ is a nonnegative measurable function defined for $E \in (0, \infty)$. (It is important to the physics of the problem that v not be required to be bounded either above or away from zero.)

(H2) $D = D(x, t, E)$ is a positive measurable function defined on $\tilde{Q}_T$ such that the product $v(E)D(x, t, E)$ is bounded above and away from zero on $\tilde{Q}_T$.

(H3) $\Sigma = \Sigma(x, t, E)$ is a nonnegative measurable function on $\tilde{Q}_T$ such that the product $v(E)\Sigma(x, t, E)$ is bounded above on $\tilde{Q}_T$.

(H4) $k = k(x, t, E, E')$ is a nonnegative measurable function defined on $\tilde{Q}_T x(0, \infty)$ such that

$$\operatorname*{ess\,sup}_{(x,t)\in Q_T} \int_{(0,\infty)\times(0,\infty)} [v(E)k(x, t, E, E')]^2 \, dE \, dE' = K < \infty. \tag{2.4}$$

(H5) $\varphi_0 = \varphi_0(x, E)$ is a nonnegative measurable square-integrable function on $\tilde{\Omega} = \Omega \times (0, \infty)$.

(H6) $q = q(x, t, E)$ is a nonnegative measurable function defined on $\tilde{Q}_T$ such that $v(E)q(x, t, E)$ is square-integrable over $\tilde{Q}_T$.

DEFINITION 1. We denote by $U(T)$ the Hilbert space which is the closure under the norm defined by

$$\|\varphi\|^2_{U(T)} = \int_{\tilde{Q}_T} [\varphi^2 + \nabla\varphi \cdot \nabla\varphi] \, dx \, dt \, dE \tag{2.5}$$

of the functions $\varphi \in C^\infty(\tilde{Q}_T)$ which vanish on $\tilde{S}_T$.

DEFINITION 2. By $U_1(T)$ we denote the subspace of U consisting of those functions with a square integrable generalized t-derivative on $\tilde{Q}_T$.

DEFINITION 3. A *weak solution* of the system (2.1)–(2.3) is a function $\varphi \in U(T)$ such that

$$\begin{aligned} &-\int_{\tilde{Q}_T} \eta_t \varphi \, dx \, dt \, dE + \int_{\tilde{Q}_T} vD\nabla\eta \cdot \nabla\varphi \, dx \, dt \, dE + \int_{\tilde{Q}_T} v\Sigma\eta\varphi \, dx \, dt \, dE \\ &= \int_{\tilde{Q}_T\times(0,\infty)} \eta(x, t, E)v(E)k(x, t, E, E')\varphi(x, t, E') \, dE' \, dx \, dt \, dE \\ &\quad + \int_{\tilde{Q}_T} \eta vq \, dx \, dt \, dE + \int_{\tilde{\Omega}} \varphi_0(x, E)\eta(x, 0, E) \, dx \, dE \end{aligned} \tag{2.6}$$

for all functions $\eta \in U_1(T)$ such that $\eta(x, T, E) \equiv 0$, $(x, E) \in \tilde{\Omega}$.

Remark. Equation (2.6) is obtained from (2.1)–(2.3) in the usual manner.

3. A parabolic partial differential equation with a parameter. We consider first the problem of the determination of the weak solution $u = u(x, t, E)$ of

$$\begin{aligned} &\frac{\partial u}{\partial t} - v(E)\nabla \cdot (D\nabla u) + v(E)\Sigma u = f(x, t, E), \qquad (x, t, E) \in \tilde{Q}_T, \\ &u(x, t, E) = 0, \qquad (x, t, E) \in \tilde{S}_T, \\ &u(x, 0, E) = g(x, E), \qquad (x, E) \in \tilde{\Omega}, \end{aligned} \tag{3.1}$$

where v, D and Σ satisfy H1, H2, and H3 respectively, g replaces φ_0 in (H5), and f replaces both q and vq in (H6). By the usual integration by parts, we make the following

DEFINITION 4. A weak solution of (3.1) is a function $u \in U(T)$ such that

$$\begin{aligned} &-\int_{\tilde{Q}_T} \eta_t u \, dx \, dt \, dE + \int_{\tilde{Q}_T} vD\nabla\eta \cdot \nabla u \, dx \, dt \, dE + \int_{\tilde{Q}_T} v\Sigma\eta u \, dx \, dt \, dE \\ &= \int_{\tilde{Q}_T} \eta f \, dx \, dt \, dE + \int_{\tilde{\Omega}} g(x, E)\eta(x, 0, E) \, dx \, dE \end{aligned} \tag{3.2}$$

for all functions $\eta \in U_1(T)$ such that $\eta(x, T, E) \equiv 0$, $(x, E) \in \tilde{\Omega}$.

We begin our study of weak solutions of (3.1) by consideration of the method of Galerkin approximations.

Let $\mathring{W}_2^{(1)}(\Omega)$ be the closure of $C_0^\infty(\Omega)$ (C^∞-functions with compact support in Ω) under the norm defined by

$$[\|\psi\|_{2,\Omega}^{(1)}]^2 = \int_\Omega [\psi^2(x) + \nabla\psi(x) \cdot \nabla\psi(x)] \, dx. \tag{3.3}$$

Obviously $\mathring{W}_2^{(1)}(\Omega)$ is a Hilbert space with the norm defined by (3.3) and generated by the inner product

$$(f, g)_{2,\Omega}^{(1)} = \int_\Omega [f(x)g(x) + \nabla f(x) \cdot \nabla g(x)] \, dx. \tag{3.4}$$

Denote by $\{\psi_i(x)\}_{i=1}^\infty$ a countable basis in $\mathring{W}_2^{(1)}(\Omega)$. (By *basis* we intend a set whose linear combinations are dense.) For convenience it will be assumed the collection $\{\psi_i\}$ is orthonormalized in $L^2(\Omega)$. We look for approximate solutions of (3.1) in the form

$$u^N(x, t, E) = \sum_{i=1}^N \beta_i(t, E)\psi_i(x) \tag{3.5}$$

where the β_i are determined by the system

$$\begin{aligned} &\frac{d\beta_i}{dt} + \int_\Omega v(E)D \sum_{j=1}^N \beta_j \nabla\psi_j \cdot \nabla\psi_i \, dx + \int_\Omega v(E)\Sigma \sum_{j=1}^N \beta_j\psi_j\psi_i \, dx \\ &\qquad = f_i(t, E), \quad (t, E) \in (0, T) \times (0, \infty), \quad i = 1, \cdots, N, \end{aligned} \tag{3.6}$$

$$\beta_i(0, E) = g_i(E), \tag{3.7}$$

where

$$f_i = \int_\Omega \psi_i f\,dx \quad \text{and} \quad g_i = \int_\Omega \psi_i g\,dx. \tag{3.8}$$

By a solution of the system (3.6) we mean a sequence $\beta_i(t, E)$, $i = 1, \cdots, N$, of jointly measurable functions on $[0, T]\times(0, \infty)$ such that:

(i) for almost all $E\in(0, \infty)$, each $\beta_i(t, E)$ is absolutely continuous in $t\in[0, T]$;

(ii) for almost every $t\in[0, T]$, each $\beta_i(t, E)$ is in $L^2(0, \infty)$ as a function of E;

(iii) for almost all $E\in(0, \infty)$ equation (3.6) holds almost everywhere in $t\in[0, T]$;

(iv) equation (3.7) holds for almost all $E\in(0, \infty)$.

We denote by $L_n^2(0, \infty)$ (respectively $L_n^2([0, T]\times(0, \infty))$) the set of n-vectors whose components are functions square-integrable over $(0, \infty)$ (respectively $[0, T]\times(0, \infty)$). These function classes are Hilbert spaces under the respective inner products

$$(w, y)_{L_n^2(0,\infty)} = \sum_{i=1}^{n} \int_{(0,\infty)} w_i(E)y_i(E)\,dE, \tag{3.9}$$

and

$$(w, y)_{L_n^2([0,T]\times(0,\infty))} = \sum_{i=1}^{n} \int_{[0,T]\times(0,\infty)} w_i(\tau, E)y_i(\tau, E)\,d\tau\,dE. \tag{3.10}$$

We further denote by $V_n(T)$ the subclass of $w\in L_n^2([0, T]\times(0, \infty))$ such that the mapping $t\to w(t, \cdot)$ is continuous from $[0, T]$ to $L_n^2(0, \infty)$. This set is a Banach space under the norm

$$|w|_{V_n(T),\sigma} = \max_{0\le t\le T} e^{-\sigma t}\|w(t, \cdot)\|_{L_n^2(0,\infty)} \tag{3.11}$$

for any real σ. We also consider the class $B_n(T)$ of $w\in L_n^2([0, T]\times(0, \infty))$ such that the mapping $t\to w(t, \cdot)$ is essentially bounded from $[0, T]$ to $L_n^2(0, \infty)$. The set $B_n(T)$ is a Banach space under the norm

$$|w|_{B_n(T),\sigma} = \operatorname*{ess\,sup}_{0\le t\le T} e^{-\sigma t}\|w(t, \cdot)\|_{L_n^2(0,\infty)} \tag{3.12}$$

for any real σ. The main result of this section can now be stated as follows.

THEOREM 1. *Under the above assumptions and notation, the system* (3.6) *and* (3.7) *has a unique solution* $\beta = (\beta_i)$. *Furthermore* $\beta\in V_N(T)$, *and the correspondence* $(f, g)\to\beta$ *taking the data into the solution is continuous from* $(L_N^2([0, T]\times(0, \infty)))\times L_N^2(0, \infty)$ *to* $V_N(T)$.

Proof. If $\beta = (\beta_i)$ is a solution of (3.6) and (3.7), then clearly the equations

$$\begin{aligned}\beta_i(t, E) = \int_0^t \Big\{ f_i(\tau, E) - \int_\Omega v(E)D \sum_{j=1}^{N} \beta_j(\tau, E)\nabla\psi_j\cdot\nabla\psi_i\,dx \\ - \int_\Omega v(E)\Sigma \sum_{j=1}^{N} \beta_j(\tau, E)\psi_j\psi_i\,dx \Big\}\,d\tau + g_i(E), \qquad i = 1, \cdots, N,\end{aligned} \tag{3.13}$$

hold everywhere in $t \in [0, T]$ except possibly for those E in a null subset of $(0, \infty)$. Conversely if (3.13) holds for all $t \in [0, T]$ except for E in a null subset of $(0, \infty)$, and each $\beta_i(t, E)$ is known a priori to be square-integrable in E for almost all $t \in [0, T]$, then β is a solution of (3.6)–(3.7). Thus the system (3.13) of integral equations is completely equivalent to the initial-value problem (3.6)–(3.7) within the framework of our definition of solution.

We denote the right-hand side of (3.13) by $(\mathscr{L}\beta)_i(t, E)$, when it exists. Our immediate aim is to prove that $\mathscr{L}$ is a contractive mapping on $B_N(T)$. Let M be an upper bound for the integrals

$$\int_\Omega vD\nabla\psi_j \cdot \nabla\psi_i \, dx, \int_\Omega v\Sigma\psi_j\psi_i \, dx, \qquad i, j = 1, \cdots, N.$$

Suppose $\gamma = \gamma(t, E) \in B_N(T)$. For $t \in [0, T]$ we then compute

$$\begin{aligned} \|\mathscr{L}\gamma(t, \cdot)\|_{L^2_N(0,\infty)} &\leqq \sqrt{t}\|f\|_{L^2_N([0,T]\times(0,\infty))} \\ &\quad + 2MN\sqrt{t}\|\gamma\|_{L^2_N([0,T]\times(0,\infty))} + \|g\|_{L^2_N(0,\infty)}, \end{aligned} \tag{3.14}$$

which shows $\mathscr{L}\gamma$ is also in $B_N(T)$. If β, γ are both functions in $B_N(T)$ we estimate

$$|\mathscr{L}\beta - \mathscr{L}\gamma|_{B_N(T),\sigma} \leqq 2MNT^{1/2}\sigma^{-1/2}|\beta - \gamma|_{B_N(T),\sigma}.$$

This shows $\mathscr{L}$ is a contraction mapping in $B_N(T)$ relative to the norm (3.12) for sufficiently large σ.

It now follows from the contraction mapping theorem that (3.13) has a unique solution, β, in $B_N(T)$. In order to see that $\beta \in V_N(T)$ note that the inequality

$$\begin{aligned} \|\beta(t + \Delta t, \cdot) - \beta(t, \cdot)\|_{L^2_N(0,\infty)} &\leqq \sqrt{\Delta t}\,[\|f\|_{L^2_N([0,T]\times(0,\infty))} \\ &\quad + 2MN\|\beta\|_{L^2_N([0,T]\times(0,\infty))}] \end{aligned} \tag{3.15}$$

follows from (3.13) and (3.14). It follows immediately from inequality (3.15) that the mapping $t \to \beta(t, \cdot)$ is continuous from $[0, T]$ to $L^2_N(0, \infty)$.

It remains only to show continuous dependence on the data as asserted in Theorem 1. Note, as observed above, that $\beta \in B_N(T)$ and satisfying (3.13) implies it is a solution of the initial-value system (3.6) and (3.7) in the sense defined previously. If we multiply (3.6) by β_i, sum over i from 1 to N, and integrate on time from 0 to t, we get the equation

$$\begin{aligned} &\sum_{i=1}^N \beta_i(t, E)^2 + 2\int_0^t \int_\Omega v(E)D(x, \tau, E)\left(\sum_{i=1}^N \beta_i(\tau, E)\nabla\psi_i\right) \cdot \left(\sum_{i=1}^N \beta_i(\tau, E)\nabla\psi_i\right) dx\, d\tau \\ &\qquad + 2\int_0^t \int_\Omega v(E)\Sigma(x, \tau, E)\left(\sum_{i=1}^N \beta_i(\tau, E)\psi_i\right)^2 dx \\ &\qquad\qquad = \sum_{i=1}^N g_i^2(E) + 2\int_0^t \sum_{i=1}^N f_i(\tau, E)\beta_i(\tau, E)\, d\tau, \end{aligned} \tag{3.16}$$

which holds for all (t, E) such that $r \in [0, T]$ and $E \in (0, \infty) \sim A$, where the exceptional set A has measure zero. From (3.16) and the nonnegativity of vD and $v\Sigma$ we obtain the estimate

$$\|\beta(t, \cdot)\|^2_{L^2_N(0,\infty)} \leqq \|g\|^2_{\tilde{\Omega}} + \|f\|^2_{\tilde{Q}_t} + \int_0^t \|\beta(\tau, \cdot)\|^2_{L^2_N(0,\infty)}\, d\tau. \tag{3.17}$$

From (3.17) and the Bellman–Gronwall lemma there now follows the inequality

$$\|\beta(t, \cdot)\|^2_{L^2_N(0,\infty)} \leqq e^t\{\|g\|^2_{\tilde{\Omega}} + \|f\|^2_{\tilde{Q}_t}\}. \tag{3.18}$$

The continuous dependence upon data asserted in Theorem 1 follows immediately from the inequality (3.18) and linearity of the problem (3.6)–(3.7) (or (3.13)). Hence, Theorem 1 is finished.

Set

$$\eta^P = \sum_{i=1}^{P} \alpha_i(t, E)\psi_i(x), \qquad P \leqq N. \tag{3.19}$$

where α_i are smooth functions which vanish when $t = T$ and $\partial\alpha_i/\partial t \in L^2((0, T) \times (0, \infty))$. Multiplying (3.6) by α_i and summing, integrating with respect to t over $(0, T]$ and performing an integration by parts on the first term, and finally integrating with respect to E over $(0, \infty)$ yields

$$\begin{aligned} -\int_{\tilde{Q}_T} \eta_t^P u^N\, dx\, dt\, dE + \int_{\tilde{Q}_T} vD\nabla\eta^P \cdot \nabla u^N\, dx\, dt\, dE + \int_{\tilde{Q}_T} v\Sigma\eta^P u^N\, dx\, dt\, dE \\ = \int_{\tilde{Q}_T} \eta^P f\, dx\, dt\, dE + \int_{\tilde{\Omega}} g(x, E)\eta^P(x, 0, E)\, dx\, dE. \end{aligned} \tag{3.20}$$

for all N and P with $P \leqq N$. The inequality (3.18) and (H2) applied to (3.16) yields the fact that

$$\|u^N\|_{U(T)} \leqq C_1, \tag{3.21}$$

where C_1 is a positive constant independent of N. From the weak compactness of the Hilbert space $U(T)$, there exists a $u \in U(T)$ which satisfies (3.20) for each P. Since the η^P are dense in $U_1(T)$, it follows that u is a weak solution of (3.1). Moreover, since (3.18) yields

$$\operatorname*{ess\,sup}_{0 \leqq t \leqq T} \|u^N(\cdot, t, \cdot)\|_{L^2(\tilde{\Omega})} \leqq C_2 \tag{3.22}$$

where C_2 is a positive constant independent of N, clearly

$$\operatorname*{ess\,sup}_{0 \leqq t \leqq T} \|u(\cdot, t, \cdot)\|_{L^2(\tilde{\Omega})} \leqq C_2. \tag{3.23}$$

In addition, it is clear that the argument of [4, pp. 156–159] can be utilized to show that the mapping $t \to u(\cdot, t, \cdot) \in L^2(\tilde{\Omega})$ is continuous for all $t \in [0, T]$. From the

continuity it follows from [4, pp. 141–143] that

$$\frac{1}{2}\int_{\tilde{\Omega}} [u(x,t,E)]^2\,dx\,dE + \int_0^t\int_{\tilde{\Omega}} vD\nabla u\cdot\nabla u\,dx\,dE\,dt + \int_0^t\int_{\tilde{\Omega}} v\Sigma u^2\,dx\,dE\,dt = \frac{1}{2}\int_{\tilde{\Omega}} g^2\,dx\,dE + \int_0^t\int_{\tilde{\Omega}} fu\,dx\,dE\,dt. \tag{3.24}$$

An application of the Bellman–Gronwall inequality yields

$$\|u(\cdot,t,\cdot)\|^2_{L^2(\tilde{\Omega})} \leqq e^t\{\|g\|^2_{L^2(\tilde{\Omega})} + \|f\|^2_{L^2(\tilde{Q}_T)}\}. \tag{3.25}$$

Clearly,

$$\operatorname*{ess\,sup}_{0\leqq t\leqq T} \|u(\cdot,t,\cdot)\|^2_{L^2(\tilde{\Omega})} \leqq e^T\{\|g\|^2_{L^2(\tilde{\Omega})} + \|f\|^2_{L^2(\tilde{Q}_T)}\}, \tag{3.26}$$

$$\|u\|^2_{L^2(\tilde{Q}_T)} \leqq Te^T\{\|g\|^2_{L^2(\tilde{\Omega})} + \|f\|^2_{L^2(\tilde{Q}_T)}\}, \tag{3.27}$$

and

$$\|u\|^2_{U(T)} \leqq \text{const.}\cdot\{\|g\|^2_{L^2(\tilde{\Omega})} + \|f\|^2_{L^2(\tilde{Q}_T)}\}. \tag{3.28}$$

where the constant depends on T.

In summation, we can state the following result.

THEOREM 2. *Under the assumptions upon the data which are described following* (3.1), *there exists a unique weak solution of* (3.1) *which satisfies* (3.26)–(3.28).

4. Existence and uniqueness of the solution to the problem. Let $\psi(x,t,E)\in L^2(\tilde{Q}_T)$ and consider the function

$$F(x,t,E) = \int_0^\infty v(E)k(x,t,E,E')\psi(x,t,E')\,dE'. \tag{4.1}$$

From Fubini's theorem and Schwarz's lemma we can calculate

$$\begin{aligned}
\|F\|^2_{L^2(\tilde{Q}_T)} &= \int_0^T\int_\Omega\int_0^\infty\left\{\int_0^\infty v(E)k(x,t,E,E')\psi(x,t,E')\,dE'\right\}^2 dE\,dx\,dt\\
&\leqq \int_0^T\int_\Omega\int_0^\infty\left(\int_0^\infty v^2k^2\,dE'\right)\left(\int_0^\infty \psi^2\,dE'\right)dE\,dx\,dt\\
&= \int_0^T\int_\Omega\left\{\int_0^\infty\int_0^\infty v^2k^2\,dE'\,dE\right\}\left\{\int_0^\infty\psi^2\,dE'\right\}dx\,dt\\
&\leqq \operatorname*{ess\,sup}_{\substack{x\in\Omega\\ 0\leqq t\leqq T}}\left\{\int_0^\infty\int_0^\infty v^2k^2\,dE'\,dE\right\}\|\psi\|^2_{L^2(\tilde{Q}_T)}\\
&= K\|\psi\|^2_{L^2(\tilde{Q}_T)}\quad\text{(see (2.4)).}
\end{aligned} \tag{4.2}$$

By defining $f = vq + F$ and $g = \varphi_0$, the results of § 3 yield a unique weak solution u of (3.1). If $u = \mathcal{M}\psi$, this defines the map $\mathcal{M}: L^2(Q_T)\to \tilde{U}(T)\subset L^2(Q_T)$. Obviously, the map $\mathcal{M}$ is linear for $\varphi_0 = vq = 0$. For $u_i = \mathcal{M}\psi_i$, $i = 1, 2$, inequality

(3.27) yields

$$\|u_1-u_2\|^2_{L^2(\tilde{Q}_T)} \leqq T e^T K \|\psi_1-\psi_2\|^2_{L^2(\tilde{Q}_T)}. \tag{4.3}$$

Consequently, for T sufficiently small, $\mathcal{M}$ is a contraction of the Banach space $L^2(\tilde{Q}_T)$ into itself. Thus, there exists a unique fixed point $\varphi = \mathcal{M}\varphi$ for T sufficiently small but positive. Since φ is a weak solution of (3.1) and thus continuous as a map $t \to \varphi(\cdot, t, \cdot) \in L^2(\tilde{Q}_T)$, it follows that the solution φ can be continued as a fixed point of $\mathcal{M}$ to any positive T.

Since a fixed point of $\mathcal{M}$ is a weak solution of (2.1)–(2.3), we have demonstrated the following result.

THEOREM 3. *Under the assumptions* (H1)–(H6), *there exists a unique weak solution of* (2.1)–(2.3).

5. Properties of the solution. In this section we shall derive stability estimates for the solution and we shall discuss the continuous dependence of the solution upon the data and the nonnegativity of the solution.

We consider the stability estimate first. From (3.24), (4.1) and (4.2) with $f = vq + F$ and $g = \varphi_0$, it follows that

$$\begin{aligned}\|\varphi(\cdot, t, \cdot)\|^2_{L^2(\tilde{\Omega})} &\leqq \|\varphi_0\|^2_{L^2(\tilde{\Omega})} + \|vq\|^2_{L^2(\tilde{Q}_t)} \\ &\quad + (K+2)\int_0^t \|\varphi(\cdot, \tau, \cdot)\|^2_{L^2(\tilde{\Omega})}\, d\tau.\end{aligned} \tag{5.1}$$

From the Bellman–Gronwall inequality, we obtain

$$\|\varphi(\cdot, t, \cdot)\|^2_{L^2(\tilde{\Omega})} \leqq e^{(K+2)t}\{\|\varphi_0\|^2_{L^2(\tilde{\Omega})} + \|vq\|^2_{L^2(\tilde{Q}_t)}\} \tag{5.2}$$

and

$$\|\varphi\|^2_{L^2(\tilde{Q}_t)} \leqq t\, e^{(K+2)t}\{\|\varphi_0\|^2_{L^2(\tilde{\Omega})} + \|vq\|^2_{L^2(\tilde{Q}_t)}\}. \tag{5.3}$$

By using (3.24) coupled with (5.1), it is clear that

$$\|\varphi\|^2_{U(T)} \leqq \text{const.} \cdot \{\|\varphi_0\|^2_{L^2(\tilde{\Omega})} + \|vq\|^2_{L^2(\tilde{Q}_t)}\}, \tag{5.4}$$

where the constant depends upon T and K. Thus far, we have shown that the $L^2(\tilde{\Omega})$ norm of φ grows at most exponentially as t tends to infinity provided that $\|vq\|_{L^2(\tilde{Q}_t)}$ grows at most exponentially as t tends to infinity. Also, from the linearity of (2.1)–(2.3), we have shown that φ in the norm of $U(T)$ depends continuously upon the initial and source data. We shall consider next the dependence upon the coefficients vD and $v\Sigma$ and the kernel k.

Let φ_i, $i = 1, 2$ denote the weak solutions of (2.1)–(2.3) corresponding respectively to the data v_i, D_i, Σ_i, q_i, k_i, and φ_{0i}, $i = 1, 2$. It follows from elementary calculations that $z = \varphi_1 - \varphi_2$ satisfies (2.1)–(2.3) where vD, $v\Sigma$, vk, vq, and φ_0 are replaced by v_1D_1, $v_1\Sigma_1$, v_1k_1, $(\tilde{vq})$, and z_0, respectively, and

$$\begin{aligned}(\tilde{vq}) &= \nabla \cdot (v_1D_1 - v_2D_2)\nabla\varphi_2 + (v_1\Sigma_1 - v_2\Sigma_2)\varphi_2 \\ &\quad + (v_1q_1 - v_2q_2) + \int_0^\infty (v_1k_1 - v_2k_2)\varphi_2\, dE'.\end{aligned} \tag{5.5}$$

and $z_0 = \varphi_{0,1} - \varphi_{0,2}$.

In estimating the $(\tilde{vq})$ term in the energy relation for (2.1)–(2.3) which is analogous to (3.24), we must estimate the following term

$$
\begin{aligned}
(\tilde{vq}, z)_{L^2(\tilde{Q}_T)} = &-\int_{\tilde{Q}_T} (v_1 D_1 - v_2 D_2)\nabla\varphi_2 \cdot \nabla z \, dE \, dx \, dt \\
&+\int_{\tilde{Q}_T} (v_1\Sigma_1 - v_2\Sigma_2)\varphi_2 z \, dE \, dx \, dt \\
&+\int_{\tilde{Q}_T} (v_1 q_1 - v_2 q_2) z \, dE \, dx \, dt \\
&+\int_{\tilde{Q}_T} z\left\{\int_0^\infty (v_1 k_1 - v_2 k_2)\varphi_2 \, dE'\right\} dE \, dx \, dt \\
&= I_1 + I_2 + I_3 + I_4.
\end{aligned} \tag{5.6}
$$

First,

$$
\begin{aligned}
|I_1| &\leqq \operatorname*{ess\,sup}_{\substack{x\in\Omega \\ 0\leqq t\leqq T \\ E\in[0,\infty]}} |v_1 D_1 - v_2 D_2| \cdot \left| \int_{\tilde{Q}_T} \nabla\varphi_2 \cdot \nabla z \, dE \, dx \, dt \right| \\
&\leqq \text{const.} \cdot \operatorname*{ess\,sup}_{\substack{x\in\Omega \\ 0\leqq t\leqq T \\ E\in[0,\infty]}} |v_1 D_1 - v_2 D_2|,
\end{aligned} \tag{5.7}
$$

where an application of (5.4) shows that the constant depends upon T, K_i, $\|\varphi_{0i}\|_{L^2(\tilde{\Omega})}$ and $\|v_i q_i\|_{L^2(\tilde{Q}_T)}$, $i = 1, 2$. Likewise,

$$
|I_2| \leqq \text{const.} \cdot \operatorname*{ess\,sup}_{\substack{x\in\Omega \\ 0\leqq t\leqq T \\ E\in[0,\infty]}} |v_1\Sigma_1 - v_2\Sigma_2|, \tag{5.8}
$$

I_3 is handled like the original vq term in (5.1), and

$$
|I_4| \leqq \text{const.} \cdot \operatorname*{ess\,sup}_{\substack{x\in\Omega \\ 0\leqq t\leqq T}} \left\{\int_0^\infty \int_0^\infty (v_1 k_1 - v_2 k_2)^2 \, dE' \, dE\right\}. \tag{5.9}
$$

An application of the Bellman–Gronwall lemma and elementary estimates yields

$$
\begin{aligned}
\|\varphi_1 - \varphi_2\|^2_{U(T)} \leqq \text{const.} \Big\{ &\|\varphi_{01} - \varphi_{02}\|^2_{L^2(\tilde{\Omega})} \\
&+ \|v_1 q_1 - v_2 q_2\|_{L^2(\tilde{Q}_T)} + \operatorname*{ess\,sup}_{\substack{x\in\Omega \\ 0\leqq t\leqq T \\ E\in[0,\infty]}} |v_1 D_1 - v_2 D_2| \\
&+ \operatorname*{ess\,sup}_{\substack{x\in\Omega \\ 0\leqq t\leqq T \\ E\in[0,\infty]}} |v_1\Sigma_1 - v_2\Sigma_2| \\
&+ \operatorname*{ess\,sup}_{\substack{x\in\Omega \\ 0\leqq t\leqq T}} \left\{\int_0^\infty \int_0^\infty (v_1 k_1 - v_2 k_2)^2 \, dE' \, dE\right\}\Big\}
\end{aligned} \tag{5.10}
$$

where the constant in (5.10) depends upon K_1, K_2, $\|\varphi_{01}\|_{L^2(\tilde{\Omega})}$, $\|\varphi_{02}\|_{L^2(\tilde{\Omega})}$,

$\|v_1q_1\|_{L^2(\tilde{Q}_T)}$, $\|v_2q_2\|_{L^2(\tilde{Q}_T)}$, T, and the lower bounds on v_1D_1 and v_2D_2. Hence, the solution φ of (2.1)–(2.3) depends continuously in the norm of $U(T)$ upon all of the data in the manner displayed in (5.10), where the constant on the right side of (5.10) is independent of the particular data when the norms of all the data are uniformly bounded.

We shall conclude this section with a discussion of the nonnegativity of φ almost everywhere. From the maximum principle for parabolic partial differential equations [4], it follows that for positive smooth vD, $v\Sigma$, f, and g and compact Ω with smooth boundary in (3.1), the solution $u(x, t, E)$ is nonnegative for each E. For smooth coefficients which are continuous with respect to a parameter in a linear parabolic partial differential equation, it follows that the parametrix and its x and t derivatives are continuous in the parameter. Consequently, the Green's function and its x and t derivatives are continuous. Thus, u and its x derivatives are continuous with respect to E. Consequently, u and its x derivatives are jointly measurable in x, t and E. Thus, for compact $\tilde{\Omega}$, u is also the weak solution in the sense of (3.2) for (3.1). Now, let u_m, $m = 1, 2, \cdots$, denote weak solutions of (3.1) which correspond respectively to the data $(vD)_m$, $(v\Sigma)_m$, f_m, and g_m, where each data function satisfies the relevant hypothesis set forth in § 2 and $(vD)_m$ and $(v\Sigma)_m$ are bounded uniformly by the respective bounds for (vD) and $(v\Sigma)$. Suppose $(vD)_m$ and $(v\Sigma)_m$ converge pointwise to (vD) and $(v\Sigma)$, respectively, as m tends to ∞. Also, suppose that

$$\lim_{m\to\infty} \|f-f_m\|_{L^2(\tilde{Q}_T)} = 0 \tag{5.11}$$

and

$$\lim_{m\to\infty} \|g-g_m\|_{L^2(\tilde{\Omega})} = 0. \tag{5.12}$$

Let u denote the weak solution of (3.1) corresponding to (vD), $(v\Sigma)$, f and g. Simple calculations yield

$$\begin{aligned} &\frac{1}{2}\int_{\tilde{\Omega}} [(u-u_m)(x, t, E)]^2\, dx\, dE + \int_{\tilde{Q}_T} (vD)_m |\nabla(u-u_m)|^2\, dx\, dt\, dE \\ &+ \int_{\tilde{Q}_T} (v\Sigma)_m (u-u_m)^2\, dx\, dt\, dE \\ &= \int_{\tilde{Q}_T} (u-u_m)(f-f_m)\, dx\, dt\, dE \\ &\quad + \int_{\tilde{\Omega}} (g-g_m)^2\, dx\, dE - \int_{\tilde{Q}_T} \{(vD)-(vD)_m\}\nabla u \cdot \nabla(u-u_m)\, dx\, dt\, dE \\ &\quad - \int_{\tilde{Q}_T} \{(v\Sigma)-(v\Sigma)_m\}u(u-u_m)\, dx\, dt\, dE. \end{aligned} \tag{5.13}$$

Using $ab \leqq \varepsilon a^2 + (1/\varepsilon)b^2$ in the third integral on the right side of (5.13) and the boundedness of $(vD)_m$ and (vD), it follows from the Bellman–Gronwall lemma

that

$$\|u-u_m\|^2_{L^2(\tilde{Q}_T)} \leqq \text{const.}\Big\{\|g-g_m\|_{L^2(\tilde{\Omega})}+\|f-f_m\|_{L^2(\tilde{Q}_T)} + \int_{\tilde{Q}_T} |(vD)-(vD)_m| \cdot |\nabla u|^2 \, dx\, dt\, dE + \int_{\tilde{Q}_T} |(v\Sigma)-(v\Sigma)_m| u^2 \, dx\, dt\, dE\Big\} \tag{5.14}$$

where the constant depends only on the bounds for vD, $(v\Sigma)$, and T. Consequently, from Lesbesgue's dominated convergence theorem, we see that

$$\lim_{m\to\infty} \|u-u_m\|_{L^2(\tilde{Q}_T)} = 0. \tag{5.15}$$

This implies that a subsequence of the u_m tends to u pointwise almost everywhere. As the $u_m \geqq 0$, $m = 1, 2, \cdots$, we have that $u \geqq 0$ almost everywhere. Thus, for compact $\tilde{Q}_T$ the solution u of (3.1) is nonnegative almost everywhere when $f \geqq 0$, $g \geqq 0$, and (vD) and $(v\Sigma)$ are pointwise limits almost everywhere of positive continuous functions that are respectively uniformly bounded by the bounds for (vD) and $(v\Sigma)$. Since simple functions are almost everywhere pointwise limits of continuous functions and simple functions are dense in L^1, it is clear that for measurable (vD) and $(v\Sigma)$ sequences of simple functions $(vD)_m$ and $(v\Sigma)_m$, $m = 1, 2, \cdots$, can be chosen such that the third and fourth terms in (5.14) can be made arbitrarily small as m tends to infinity. For example consider the integral involving (vD). Partition $\tilde{Q}_T$ into the sets $\{|\nabla u|^2 \leqq N\}$ and $\{|\nabla u|^2 > N\}$. Use the boundedness of (vD) and $(vD)_m$ for the integral of $|\nabla u|^2$ over $\{|\nabla u|^2 > N\}$. Fix N. Then use the L^1 convergence of $(vD)_m$ to (vD) to handle the integral over the set $\{|u|^2 \leqq N\}$. Consequently, for compact $\tilde{Q}_T$, the solution of (3.1) is nonnegative. As it is not difficult to show that the solution of (3.1) for noncompact $\tilde{Q}_T$ is the weak $U(T)$ limit of solutions of (3.1) for compact $\tilde{Q}_T$, we conclude that solutions of (3.1) are nonnegative almost everywhere. Since k is nonnegative and an iteration process for producing the fixed point of $\mathcal{M}$ can be initiated with $\varphi_0 \geqq 0$, it follows immediately that the weak solution of (2.1)–(2.3) is nonnegative.

In summary we have shown the following result.

THEOREM 4. *Under the hypotheses* (H1)–(H6), *the solution of* (2.1)–(2.3) *is nonnegative almost everywhere, depends continuously upon the data as shown in* (5.10) *and exhibits a growth of* $L^2(\tilde{\Omega})$ *norm as depicted in* (5.2).

REFERENCES

[1] C. V. PAO, *The energy-dependent diffusion system in nuclear reactor dynamics*, SIAM J. Appl. Math., 29 (1975), pp. 40–59.

[2] G. I. BELL AND S. GLASSTONE, *Nuclear Reactor Theory*, Van Nostrand Reinhold, New York, 1970.

[3] I. HLÁVAČEK, *On the existence and uniqueness of solutions of the Cauchy problem for linear integro-differential equations with operator coefficients*, Apl. Mat., 16 (1971), pp. 64–80.

[4] O. A. LADYŽENSKAJA, V. A. SOLONNIKOV AND N. N. URAL'CEVA, *Linear and Quasilinear Equations of Parabolic Type*, vol. 23, Translations of Mathematical Monographs, American Mathematical Society, Providence, R.I., 1968.

SIAM J. MATH. ANAL.
Vol. 8, No. 1, February 1977

INEQUALITIES FOR THE ZEROS OF BESSEL FUNCTIONS*

ROGER C. McCANN†

Abstract. Let $j_{p,n}$, $j'_{p,n}$ denote the nth positive zeros of J_p, J'_p respectively. It is shown that both $p^{-1}j_{p,n}$ and $p^{-1}j'_{p,n}$ are strictly decreasing functions of p.

1. Preliminary lemmas. We begin by considering the eigenvalue problem

$$-(xy')'+x^{-1}y=\lambda x^{2p-1}y, \qquad p>0, \tag{1}$$

$$y(a)=y(1)=0, \qquad 0<a<1. \tag{2}$$

It is easily verified that the general solution of (1) is $y(x)=C_1J_{1/p}(\lambda^{1/2}x^p/p)+C_2Y_{1/p}(\lambda^{1/2}x^p/p)$ and that the eigenvalues λ are solutions of

$$J_{1/p}(\lambda^{1/2}/p)-\frac{J_{1/p}(\lambda^{1/2}a^p/p)}{Y_{1/p}(\lambda^{1/2}a^p/p)}Y_{1/p}(\lambda^{1/2}/p)=0. \tag{3}$$

In particular, the nth eigenvalue of (1), (2) is the nth positive zero of (3).

LEMMA 1. *For $q>0$ and $0<a<1$ set*

$$f_{a,q}(x)=J_q(x)-\frac{J_q(a^{1/q}x)}{Y_q(a^{1/q}x)}Y_q(x).$$

Then

(i) *$f_{a,q}\to J_q$ uniformly as $a\to 0^+$ on any interval of the form $[\alpha,\beta]$ with $0<\alpha<\beta\leqq 1$.*

(ii) *$f'_{a,q}\to J'_q$ uniformly as $a\to 0^+$ on any interval of the form $[\alpha,\beta]$ with $0<\alpha<\beta\leqq 1$.*

(iii) *There exist $\varepsilon,\delta>0$ such that $f_{a,q}(x)>0$ and $J_q(x)>0$ for $a\in(0,\delta)$ and $x\in(0,\varepsilon]$.*

Proof. For $z\to 0^+$

$$J_q(z)\cong(2^q\Gamma(q+1))^{-1}z^q, \qquad Y_q(z)\cong-\frac{2^q\Gamma(q)}{\pi}z^{-q},$$

$$J'_q(z)\cong q(2^q\Gamma(q+1))^{-1}z^{q-1}, \qquad Y'_q(z)\cong\frac{q2^q\Gamma(q)}{\pi}z^{-q-1}.$$

Since $J_q(0)=0$ and $|Y_q(a^{1/q}x)|\to\infty$ as $a\to 0^+$ for every $x>0$, statement (i) is valid. Using the above asymptotic expansions it is easy to verify that

$$f'_{a,q}(x)=J'_q(x)-\frac{a^{1/q}J'_q(a^{1/q}x)}{Y_q(a^{1/q}x)}Y_q(x)$$
$$-\frac{Y'_q(x)Y_q(a^{1/q}x)-a^{1/q}Y'_q(a^{1/q}x)Y_q(x)}{[Y_q(a^{1/q}x)]^2}J_q(a^{1/q}x)$$
$$\to J'_q(x)$$

* Received by the editors October 29, 1974, and in final revised form February 2, 1976.

† Department of Mathematics and Statistics, Case Western Reserve University, Cleveland, Ohio 44106.

uniformly on intervals of the form $[\alpha, \beta]$, with $0<\alpha<\beta\leqq 1$. Finally, it is easily verified that

$$f_{a,q}(x)\sim(2^q\Gamma(q+1))^{-1}x^q(1-a^2)>\tfrac{1}{2}(2^q\Gamma(q+1))^{-1}x^q\sim\tfrac{1}{2}J_q(x)$$

whenever a is sufficiently small. Thus, there exist $\varepsilon, \delta>0$ such that $f_{a,q}(x)>0$ and $J_p(x)>0$ whenever $x\in(0,\varepsilon)$ and $a\in(0,\delta)$. This proves (iii).

LEMMA 2. *Let $z_n(a,q)$ denote the n-th positive zero of $f_{a,q}$ and $j_{q,n}$ denote the n-th positive zero of J_q. Then $z_n(a,q)\to j_{q,n}$ as $a\to 0^+$ whenever $q>0$.*

Proof. The proof proceeds by induction. Let ε and δ be as in Lemma 1 (iii). Assume that a is so small that $Y_q(a^{1/q}x)\neq 0$ for $x\in(0,j_{q,2})$. Then $J_q(\varepsilon)>0$ and $f_{a,q}(\varepsilon)>0$. Moreover, we also have that $z_1(a,q)>\varepsilon$ and $j_{q,1}>\varepsilon$. It is well known that J_q changes sign at each of its positive zeros. Let $0<\varepsilon_1<\frac{1}{2}(j_{q,2}-j_{q,1})$. Then $J_q(j_{q,1}+\varepsilon_1)<0$ so that for a sufficiently small $f_{a,q}(j_{q,1}+\varepsilon_1)<0$. Thus, we must have $z_1(a,q)\in(\varepsilon, j_{1,q}+\varepsilon_1)$ for all a sufficiently small, say $a<\delta_1$. Let z be any accumulation point of $\{z_1(a,q): 0<a<\delta_1\}$ and let $\{a_i\}$ be a sequence with $a_i\to 0$ and such that $z_1(a_i,q)\to z$. Then $0=f_{a_i,q}(z_1(a_i,q))\to J_q(z)$. The only zero of J_q in the interval $[\varepsilon, j_{1,q}+\varepsilon_1]$ is $j_{1,q}$. It follows that $z_1(a,q)\to j_{1,q}$. Now suppose that $z_n(a,q)\to j_{j,n}$. Let $0<\varepsilon<\frac{1}{2}\min\{j_{q,n}-j_{q,n-1}, j_{q,n+1}-j_{q,n}, j_{q,n+2}-j_{q,n+1}\}$. Then for a sufficiently small $z_n(a,q)\in(j_{q,n}-\varepsilon, j_{q,n}+\varepsilon)$ and $Y_q(a^{1/q}x)\neq 0$ for $x\in(0,j_{q,n+2})$. Since J_q changes sign at each of its zeros, $J_q(j_{q,n}+\varepsilon)$ and $J_q(j_{q,n+1}+\varepsilon)$ have opposite signs. For a sufficiently small $f_{a,q}(j_{q,n}+\varepsilon)$ and $f_{a,q}(j_{q,n+1}+\varepsilon)$ have opposite signs. Thus, $z_{n+1}(a,q)\in(j_{q,n}-\varepsilon, j_{q,n+1}+\varepsilon)$. Since $f_{a,q}\to J_q$ uniformly, any accumulation point of $\{z_{n+1}(a,q)\}$ is a zero of J_q. The only zeros of J_q in $[j_{q,n}-\varepsilon, j_{q,n+1}+\varepsilon]$ are $j_{q,n}$ and $j_{q,n+1}$. Suppose there is a sequence $\{a_i\}$ with $a_i\to 0$ and such that $z_{n+1}(a_i,q)\to j_{q,n}$. Then $f_{a_i,q}(z_{n+1}(a_i,q))=f_{a_i,q}(z_n(a_i,q))=0$. By the mean value theorem there is a $b_i\in(z_n(a_i,q), z_{n+1}(a_i,q))$ such that $f'_{a_i,q}(b_i)=0$. Notice that $b_i\to j_{q,n}$ since both $\{z_n(a_i,q)\}$ and $\{z_{n+1}(a_i,q)\}$ converge to $j_{q,n}$. Thus, $0=f'_{a_i,q}(b_i)\to J'_q(j_{q,n})$. This is impossible since J_q and J'_q do not vanish simultaneously. It follows that $z_{n+1}(a,q)\to j_{q,n}$. This completes the proof.

2. The inequalities. Let $R[p,y]$ denote the Rayleigh quotient

$$R[p,y]=\frac{\int_a^1(-(xy')'+x^{-1}y)y\,dx}{\int_a^1 x^{2p-1}y^2\,dx}.$$

It is well known that the eigenvalues $\{\lambda_n(p)\}$ of (1), (2) can be obtained from the Rayleigh quotient, [2, §§ 31 and 35]. Let V denote the linear space of all functions in $C^2((a,1))$ which satisfy the boundary conditions (2). Then

$$\lambda_1(p)=\min_{\substack{y\in V\\ y\neq 0}} R[p,y].$$

Let $y_1, y_2, \cdots, y_n$ be n functions in V, let A denote the subspace of V spanned by $y_1, y_2, \cdots, y_n$ and let $A^\perp$ denote the orthogonal complement of A relative to V. Then

$$\lambda_{n+1}(p)=\max_A \min_{\substack{y\in A^\perp\\ y\neq 0}} R[p,y]$$

where the maximum is taken over all sets of n functions in V.

Notice that $\lambda_n(p) = (pz_n(a, 1/p))^2$ where $z_n(a, 1/p)$ is the nth positive zero of $f_{a,1/p}$. Whenever $p \geqq q$ we have that $x^{2p-1} \leqq x^{2q-1}$ for $x \in [a, 1]$ and, hence, $R[p, y] \geqq R[q, y]$. It follows that $\lambda_n(p) \geqq \lambda_q(q)$ whenever $p \geqq q$ or, equivalently,

$$pz_n(a, 1/p) \geqq qz_n(a, 1/q)$$

whenever $p \geqq q$. Setting $t = q^{-1}$ and $s = p^{-1}$ we have that

$$\frac{z_n(a, s)}{s} \geqq \frac{z_n(a, t)}{t} \tag{4}$$

whenever $t \geqq s$. If we now let $a \to 0^+$ in (4) and apply Lemma 2 we obtain

$$\frac{t}{s} j_{s,n} \geqq j_{t,n} \tag{5}$$

whenever $t \geqq s > 0$.

THEOREM 3. *$(t/s)j_{s,n} > j_{t,n} > j_{s,n}$ whenever $t > s > 0$.*

Proof. The second inequality is well known. A proof using the Sturm comparison theorem may be found in [1]. An alternate proof is in [5, p. 508]. From (5) we have $s^{-1}j_{s,n} \geqq t^{-1}j_{t,n}$. Suppose that there are numbers t and s, $0 < s < t$, such that $s^{-1}j_{s,n} = t^{-1}j_{t,n}$. Then

$$\frac{j_{s,n}}{s} = \frac{j_{p,n}}{p}, \qquad s \leqq p \leqq t.$$

For simplicity set $s^{-1}j_{s,n} = k > 0$. Then $j_{p,n} = kp$ for $p \in [s, t]$ and we have

$$0 = J_p(j_{p,n}) = J_p(kp) \tag{6}$$

for $p \in [s, t]$. References to proofs of the following properties of J_p may be found on page 44 of [5]. $J_p(z)$ is an analytic function of z for all values of z ($z = 0$ possibly being excepted) and it is an analytic function of p for all values of p. Moreover, the series which defines $J_p(z)$ converges absolutely and uniformly in any closed domain of values of z [the origin not being a point of the domain when $R(p) < 0$], and in any bounded domain of values of p. It follows that $J_p(kp)$ is an analytic function of p on any compact interval not containing $p = 0$. Hence, $J_p(kp) \equiv 0$ on any compact interval not containing $p = 0$. It is known [5, p. 508] that $j_{t,1}$ is an increasing function of t so that $j_{t,n} \geqq j_{t,1} > j_{0,1} > 2.4$ whenever $t > 0$. If $p = 1/k$, then $J_{1/k}(1) = 0$. Thus, $j_{1/k,n} = 1$ for some n. This impossibility leads us to the conclusion $s^{-1}j_{s,n} \neq t^{-1}j_{t,n}$ whenever $s \neq t$. The desired inequality follows.

Theorem 3 states that $t^{-1}j_{t,n}$ is a strictly decreasing function of t. This is in contrast to the classical result that $j_{t,n}$ is a strictly increasing function of t, for $t > 0$ [5, p. 508]. Since $t^{-1}j_{t,n}$ is a strictly decreasing function of t it is natural to evaluate $\lim_{t\to\infty} t^{-1}j_{t,n}$. In [4] it is shown that

$$j_{t,n} = t + i_{1,n}6^{-1/3}t^{1/3} + \tfrac{3}{10}(i_{1,n})^2 6^{-2/3}t^{-1/3} + O(t^{-1}) \qquad (n = 1, 2, \cdots)$$

where $i_{1,n}$ is independent of t. It follows that $\lim_{t\to\infty} t^{-1}j_{t,n} = 1$ for $n = 1, 2, \cdots$.

COROLLARY 4. *For each n the function $j_{t,n}$ satisfies a uniform Lipschitz condition of order one in any interval $0 < a \leqq t < \infty$.*

Proof. Let $t, s \in [a, \infty)$ with $t \geqq s$. Then

$$0 \leqq j_{t,n} - j_{s,n} \leqq \frac{t}{s} j_{s,n} - j_{s,n} = \frac{t-s}{s} j_{s,n}$$

and $j_{s,n} < (s/a) j_{a,n}$ so that

$$0 \leqq j_{t,n} - j_{s,n} \leqq \frac{j_{a,n}}{a}(t-s).$$

The best known upper bound for $j_{t,1}$ is [4, p. 486]

$$j_{t,1} < (2(t+1)(t+3))^{1/2}.$$

Using the data in [3] it can be shown that $(11.5)^{-1} j_{11.5,1} < \sqrt{2}$. Hence

$$j_{t,1} < (11.5)^{-1} j_{11.5,1} < \sqrt{2t} < (2(t+1)(t+3))^{1/2}$$

whenever $t > 11.5$. This illustrates just one of the inequalities for $j_{t,n}$ which may be obtained from Theorem 3.

If (2) is replaced by

$$y'(a) = y'(0) = 0, \qquad 0 < a < 1, \tag{2'}$$

then the above procedure may be modified to show that $t^{-1} j'_{t,n}$ is a strictly decreasing function of t where $j'_{t,n}$ is the nth positive zero of J'_t. If we set

$$g_{a,q}(x) = J'_q(x) - \frac{J'_q(a^{1/q}x)}{Y'_q(a^{1/q}x)} Y'_q(x)$$

analogies to Lemmas 1 and 2 may be proved. In this case $z_n(a, q) \to j'_{q,n}$ as $a \to 0^+$. Line (4) remains valid with (2) replaced by (2') so that

$$\frac{t}{s} j'_{s,n} \geqq j'_{t,n}$$

whenever $t \geqq s > 0$. Proceeding in analogy to the proof of Theorem 3 we are able to verify

THEOREM 5. *$(t/s) j'_{s,n} > j'_{t,n}$ whenever $t > s > 0$.*

COROLLARY 6. *For each n the function $j'_{p,n}$ satisfies a uniform Lipschitz condition of order one in any interval $0 < a \leqq t < \infty$.*

Proof. The proof is analogous to that of Corollary 4.

COROLLARY 7. $\lim_{t \to \infty} t^{-1} j'_{t,n} = 1$.

Proof. On pages 485 and 487 of [5] it is shown that $t < j'_{t,1} < j_{t,1}$. Between any two consecutive zeros of J_t there is a zero of J'_t (the mean value theorem). A simple counting argument shows that $j'_{t,n} < j_{t,n}$ so that $1 < t^{-1} j'_{t,n} < t^{-1} j_{t,n} \to 1$ as $t \to \infty$.

Acknowledgment. The author would like to thank the referees for their comments and suggestions. In particular, I would like to thank them for bringing the paper, [4], by Tricomi to my attention and for suggesting the strengthening of the inequality in Theorem 3.

REFERENCES

[1] M. BÔCHER, *On certain methods of Sturm and their application to the roots of Bessel functions*, Bull. Amer. Math. Soc., 3 (1897), pp. 205–213.

[2] S. G. MIKHLIN, *Variational Methods of Mathematical Physics*, Macmillan, New York, 1964.

[3] *Royal Society Mathematical Tables, Vol. 7, Bessel Functions (Part III)*, University Press, Cambridge, England, 1960.

[4] F. TRICOMI, *Sulle funzioni di Bessel di ordine e argomenta pressoche uguali*, Atti. Accad. Sci. Torino Cl. Sci. Fis. Math. Nat., 83 (1949), pp. 3–20.

[5] G. N. WATSON, *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, Cambridge, England, 1958.

SIAM J. MATH. ANAL.
Vol. 8, No. 1, February 1977

MONOTONICITY AND CONVEXITY PROPERTIES OF ZEROS OF BESSEL FUNCTIONS*

J. T. LEWIS† AND M. E. MULDOON‡

Abstract. It is shown that $j_{\nu k}/\nu$ decreases as ν increases, $0<\nu<\infty$, and that $j_{\nu k}^2/\nu$ and $dj_{\nu k}^2/d\nu$ increase with ν for sufficiently large ν, where $j_{\nu k}$ is the kth positive zero of the Bessel function $J_\nu(x)$. In particular, $j_{\nu 1}^2/\nu$ and $dj_{\nu 1}^2/d\nu$ increase for $3\leqq\nu<\infty$. Some related results are proved for zeros of $J_\nu'(x)$, of cross-product Bessel functions and of modified Bessel functions of purely imaginary order.

1. Introduction. Putterman, Kac and Uhlenbeck [15] have proposed a purely quantum mechanical explanation for the origin of the vortex lines which are produced in superfluid helium when its container is rotated. It is based on the results of Blatt and Butler [2] (see also [9]) who showed that a rotating ideal Boson gas in a cylindrical bucket undergoes phase transitions. The total angular momentum Ω increases linearly with the angular velocity ω of the bucket in between successive critical values $\omega_1, \omega_2, \cdots$ where it jumps by an amount $N_0\hbar$, where N_0 is the number of condensed particles:

$$\Omega=\tfrac{1}{2}(N-N_0)mR^2\omega+N_0\hbar n, \qquad \omega_n<\omega<\omega_{n+1},$$

where N is the total number of particles, m the mass of a particle, and R the radius of the bucket. If the wave-function of a particle is assumed to satisfy Dirichlet boundary conditions on the walls of the bucket, the nth critical velocity is given by

$$\omega_n=(\tfrac{1}{2}\hbar mR^2)(j_{n1}^2-j_{n-1,1}^2),$$

where $j_{\nu k}$ or $j_{\nu,k}$ denotes the kth positive zero of the Bessel function $J_\nu(x)$. On physical grounds one expects that $\omega_1, \omega_2, \cdots$ is an increasing sequence and so it was conjectured that

$$j_{n+2,1}^2-2j_{n+1,1}^2+j_{n,1}^2>0, \qquad n=0,1,2,\cdots. \tag{1.1}$$

The results in this note were motivated by this. One of our principal results is that

$$dj_{\nu k}^2/d\nu \text{ increases with } \nu, \qquad 3\leqq\nu<\infty. \tag{1.2}$$

The conjecture (1.1) follows easily from this and the tables of $j_{\nu 1}$ for $\nu=1,2,3,4$ [19, pp. 748–750]. In order to prove (1.2) we show first that $j_{\nu 1}/\nu$ decreases with ν $(0<\nu<\infty)$ and that $j_{\nu 1}^2/\nu$ increases with ν $(3\leqq\nu<\infty)$.

Apart from the necessity to modify the ν-interval of validity we show that all of the above results hold, not only for $j_{\nu 1}$, but also for $j_{\nu k}$ $(k=2,3,\cdots)$.

* Received by the editors October 16, 1975.

† School of Theoretical Physics, Dublin Institute for Advanced Studies, Dublin 4, Republic of Ireland.

‡ Department of Mathematics, York University, Downsview, Ontario M3J 1P3, Canada. The work of the second author was supported by the Canada Council and by the National Research Council (Canada).

It is well known [3] that, for fixed k, $j_{\nu k}$ increases with ν, $0<\nu<\infty$. Watson [19, pp. 507–508] proves this by using Schläfli's formula

$$dj_{\nu k}/d\nu = 2\nu[j_{\nu k}J_{\nu+1}^2(j_{\nu k})]^{-1}\int_0^{j_{\nu k}} t^{-1}J_\nu^2(t)\,dt. \tag{1.3}$$

We show here that (1.3), together with the results on the monotonicity of $j_{\nu k}/\nu$, follow easily from a general result on the derivative of an eigenvalue with respect to a parameter, a special case of the Hellmann-Feynman theorem of quantum chemistry. However, in order to get (1.2) we need, in addition, another formula given by Watson [19, p. 508, (3)] which, in the special case needed here, says

$$dj_{\nu k}/d\nu = 2j_{\nu k}\int_0^\infty K_0(2j_{\nu k}\sinh t)\,e^{-2\nu t}\,dt, \tag{1.4}$$

where K_0 denotes the modified Bessel function of order zero.

There is a conjecture analogous to (1.1) for the zeros $j'_{\nu k}$ of fixed rank of $J'_\nu(x)$ which arises in the physical problem when Neumann boundary conditions replace the Dirichlet ones. We have not been able to prove this conjecture nor the analogue of (1.2) because the analogue of (1.4) [19, p. 510, (4)] is not amenable to any obvious simple treatment. However, we prove analogues of our *monotonicity* results for the zeros of $J'_\nu(x)$. Some of our monotonicity results are also applied to the zeros of cross-product Bessel functions (§ 5) and to zeros of modified Bessel functions of purely imaginary order (§ 6).

2. Some preliminary results. We consider a family of boundary value problems depending on a real parameter ν. They are given by a differential equation

$$-(d/dx)[p(x)\,dy/dx]+\nu^2q(x)y=\lambda\varphi(x)y, \tag{2.1}$$

and by boundary conditions

$$\lim_{x\to a+} p(x)y(x)y'(x)=p(b)y(b)y'(b). \tag{2.2}$$

It is supposed that $-\infty\leqq a<b<\infty$, that $p(x)>0$, and that $p'(x)$, $q(x)$ and $\varphi(x)$ are continuous for $a<x\leqq b$. We then have the following result.

LEMMA 2.1. *Suppose that, for each $\nu>0$, the boundary value problem* (2.1), (2.2) *has a discrete set of real eigenvalues. Let λ_ν be an eigenvalue of fixed rank and let there be a corresponding eigenfunction $\psi_\nu(x)$ satisfying the normalization condition*

$$\int_a^b \varphi(x)[\psi_\nu(x)]^2\,dx=1. \tag{2.3}$$

Suppose also that, for each ν,

$$\lim_{\mu\to\nu}\int_a^b q(x)\psi_\mu(x)\psi_\nu(x)\,dx=\int_a^b q(x)[\psi_\nu(x)]^2\,dx \tag{2.4}$$

and

$$\lim_{\mu\to\nu}\int_a^b \varphi(x)\psi_\mu(x)\psi_\nu(x)\,dx=\int_a^b \varphi(x)[\psi_\nu(x)]^2\,dx=1. \tag{2.5}$$

Suppose, also, that for each $\nu > 0$,

$$\lim_{x \to a+} p(x)\psi'_\nu(x)$$

exists, that $\psi_\nu(a^+) = 0$ *and either*

$$\psi_\nu(b) = 0 \quad \text{or} \quad \psi'_\nu(b) = 0. \tag{2.6}$$

Then

$$\frac{d\lambda_\nu}{d\nu} = 2\nu \int_a^b q(x)[\psi_\nu(x)]^2\, dx, \tag{2.7}$$

$$\frac{d}{d\nu}\frac{\lambda_\nu}{\nu} = \int_a^b [2q(x) - \nu^{-2}\lambda_\nu \varphi(x)][\psi_\nu(x)]^2\, dx \tag{2.8}$$

and

$$\frac{d}{d\nu}\frac{\lambda_\nu}{\nu^2} = -2\nu^{-3} \int_a^b p(x)[\psi'_\nu(x)]^2\, dx. \tag{2.9}$$

Proof. To prove (2.7), we multiply the equations

$$-(p\psi'_\mu)' + \mu^2 q\psi_\mu = \lambda_\mu \varphi\psi_\mu,$$
$$-(p\psi'_\nu)' + \nu^2 q\psi_\nu = \lambda_\nu \varphi\psi_\nu$$

by ψ_ν, ψ_μ respectively, subtract and integrate between a and b to get

$$(\lambda_\nu - \lambda_\mu) \int_a^b \varphi(x)\psi_\mu(x)\psi_\nu(x)\, dx = (\nu^2 - \mu^2) \int_a^b q(x)\psi_\mu(x)\psi_\nu(x)\, dx.$$

Dividing by $\nu - \mu$, letting $\mu \to \nu$, and using (2.3), (2.4) and (2.5) we get (2.7).

We have

$$\frac{d}{d\nu}\frac{\lambda_\nu}{\nu} = \nu^{-1}\frac{d\lambda_\nu}{d\nu} - \nu^{-2}\lambda_\nu = 2\int_a^b q(x)[\psi_\nu(x)]^2\, dx$$
$$-\lambda_\nu \nu^{-2} \int_a^b \varphi(x)[\psi_\nu(x)]^2\, dx$$

and this gives (2.8).

We also have

$$\frac{d}{d\nu}\frac{\lambda_\nu}{\nu^2} = \nu^{-2}\frac{d\lambda_\nu}{d\nu} - 2\nu^{-3}\lambda_\nu.$$

If we use (2.7) and the consequence

$$\lambda_\nu = -\int_a^b [p\psi'_\nu]'\psi_\nu\, dx + \nu^2 \int_a^b q(x)[\psi_\nu(x)]^2\, dx$$

of (2.1) and (2.3) we get

$$\frac{d}{d\nu}\frac{\lambda_\nu}{\nu^2} = 2\nu^{-3} \int_a^b \frac{d}{dx}\{p(x)\psi'_\nu(x)\}\psi_\nu(x)\, dx.$$

Integrating by parts and using the boundary conditions (2.2) we get (2.9).

This completes the proof of Lemma 2.1.

We remark that results of the type (2.7) are well known in the mathematical literature on perturbation theory and are given with various degrees of generality in [17, pp. 36–37], [18, pp. 233–235], [16, pp. 373–376] and [8]. In the present simple case we have found it more convenient to prove (2.7) directly than to deduce it from the more general results. In quantum chemistry results of this and of a more general kind are known as the "Hellmann-Feynman theorem". Feynman's work is in [6]; see [13] for an account of earlier versions of this result.

3. Application to zeros of Bessel functions. For $\nu > 0$, the boundary value problem

$$-\frac{d}{dx}\left(\frac{x\,dy}{dx}\right)+\nu^2x^{-1}y=\lambda xy, \tag{3.1}$$

$$y(0)=y(1)=0 \tag{3.2}$$

has eigenvalues $j_{\nu k}^2$, $k=1, 2, \cdots$, and corresponding eigenfunctions $J_\nu(j_{\nu k}x)$. The normalized eigenfunctions, in accordance with (2.3), are given by

$$\psi_\nu(x)=2^{1/2}[J_{\nu+1}(j_{\nu k})]^{-1}J_\nu(j_{\nu k}x), \tag{3.3}$$

where we have used a formula [19, p. 135, (11)] for the indefinite integral of $xJ_\nu^2(x)$. Now,

$$J_\nu(x)=O[x^\nu], \qquad x\to 0+,$$

so the hypotheses of Lemma 2.1 are satisfied. Hence (2.7) gives (1.3), a result which was proved in a different way by Watson [19, pp. 507–508].

Our principal results on Bessel function zeros are the following.

THEOREM 3.1. (I) *For each fixed k, $j_{\nu k}/\nu$ decreases as ν increases, $0<\nu<\infty$.*

(ii) For each fixed k, $j_{\nu k}^2/\nu$ increases with ν for sufficiently large ν; in particular, $j_{\nu 1}^2/\nu$ increases with ν for $3\leqq\nu\leqq\infty$.

(iii) *For each fixed k, $dj_{\nu k}^2/d\nu$ increases with ν, for sufficiently large ν; in particular, $dj_{\nu 1}^2/d\nu$ increases for $3\leqq\nu<\infty$.*

Proof. Part (i) is an immediate consequence of the result (2.9) of Lemma 2.1. From (2.8), we have

$$\frac{d}{d\nu}\frac{j_{\nu k}^2}{\nu}=\int_0^1 [2-j_{\nu k}^2\nu^{-2}x^2]\psi_\nu^2(x)x^{-1}\,dx,$$

where $\psi_\nu(x)$ is given by (3.3). The integrand here is positive provided that $j_{\nu k}^2/\nu^2<2$; this is so for sufficiently large ν since $j_{\nu k}/\nu\to 1$ as $\nu\to\infty$ [14, Exer. 6.4, p. 408]. Thus the first part of (ii) is proved.

However, we give an alternative proof which enables us to get a sharper result in the second part of (ii). We use (1.4) to get

$$\frac{d}{d\nu}\frac{j_{\nu k}^2}{\nu}=4\nu^{-1}j_{\nu k}^2\int_0^\infty K_0(2j_{\nu k}\sinh t)e^{-2\nu t}\,dt-\nu^{-2}j_{\nu k}^2.$$

This is positive provided $I(\nu, k) > 1$, where

$$I(\nu, k) = 4\nu \int_0^\infty K_0(2j_{\nu k} \sinh t)\, e^{-2\nu t}\, dt. \tag{3.4}$$

We have

$$I(\nu, k) = 2 \int_0^\infty K_0[2j_{\nu k} \sinh \{u/(2\nu)\}]\, e^{-u}\, du.$$

From part (i), $j_{\nu k}/\nu$ decreases with ν, $0 < \nu < \infty$; so, for each $u > 0$, does $2\nu \sinh \{u/(2\nu)\}$. Hence, so does their product $2j_{\nu k} \sinh \{u/(2\nu)\}$. Since K_0 is a decreasing function of its argument and since a decreasing function of a decreasing function is increasing, we find that $I(\nu, k)$ increases with ν, $0 < \nu < \infty$. Furthermore,

$$\lim_{\nu \to \infty} I(\nu, k) = 2 \int_0^\infty K_0(u)\, e^{-u}\, du = 2,$$

[19, p. 388]. Thus, for each k, there is a unique $\nu(k)$ such that $I(\nu(k), k) = 1$ and $I(\nu, k) > 1$ for $\nu > \nu(k)$. Thus we find again that $j_{\nu k}^2/\nu$ increases with ν for sufficiently large ν.

Next, we estimate $\nu(1)$. Using, [19, p. 444, (2)],

$$-(4/\pi) \int_0^\infty K_0(2z \sinh t)\, e^{-2\nu t}\, dt = J_\nu(z) \frac{\partial Y_\nu(z)}{\partial \nu} - Y_\nu(z) \frac{\partial J_\nu(z)}{\partial \nu},$$

the formula [14, p. 244, Exer. 5.6]

$$\left[\frac{\partial J_\nu(z)}{\partial \nu}\right]_{\nu = n} = \frac{\pi}{2} Y_n(z) + \frac{1}{2} n! \left(\frac{z}{2}\right)^{-n} \sum_{s=0}^{n-1} \left(\frac{z}{2}\right)^s [s!(n-s)]^{-1} J_s(z)$$

and the tables of values of $J_\nu(z)$ and $Y_\nu(z)$ in [19, pp. 666–733] and [1, Chap. 9], we find $I(2, 1) < 1$ and $I(3, 1) > 1$. Hence $2 < \nu(1) < 3$, $I(\nu, 1) > 1$ for $3 \leqq \nu < \infty$ and so $j_{\nu 1}^2/\nu$ increases for these values of ν. This completes the proof of part (ii).

To prove part (iii) we consider that, from (1.4),

$$\frac{d(j_{\nu k}^2)}{d\nu} = 4j_{\nu k}^2 \int_0^\infty K_0(2j_{\nu k} \sinh t)\, e^{-2\nu t}\, dt = \frac{j_{\nu k}^2}{\nu} I(\nu, k),$$

where $I(\nu, k)$ is given by (3.4). Since $I(\nu, k)$ increases with ν, $0 < \nu < \infty$, we see that $d(j_{\nu k}^2)/d\nu$ increases for at least those values of ν for which $j_{\nu k}^2/\nu$ increases, and part (iii) follows from part (ii).

Remark 1. In view of numerical evidence it seems likely that (iii) holds for all $\nu > 0$ at least in the case $k = 1$. Of course part (ii) does not hold for small positive ν since $j_{\nu k}^2/\nu \to \infty$ as $\nu \to 0+$.

Remark 2. We conjecture, but have not been able to prove that $dj_{\nu k}/d\nu$ decreases as ν increases, $0 < \nu < \infty$. The weaker result that $d(\log j_{\nu k})/d\nu$ decreases, $0 < \nu < \infty$, follows easily from (1.4). In fact, for each fixed k, we get $d^2(\log c_{\nu k})/d\nu^2 < 0$, $0 < \nu < \infty$, where $c_{\nu k}$ is a zero of fixed rank of any solution $AJ_\nu(x) + BY_\nu(x)$ of the Bessel equation. This was implicit in [11, p. 389] but it was stated there only for the first positive zero of $Y_\nu(x)$.

Remark 3. L. Lorch and P. Szego [12] have considered monotonicity with respect to order of the difference (and the higher differences) of consecutive zeros of Bessel functions. For fixed ν, higher monotonicity with respect to the rank k has also been considered; see [10] and the references contained therein.

Remark 4. The fact that $j_{\nu k}$ increases with ν for $\nu > 0$ was proved by Bôcher [3] using the Sturm comparison theorem. Watson's result, quoted as (1.4) above, has the advantage that it shows $j_{\nu k}$ to be increasing for $\nu > -1$ and, in fact, for all real ν, when the rank of a zero is suitably construed. We remark that in case $\nu > -1$ this follows also from the "fractional integral" formulation

$$x^{(\nu+\varepsilon)/2} J_{\nu+\varepsilon}(x^{1/2}) = [2^{\varepsilon}\Gamma(\varepsilon)]^{-1} \int_0^x (x-t)^{\varepsilon-1} t^{\nu/2} J_{\nu}(t^{1/2})\, dt,$$

$$\nu > -1, \quad \varepsilon > 0.$$

of Sonine's first integral [19, p. 373].

4. Zeros of derivatives of Bessel functions, If we consider the differential equation (3.1) again but with the boundary conditions

$$y(0) = y'(1) = 0,$$

we get the following results: the proof is similar to that of Theorem 3.1.

THEOREM 4.1. *Let $j'_{\nu k}$ denote the k-th positive zero of $J'_{\nu}(x)$. Then*

$$\frac{dj'_{\nu k}}{d\nu} = 2\nu[j'_{\nu k}\{(j'_{\nu k}/\nu)^2 - 1\}J^2_{\nu+1}(j'_{\nu k})]^{-1} \int_0^{j_{\nu k}} x^{-1} J^2_{\nu}(x)\, dx; \tag{4.1}$$

$j'_{\nu k}/\nu$ decreases to 1 as ν increases, $0 < \nu < \infty$, and $(j'_{\nu k})^2/\nu$ increases at least for those values of ν for which $j'_{\nu k} < 2^{1/2}\nu$.

The result (4.1) is not given by Watson [19]; its obvious consequence—that $j'_{\nu k}$ increases with ν is proved in a different way in [19, p. 510].

5. Zeros of cross-product Bessel functions. Again we consider the differential equation (3.1) but now with the boundary conditions

$$y(a) = y(1) = 0,$$

where $0 < a < 1$. The eigenvalues are the numbers $t^2_{\nu k}$ $(k = 1, 2, 3, \cdots)$ where $t_{\nu k}$ is the kth positive zero of the cross product

$$J_{\nu}(ax)Y_{\nu}(x) - Y_{\nu}(ax)J_{\nu}(x).$$

These zeros are real and simple [7, p. 82, Thm. X]. The corresponding unnormalized eigenvalues are

$$\psi_{\nu}(x) = J_{\nu}(t_{\nu k})Y_{\nu}(t_{\nu k}x) - Y_{\nu}(t_{\nu k})J_{\nu}(t_{\nu k}x).$$

Arguments like those in § 3 then give the following results.

THEOREM 5.1. *With the notation given above,*

$$\frac{dt^2_{\nu k}}{d\nu} = \frac{2\nu \int_a^1 [\psi_{\nu}(x)]^2 x^{-1}\, dx}{\int_a^1 x[\psi_{\nu}(x)]^2\, dx}, \qquad 0 < \nu < \infty. \tag{5.1}$$

Moreover, $t_{\nu k}/\nu$ decreases as ν increases, $0<\nu<\infty$, and $t_{\nu k}^2/\nu$ increases for those values of ν for which $t_{\nu k} \leqq 2^{1/2}\nu$.

Remark. In particular, it follows from (5.1) that each $t_{\nu k}$ increases with ν; this was proved in a more complicated ad hoc way by D. M. Willis [20].

6. Modified Bessel functions of purely imaginary order. We consider the modified Bessel function

$$K_{i\nu}(x)=\int_0^\infty e^{-x\cosh t}\cos \nu t\, dt$$

which satisfies the differential equation

$$x^2y''+xy'-(x^2-\nu^2)y=0.$$

$K_{i\nu}(x)$ vanishes at $+\infty$ and has infinitely many positive zeros whose only point of accumulation is $x=0$; see [5] for references and a good deal of further information about these zeros. We denote the zeros in decreasing order by $\varkappa_{\nu 1}, \varkappa_{\nu 2}, \cdots$. We find then that the boundary value problem

$$-(d/dx)(xdy/dx)-(\nu^2/x)y=-\lambda xy, \qquad y(-\infty)=y(-1)=0$$

has eigenvalues $\lambda=\varkappa_{\nu k}^2$ $(k=1,2,\cdots)$ and corresponding eigenfunctions $K_{i\nu}(-\varkappa_{\nu k}x)$. From the asymptotic behavior of $K_{i\nu}(x)$ [4, p. 87, (18)] it is seen that Lemma 2.1 is applicable and gives

$$\frac{d\varkappa_{\nu k}^2}{d\nu}=\frac{2\nu\int_1^\infty t^{-1}K_{i\nu}^2(\varkappa_{\nu k}t)\,dt}{\int_1^\infty tK_{i\nu}^2(\varkappa_{\nu k}t)\,dt},$$

$$\frac{d(\varkappa_{\nu k}^2/\nu^2)}{d\nu}=2\nu^{-3}\varkappa_{\nu k}^2\frac{\int_1^\infty t[K_{i\nu}'(\varkappa_{\nu k}t)]^2\,dt}{\int_1^\infty tK_{i\nu}^2(\varkappa_{\nu k}t)}.$$

Hence we have the following result.

THEOREM 6.1. *For each fixed k, $\varkappa_{\nu k}$ (the k-th positive zero in decreasing order of $K_{i\nu}(x)$) increases with ν, $0<\nu<\infty$, and $\varkappa_{\nu k}/\nu$ increases with ν, $0<\nu<\infty$.*

Acknowledgment. The authors are grateful to Professor Lee Lorch for his interest and for his useful comments.

REFERENCES

[1] M. ABRAMOVITZ AND I. A. STEGUN, eds., *Handbook of Mathematical Functions*, Applied Mathematics Series, vol. 55, National Bureau of Standards, Washington, 1964.

[2] J. M. BLATT AND S. T. BUTLER, *Superfluidity of an ideal Bose–Einstein gas*, Phys. Rev., 100 (1955), pp. 476–480.

[3] M. BÔCHER, *On certain methods of Sturm and their application to the roots of Bessel's functions*, Bull. Amer. Math. Soc., 3 (1897), pp. 205–213.

[4] A. ERDELYI, W. MAGNUS, F. OBERHETTINGER AND F. G. TRICOMI, *Higher Transcendental Functions*, vol. 2, McGraw-Hill, New York, 1954.

[5] E. M. FERREIRA AND J. SESMA, *Zeros of the modified Hankel function*, Numer. Math., 16 (1970), pp. 278–284.

[6] R. P. FEYNMAN, *Forces in molecules*, Phys. Rev., 56 (1939), pp. 340–343.

[7] A. GRAY, G. B. MATHEWS AND T. M. MACROBERT, *A Treatise on Bessel Functions and their Applications to Physics*, Macmillan, London, 1922.
[8] T. KATO, *Perturbation Theory for Linear Operators*, Springer-Verlag, New York, 1966.
[9] J. T. LEWIS AND J. V. PULÈ, *The free Boson gas in a rotating bucket*, Comm. Math. Phys., 45 (1975), pp. 115–131.
[10] L. LORCH, M. E. MULDOON AND P. SZEGO, *Higher monotonicity properties of certain Sturm-Liouville functions. IV*, Canad. J. Math., 24 (1972), pp. 349–365.
[11] ———, *Some monotonicity properties of Bessel functions*, this Journal, 4 (1973), pp. 385–392.
[12] L. LORCH AND P. SZEGO, *Monotonicity of the differences of zeros of Bessel functions as a function of order*, Proc. Amer. Math. Soc., 15 (1964), pp. 91–96.
[13] J. I. MUSHER, *Comment on some theorems of quantum chemistry*, Amer. J. Phys., 34 (1966), pp. 267–268.
[14] F. W. J. OLVER, *Asymptotics and Special Functions*, Academic Press, New York and London, 1974.
[15] S. J. PUTTERMAN, M. KAC AND G. E. UHLENBECK, *Possible origin of the quantized vortices in He. II*, Phys. Rev. Lett., 29 (1972), pp. 546–549.
[16] F. RIESZ AND B. SZ-NAGY, *Functional Analysis*, F. Ungar, New York, 1955.
[17] E. C. TITCHMARSH, *Some theorems on perturbation theory*, Proc. Roy. Soc. London, Ser. A., 200 (1950), pp. 34–46.
[18] ———, *Eigenfunction Expansions Associated with Second-order Differential Equations*, Part II, Clarendon Press, Oxford, 1958.
[19] G. N. WATSON, *A Treatise on the Theory of Bessel Functions*, 2nd ed., Cambridge University Press, Cambridge, England, 1944.
[20] D. M. WILLIS, *A property of the zeros of a cross-product of Bessel functions*, Proc. Cambridge Philos. Soc., 61 (1965), pp. 425–428.

SIAM J. MATH. ANAL.
Vol. 8, No. 2, April 1977

GROUP REPRESENTATION THEORY AND BRANCH POINTS OF NONLINEAR FUNCTIONAL EQUATIONS*

D. H. SATTINGER†

Abstract. In many physical applications the equations describing a system are invariant under some transformation group. When bifurcation problems arise in such a situation, the group invariance may lead to multiplicities of the branch point. The main goal of the present paper is to demonstrate in a precise way the application of group representation theory to bifurcation theory. Group representation theory is a linear one, while bifurcation theory deals with the branch points of nonlinear functional equations. Nevertheless, the theory of group representations applies to those nonlinear problems in a natural and elegant manner. The link between the two disciplines lies in the tensor character of the bifurcation equations on the one hand, and the theory of tensor products of group representations on the other.

1. Introduction. In a previous note [15] we pointed out, in a simple way, how the methods of group representation theory could be used to resolve a bifurcation problem where the dimension of the nullspace was greater than one. We wish to continue that direction of investigation further here, and to elaborate more fully on the possible applications of group theory to the analysis of branch points of nonlinear functional equations in a Banach space. Ostensibly, group representation theory is a linear one, while bifurcation theory deals with the branch points of nonlinear equations. Nevertheless, the theory of group representation applies, in a natural and elegant way, to these nonlinear problems in bifurcation theory. The link between these two disciplines lies in the tensor character of the bifurcation equations on the one hand, and the elegant theory of tensor products of group representations on the other hand.

Moreover, it is very often the case in physical applications, especially in the area of mechanics, that the given system of equations, even though nonlinear, is covariant with respect to a transformation group. For example, the Hamiltonian equations of celestial mechanics, or the partial differential equations governing the mechanics of a homogeneous continuum, are covariant under the Euclidean group. In § 2 we show that the Navier–Stokes equations governing the motion of a viscous incompressible fluid, are covariant under the Euclidean group $E(3)$. (See also [5].)

The original reason for our interest in the phenomenon of group invariance and its relation to bifurcation theory was our observation that the group invariance may in some cases account for the multiplicity of a branch point of a given nonlinear equation. This phenomenon is well known in quantum mechanics, where the invariance of the Hamiltonian under a symmetry group leads to a degeneracy of the energy levels. Group representation theory is an important tool in quantum mechanics for analyzing the splitting of these energy levels under symmetry-destroying perturbations (e.g., the Stark effect). It is our belief that the

* Received by the editors March 5, 1975, and in revised form October 9, 1975.

† School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455. This research was supported by the National Science Foundation under Grant GP-37512.

theory of group representations may also aid, and in a very elegant way, in the resolution of the bifurcation problem at a multiple eigenvalue.

Of particular interest is the phenomenon of pattern formation, or "symmetry breaking instabilities," analogous to the splitting of the energy levels in quantum mechanics (see [9], [10], [17]). For example, the appearance of convection cells in the Bénard problem or the buckling of a spherical shell under a uniform compression may be viewed as symmetry-breaking bifurcations. Prior to the onset of instability, the solutions are invariant under a continuous transformation group—$E(2)$ in the convection problem and $O(3)$ in the buckling problem. The bifurcating solution—that is, the nontrivial solution which appears at the onset of instability—is only invariant under a subgroup of the original continuous group. In the convection problem, the subgroup is necessarily a crystallographic group in the plane ("rolls" or "hexagons" seem to appear in experiments); in the buckling problem, the subgroup would be some point group of the second kind.

In such pattern formation problems, the branch point is generally of multiplicity greater than one. Mathematicians, not knowing how to treat bifurcation problems at a multiple eigenvalue, have sometimes gone to some lengths to rig the problem in such a way as to obtain a simple eigenvalue, for example, by casting the bifurcation problem in a symmetry class in which the branch point is simple. It is clear that this is not the mechanism employed by nature in selecting the symmetry patterns which in fact appear. It is our conjecture that the multiplicity of the branch point is an intrinsic aspect of the problem and that the methods of group representation theory may ultimately provide a satisfactory and elegant explanation of the mechanisms of pattern formation. In this regard, see the forthcoming paper [16].

The main goal of the present paper is to demonstrate in a precise way the application of group representation theory to bifurcation theory. To this end, we show in § 4 that the bifurcation equations are covariant under a symmetry group $\mathcal{G}$ if the original equation is covariant. We then demonstrate the tensor character of the bifurcation equations. In § 6 we summarize group representation theory and the fundamentals of tensor products of group representations. Finally, as an application of the methods, we consider a set of bifurcation equations of the form

$$\mathbf{v}+B(\mathbf{v},\mathbf{v})=0, \tag{1.1}$$

where $\mathbf{v}$ lies in a four-dimensional vector space V and B is a bilinear mapping from $V\times V$ to V which is invariant under D_3, the symmetry group of the equilateral triangle. Equations (1.1) may be written in the form

$$v_k+B_{ijk}v_iv_j=o, \qquad i,j,k=1,\cdots,4, \tag{1.2}$$

where $\mathbf{v}=v_i\mathbf{e}_i$, $\{\mathbf{e}_i\}$ forms a basis for V, and

$$B_{ijk}=\langle B(\mathbf{e}_i,\mathbf{e}_j),\mathbf{e}_k\rangle.$$

Since V is of dimension 4, there are a priori 64 possible numbers B_{ijk}. However, if B possesses a symmetry, for example, if

$$T(\sigma)B(\mathbf{v},\mathbf{w})=B(T(\sigma)\mathbf{v},T(\sigma)\mathbf{w}), \qquad \sigma\in D_3,$$

where $T(\sigma)$ is the representation of D_3 onto V given in (7.2), then the number of independent parameters is reduced to 11. It is further reduced to 7 if we assume, as we may, that B is symmetric. The equations (1.2) then take the simplified form exhibited in (8.1).

We wish to call the reader's attention to the work of Professor L. E. Scriven and his students of the Department of Chemical Engineering at the University of Minnesota. (See [1], [9], [10], [17]). Scriven and his students have already made considerable progress in applying group representation theory to nonlinear problems. The thesis of J. I. Gmitro [1] discusses the stability analysis of nonlinear initial value problems. The "formation of patterned states" is discussed, and some discussion is given of the interaction of the symmetry functions through the nonlinear terms.

The tensor character of nonlinear terms is discussed in the two papers by Othmer and Scriven [9], [10]. We hope that the present paper will clarify the mathematical structure in question, and that it will lead, in the future, to a clearer understanding of the role of group theory in nonlinear problems.

2. Invariance properties of partial differential equations. Let $\mathcal{G}$ be a group and V a linear vector space. We denote by $\mathcal{L}(V)$ the class of all linear transformations mapping V onto itself in a one-to-one fashion. A *representation* of $\mathcal{G}$ on V is a homomorphism from $\mathcal{G}$ onto the set $\mathcal{L}(V)$, considered as a group. Thus $g \to T_g$ in such a way that

$$T_{g_1} T_{g_2} = T_{g_1 g_2}.$$

For example, let W be the space of continuous functions on $\mathbb{R}^3$ and let $O(3)$ be the orthogonal group consisting of 3×3 matrices such that $OO^+ = I$. Then the family of operations defined by

$$(T_O f)(x) = f(O^{-1}x)$$

constitutes a group representation of $O(3)$ on W. We may also write this in the following way:

$$(T_O f)(x) = f(y_1, y_2, y_3),$$

where

$$y_i = O_{ji} x_j. \tag{2.1}$$

Now suppose V consists of the space of continuous vector fields on $\mathbb{R}^3$, denoted by u_i $(i = 1, 2, 3)$. A representation of $\mathcal{G}$ on V is

$$(T_O u)(x) = Ou(O^{-1}x)$$

or

$$(T_O u_i)(x) = O_{ij} u_j(y)$$

where y is given by (2.1).

Given a function $f = f(x_1, x_2, x_3)$, we denote by f_j the partial derivative of f with respect to the jth variable. Similarly the expression $u_{i,j}$ denotes the partial

derivative of the ith component of u_i with respect to the jth variable; and f_{ij} denotes the second partial derivative of f with respect to the ith and jth variables.

It is well known that the Laplacian is invariant under $O(3)$. By this is meant that $T_O\Delta = \Delta T_O$, or $T_O \Delta T_O^{-1} = \Delta$. Let us verify this using the above notation. Now in Cartesian coordinates,

$$\Delta f(x) = \delta_{ij} f_{ij}(x),$$

where the summation convention is understood. By the chain rule,

$$\frac{\partial}{\partial x^i} = \frac{\partial y_m}{\partial x_i}\frac{\partial}{\partial y_m} = O_{im}\frac{\partial}{\partial y_m}.$$

Therefore

$$\begin{aligned}(\Delta T_O f)(x) = \delta_{ij}\frac{\partial^2}{\partial x_i \partial x_j}(T_O f)(x) &= \delta_{ij}\frac{\partial^2}{\partial x_i \partial x_j} f(y_1, y_2, y_3) \\ &= \delta_{ij}\frac{\partial^2}{\partial y_m \partial y_n} f(y_1, y_2, y_3)\frac{\partial y_m}{\partial x_i}\frac{\partial y_n}{\partial x_j} \\ &= \delta_{ij}\frac{\partial^2}{\partial y_n \partial y_m} f(y_1, y_2, y_3) O_{im} O_{jn} = \delta_{mn} f_{nm}(y),\end{aligned}$$

while

$$(T_O \Delta f)(x) = (T_O \delta_{ij} f_{ij})(x) = \delta_{ij} f_{ij}(y).$$

If $\mathcal{F}$ is a nonlinear operator on a linear vector space V, we say that $\mathcal{F}$ is *covariant* under a group $\mathcal{G}$ if $\mathcal{F}$ commutes with all representations of $\mathcal{G}$ on $\mathcal{L}(V)$ – that is, if

$$\mathcal{F}(T_g u) = T_g \mathcal{F}(u)$$

for all u in V and any g in $\mathcal{G}$. A simple example of a nonlinear operation on W which is invariant under $O(3)$ is $\mathcal{F}(u) = f(u)$, where f is any continuous real-valued function. In fact,

$$T_O\{\mathcal{F}(u)(x)\} = T_O f(u(x)) = f(u(y)) = f(T_O u(x)) = \mathcal{F}(T_O u)(x).$$

Consequently, any nonlinear equation of the form $\Delta u + f(u) = 0$ is invariant under $O(3)$.

Now let us turn to a more interesting example—the Navier–Stokes equations governing the motion of a viscous incompressible fluid. These equations are

$$\Delta u_i - \frac{\partial p}{\partial x_i} - u_j \frac{\partial u_i}{\partial x_j} = 0, \qquad \frac{\partial u_j}{\partial x_i} = 0.$$

Let us show that these equations are covariant with respect to a representation of the Euclidean group. This has been done in [5]; we present the details here for the sake of completeness.

The appropriate vector space in the present case is $V \times W$, the Cartesian product of vector fields with scalar functions on $\mathbb{R}^3$. We denote an element of $V \times W$ by (u_i, p).

Associated with the Navier–Stokes equations is the nonlinear operation $\mathcal{F}$ given by

$$\mathcal{F}(u_i, p) = \left(\Delta u_i - \frac{\partial p}{\partial x_i} - u_j \frac{\partial u_i}{\partial x_j}, \frac{\partial u_\alpha}{\partial x_\alpha}\right).$$

(Summation of repeated indices is understood here.) The operation $\mathcal{F}$ transforms $V \times W$ into $V \times W$. The representation of $O(3)$ on $V \times W$ is

$$T_O(u_i, p)(x) = (O_{ij}u_j(y), p(y)).$$

Let us check the invariance of $\mathcal{F}$. We have

$$\frac{\partial}{\partial x_i} T_O p = \frac{\partial}{\partial x_i} p(y) = p_k(y) \frac{\partial y_k}{\partial x_i} = O_{ik} p_k(y),$$

$$T_O \frac{\partial p}{\partial x_i} = T_O p_i(x) = O_{ij} p_j(y),$$

hence

$$\text{grad } T_O p = T_O \text{ grad } p.$$

Now we check the nonlinear term $N(u_i) = u_j u_{i,j}$:

$$N(T_O u_i) = O_{jk}u_k(y) \frac{\partial}{\partial x_j} O_{il}u_l(y)$$

$$= O_{jk}u_k(y) O_{il}u_{l,m}(y) O_{jm} = O_{il}u_m(y)u_{l,m}(y),$$

whereas

$$T_O N(u_i) = T_O u_m(x) u_{i,m}(x) = O_{il}u_m(y)u_{l,m}(y).$$

Finally,

$$\frac{\partial}{\partial x_\alpha}(T_O u)_\alpha = \frac{\partial}{\partial x_\alpha} O_{\alpha\beta}u_\beta(y) = O_{\alpha\beta}u_{\beta,\gamma}(y)O_{\alpha\gamma} = u_{\beta,\beta}(y),$$

whereas

$$T_O u_{\alpha,\alpha}(x) = u_{\alpha,\alpha}(y).$$

We leave the remaining term, Δu_i, for the reader to check.

3. Branch points of nonlinear functional equations in a Banach space. We consider a functional equation of the form

$$L(\tau)u + N(\tau, u) = 0, \tag{3.1}$$

where $L(\tau)$ and $N(\tau, u)$ are analytic mappings of a complex Banach space $\mathbb{C} \times \mathcal{E}$ into $\mathcal{F}$. Here $\mathbb{C}$ denotes the complex numbers and $\mathcal{E}$, $\mathcal{F}$ are complex Banach spaces, with $\mathcal{E} \subseteq \mathcal{F}$. We also assume that L is an analytic function of τ with values in $\mathcal{L}(\mathcal{E}, \mathcal{F})$, the Banach space of bounded transformations from $\mathcal{E}$ to $\mathcal{F}$. In that case, L has a power series expansion

$$L(\tau) = L_0 + \tau L_1 + \cdots$$

which converges in $\mathscr{L}(\mathscr{E}, \mathscr{F})$. We also assume that N has a power series expansion in u and τ, and that all terms in N are at least quadratic in u. We write

$$N(\tau, u) = \sum_{k=2}^{\infty} \sum_{\substack{i+j=k \\ i \geqq 2}} N_{ij}(u)\tau^j,$$

where $N_{ij}(u)$ is a homogeneous operator of degree i—that is,

$$N_{ij}(\sigma u) = \sigma^i N_{ij}(u).$$

The basic problem of bifurcation theory consists of constructing all solutions of (3.1) in a neighborhood of $u = 0$, $\tau = 0$ when L_0 has a nontrivial nullspace. We shall assume, in what follows, that $u = 0$ is a solution for all values of τ and that there are no solutions of (3.1) of the form $(u, 0)$ for sufficiently small u.

Let L_0 have a nullspace of dimension n, spanned by basis vectors $\varphi_1, \cdots, \varphi_n$. We suppose the range of L_0 to be a closed subspace of $\mathscr{F}$ with co-dimension n. Since $L_0 : \mathscr{E} \to \mathscr{F}$, the adjoint operator L_0^* maps $\mathscr{F}^*$ into $\mathscr{E}^*$. We denote the n adjoint null functions by $\varphi_1^*, \cdots, \varphi_n^*$ and assume them to be chosen so that $\langle \varphi_i, \varphi_j^* \rangle = \delta_{ij}$. The range of L_0 is characterized by

$$\mathscr{F}_0 = \{f : f \in \mathscr{F}, \langle f, \varphi_i^* \rangle = 0, i = 1, \cdots, n\}.$$

Since $\mathscr{E} \subseteq \mathscr{F}$, $\mathscr{F}^* \subseteq \mathscr{E}^*$ and $\varphi_i^* \in \mathscr{E}^*$. Denote by $\mathscr{E}_0$ the subspace

$$\mathscr{E}_0 = \{u : u \in \mathscr{E}, \langle u, \varphi_i^* \rangle = 0, i = 1, \cdots, n\}.$$

Then L_0 is an isomorphism from $\mathscr{E}_0$ to $\mathscr{F}_0$. Let P be the projection in $\mathscr{F}$ given by

$$Pu = \sum_{j=1}^{n} \langle u, \varphi_j^* \rangle \varphi_j$$

and let Q be the projection onto $\mathscr{F}_0$ given by $I - P$. We can restrict P to $\mathscr{E}$, and then it is the projection onto the nullspace of L_0, while Q restricted to $\mathscr{E}$ is the projection onto $\mathscr{E}_0$.

We now summarize the Lyapounov–Schmidt procedure for reducing (3.1) to a system of n algebraic equations for n unknowns. Specifically, we shall reduce the problem to an equation of the simple form

$$\xi + M_m(\xi) = 0, \tag{3.2}$$

where $\xi \in \mathbb{C}_n$ and M_m is a homogeneous operator of degree m. If the original equation is covariant under a symmetry group, then this covariance is inherited by the m-linear mapping M_m in (3.2).

We write the solution u of (3.1) in the form

$$u = \alpha \cdot \varphi + \psi(\alpha, \tau), \tag{3.3}$$

where $\alpha = (\alpha_1, \cdots, \alpha_n)$,

$$\alpha \cdot \varphi = \alpha_1 \varphi_1 + \alpha_2 \varphi_2 + \cdots + \alpha_n \varphi_n,$$

and $\psi(\alpha, \tau) \in \mathscr{E}_0$. Substituting (3.3) into (3.1) and applying the projection Q, we get

$$QL(\tau)\psi + Q(\tau L_1 + \tau^2 L_2 + \cdots)\alpha \cdot \varphi + QN(\tau, \alpha \cdot \varphi + \psi) = 0. \tag{3.4}$$

Equation (3.4) has a solution ψ, analytic in τ and α, with values in $\mathscr{E}_0$. This is an immediate consequence of the implicit function theorem. In fact, setting

$$K(\tau, \alpha, \psi) = QL(\tau)\psi + Q(\tau L_1 + \tau^2 L_2 + \cdots)\alpha \cdot \varphi + QN(\tau, \alpha \cdot \varphi + \psi)$$

we see that K is an analytic mapping of $\mathbb{C}^{n+1} \times \mathscr{E}_0$ into $\mathscr{F}_0$. The Fréchet derivative of K with respect to ψ at $\tau = \alpha = \psi = 0$ is

$$K'_\psi(0, 0, 0) = QL_0,$$

which is a linear isomorphism between $\mathscr{E}_0$ and $\mathscr{F}_0$.

In order that (3.3) be a solution of (3.1), it is necessary and sufficient that

$$Q(L(\tau)\psi + L(\tau)\alpha \cdot \varphi + N(\tau, \alpha\varphi + \psi)) = L(\tau)(\alpha \cdot \varphi + \psi) + N(\tau, \alpha \cdot \varphi + \psi),$$

i.e.,

$$P(L(\tau)\psi + L(\tau)\alpha \cdot \varphi + N(\tau, \alpha \cdot \varphi + \psi)) = 0,$$

hence

$$\langle L(\tau)\psi + L(\tau)\alpha \cdot \varphi + N(\tau, \alpha \cdot \varphi + \psi), \varphi_j^* \rangle = 0, \qquad j = 1, \cdots, n. \tag{3.5}$$

Equations (3.5) are called the *bifurcation equations*. They consist of n equations in the $n+1$ unknowns $\tau, \alpha_1, \cdots, \alpha_n$. Let us write them concisely as $F(\alpha, \tau) = 0$, where

$$F(\alpha, \tau) = \sum_{k=1}^{\infty} \sum_{i+j=k} F_{ij}(\alpha)\tau^j. \tag{3.6}$$

A close inspection of (3.5) reveals that the lowest order term is

$$\langle L_0\alpha \cdot \varphi, \varphi_j^* \rangle = \sum_{i=1}^{n} \langle L_0\varphi_i, \varphi_j^* \rangle \alpha_i.$$

The second order terms are given by

$$\tau\langle L_1\alpha \cdot \varphi, \varphi_j^* \rangle = \tau \sum_{i=1}^{n} \langle L_1\varphi_i, \varphi_j^* \rangle \alpha_i$$

and possibly by

$$\langle N(0, \alpha \cdot \varphi), \varphi_j^* \rangle, \tag{3.7}$$

provided these are of second order and do not vanish. We assume here, as in [7], that the matrix $\langle L_1\varphi_i, \varphi_j^* \rangle$ is nonsingular.

We now proceed to reduce (3.5) further by a device employed by Graves [2] and Sather [12], namely, by constructing the Newton diagram for the bifurcation equations (3.5).

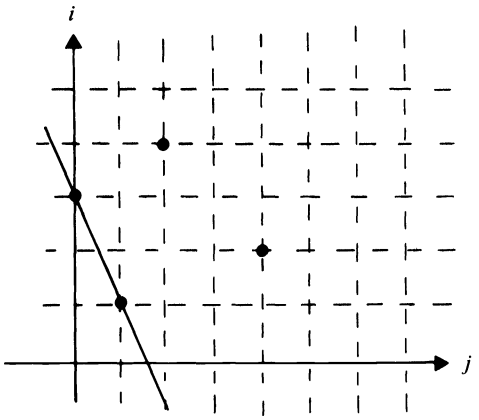


FIG. 3.1

Referring to the expansion (3.6) for $F(\alpha, \tau)$, we plot every point (i, j) on the lattice of nonnegative integer points in the first quadrant of the x-y plane for which F_{ij} does not vanish. Because the term $\tau \sum_{i=1}^{n} \langle L_1\varphi_i, \varphi_j^* \rangle \alpha_i$ appears as the lowest order term, the point (1, 1) appears on the Newton diagram. Since we have assumed that $(0, \tau)$ is always a solution of (3.1), we have $F(0, \tau) \equiv 0$; this condition becomes

$$F(0, \tau) = \sum_{k=1}^{\infty} F_{0k}(\alpha)\tau^k \equiv 0.$$

Therefore all terms of the form F_{0k} vanish and there are no points on the j-axis of the Newton diagram. On the other hand, since there are no solutions of (3.1) of the form $(u, 0)$ for sufficiently small u, $F(\alpha, 0)$ can not vanish identically, and

$$F(\alpha, 0) = \sum_{k=2}^{\infty} F_{k0}(\alpha) \neq 0.$$

Let m be the smallest integer for which $F_{m0}(\alpha) \neq 0$. There exists a point $(m, 0)$ on the i-axis of the Newton diagram.

Let l_0 be the line in the Newton diagram passing through $(m, 0)$ and (1, 1):

$$i + (m-1)j = m.$$

All other points on the Newton diagram lie above and to the right of l_0. Now let l_k be the line

$$i + (m-1)j = m + k, \qquad k = 0, 1, 2, \cdots.$$

We may rewrite F in the form

$$F(\alpha, \tau) = \sum_{k=0} \sum_{i+(m-1)j=m+k} F_{ij}(\alpha)\tau^j.$$

A new scale parameter ε is now introduced by setting

$$\alpha = \varepsilon\xi, \qquad \tau = \varepsilon^{m-1},$$

where ε is chosen so that ε is real and positive when τ is. Then the bifurcation equations become

$$F(\varepsilon\xi, \varepsilon^{m-1}) = \varepsilon^m(A\xi + F_{m0}(\xi)) + \sum_{k=1}^{\infty} \varepsilon^{m+k}\Big(\sum_{i+(m-1)j=m+k} M_{ij}(\xi)\Big) = 0,$$

where A is the matrix $A_{ij} = \langle L_1\varphi_i, \varphi_j^*\rangle$ and $A\xi = A_{ij}\xi_i$. Note that these equations are analytic in ε. If one is interested only in real solutions of (3.5), then only those solutions are considered for which ξ, ε, and ε^{m-1} are real.

Dividing by ε^m in the above expression and letting $\varepsilon \to 0$, we get

$$A\xi + F_{m0}(\xi) = 0. \tag{3.8}$$

Here $F_{m0}(\xi)$ is a homogeneous operator of degree m which maps $\mathbb{C}^n$ into itself.

In the next section we shall show that if the original nonlinear equation (3.1) is covariant under a symmetry group, then the reduced bifurcation equations (3.8) are covariant under the same symmetry group. The relationship of the reduced bifurcation equations (3.8) to the full set of equations (3.5) requires further comment. If the Jacobian of equations (3.8) at a solution is nonsingular, then a solution of (3.8) can be extended to a solution of (3.5) by the implicit function theorem (hence to a solution of the full equations (3.1)). Unfortunately, this is not always the case in applications (see [16]), and then a deeper analysis is required. The reader is referred to the papers by Kirchgässner [4] and Sather [12], [13] for further investigations of such matters.

4. Group invariance of the bifurcation equations. Let $\mathcal{G}$ be a symmetry group and let T_g be a representation of $\mathcal{G}$ on $\mathcal{F}$ which leaves equation (3.1) invariant. That is,

$$T_g L(\tau) = L(\tau) T_g \tag{4.1}$$

and

$$T_g N(\tau, u) = N(\tau, T_g u) \tag{4.2}$$

for all u in $\mathcal{E}$. First, it is clear from (4.1) that T_g leaves the subspace $\eta_0 = \{u : L_0 u = 0\}$ invariant. We also assume that $T_g\mathcal{E} \subset \mathcal{E}$; since $\mathcal{E} \subset \mathcal{F}$, T_g defines a group representation on $\mathcal{E}$ as well. Since T_g maps $\mathcal{E}$ to $\mathcal{E}$ and $\mathcal{F}$ to $\mathcal{F}$, its adjoint operator T_g^* maps $\mathcal{E}^*$ to $\mathcal{E}^*$ and $\mathcal{F}^*$ to $\mathcal{F}^*$; and, furthermore,

$$T_g^* L^* = L^* T_g^*.$$

T^* is an anti-representation on the dual space $\mathcal{E}^*$; that is,

$$T^*(g_1 g_2) = T^*(g_2) T^*(g_1).$$

However, the mapping $g \to T^*(g^{-1})$ defines a representation of $\mathcal{G}$ on the dual space.

When T_g is restricted to η_0, it is a finite-dimensional representation of $\mathcal{G}$ and therefore has a matrix representation. Thus

$$T(g)\varphi_i = T_{ji}(g)\varphi_j;$$

and, since $\langle \varphi_i, \varphi_j^* \rangle = \delta_{ij}$,

$$\langle T(g)\varphi_i, \varphi_j^* \rangle = T_{ji}(g). \tag{4.3}$$

Similarly

$$T^*(g)\varphi_i^* = T_{ji}^*(g)\varphi_j^*,$$

and

$$\begin{aligned} T_{ji}(g) = \langle T(g)\varphi_i, \varphi_j^* \rangle &= \langle \varphi_i, T_{kj}^*(g)\varphi_k^* \rangle \\ &= T_{kj}^*(g)\delta_{ik} = T_{ij}^*(g). \end{aligned}$$

Consequently, it is easily seen that $\mathscr{E}_0$ and $\mathscr{F}_0$ are invariant under $T(g)$, and that the projections P and Q introduced in § 3 commute with $T(g)$. In fact,

$$\begin{aligned} T(g)Pu &= T(g)\langle u, \varphi_i^* \rangle \varphi_i = T_{ji}(g)\langle u, \varphi_i^* \rangle \varphi_j, \\ PT(g)u &= \langle T(g)u, \varphi_i^* \rangle \varphi_i = \langle u, T^*(g)\varphi_i^* \rangle \varphi_i \\ &= \langle u, T_{ji}^*(g)\varphi_j^* \rangle \varphi_i = \overline{T_{ji}^*(g)} \langle u, \varphi_j^* \rangle \varphi_i \\ &= T_{ij}(g)\langle u, \varphi_j^* \rangle \varphi_i = T(g)Pu. \end{aligned}$$

Since $Q = I - P$, $T(g)$ commutes with Q as well.

It will be convenient to define an inner product on the kernel η_0 and to identify η_0^* (the kernel of the adjoint operator) with η_0, as follows. If φ_i, φ_j^* denote dual bases for η_0 and η_0^*, we identify η_0 with η_0^* by the correspondence

$$y_i\varphi_i \leftrightarrow \bar{y}_i\varphi_i^*.$$

An inner product on η_0 is given by

$$(x, y) = \langle x_i\varphi_i, y_j\varphi_j^* \rangle = x_i\bar{y}_i.$$

We can further assume that T is a unitary representation on this inner product space—that is, that $(Tx, Ty) = (x, y)$. If T is not unitary with respect to this inner product, then a new inner product can always be introduced relative to which T is unitary [8, p. 67]. In applications, however, this additional difficulty may not arise. The original problem may possess a Hilbert space structure and T in (4.1) may a priori be a unitary representation. The kernel η_0 then has a natural inner product inherited from the original Hilbert space and the restriction of T to η_0 will then automatically be unitary (see [16]).

We are now ready to prove

THEOREM 4.1. *Let equation* (3.1) *be covariant under a symmetry group* $\mathscr{G}$. *Let* $T_{ij}(g)$ *be the representation of* $\mathscr{G}$ *on the* n*-dimensional nullspace* η_0 *given by* (4.3). *Then the bifurcation equations* (3.5) *are covariant under* $\mathscr{G}$. *That is,*

$$T(g)F(\alpha, \tau) = F(T(g)\alpha, \tau). \tag{4.4}$$

In particular, each term in (3.6) *is covariant under* T:

$$T(g)F_{ij}(\alpha) = F_{ij}(T(g)\alpha).$$

This equation is understood in the following sense. If $\alpha \in \mathbb{C}^n$, $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ and $u = \alpha_i \varphi_i$, then

$$T(g)u = \alpha_i T(g)\varphi_i = \alpha_i T_{ji}(g)\varphi_j = (T_{ji}(g)\alpha_i)\varphi_j,$$

so

$$(T(g)\alpha_i) = T_{ij}(g)\alpha_j.$$

Similarly, $(T(g)F)_i = T_{ij}F_j$ if F is a mapping with values in $\mathbb{C}^n$, $F = (F_1, \cdots, F_n)$.

Proof. We first show that

$$T(g)\psi(\alpha, \tau) = \psi(T(g)\alpha, \tau),$$

where $\psi = Qu$. We use the fact that ψ is a solution of equation (3.4), and that this equation has a unique solution for small α and τ. We have

$$\begin{aligned} 0 &= T(g)L(\tau)\psi + T(g)QL(\tau)\alpha \cdot \varphi + T(g)QN(\tau, \alpha \cdot \varphi + \psi) \\ &= L(\tau)T(g)\psi + QL(\tau)\alpha \cdot T(g)\varphi + QN(\tau, \alpha \cdot T(g)\varphi + T(g)\psi) \\ &= L(\tau)(T(g)\psi) + QL(\tau)(T(g)\alpha) \cdot \varphi + QN(\tau, (T(g)\alpha) \cdot \varphi + (T(g)\psi)). \end{aligned}$$

Hence $(T(g)\psi)(\alpha, \tau)$ is a solution of (3.4) with α replaced by $T(g)\alpha$. Since the solution is unique, however,

$$(T(g)\psi)(\alpha, \tau) = \psi(T(g)\alpha, \tau).$$

Turning now to the bifurcation equations, we have

$$F_j(\alpha, \tau) = \langle L(\tau)\{\alpha \cdot \varphi + \psi\} + N(\tau, \alpha \cdot \varphi + \psi), \varphi_j^* \rangle,$$

so

$$\begin{aligned} (T(g)F)_i = T_{ij}F_j(\alpha, \tau) &= \langle L(\tau)\psi + L(\tau)\alpha \cdot \varphi + N(\tau, \alpha \cdot \varphi + \psi), \overline{T_{ji}(g)}\varphi_j^* \rangle \\ &= \langle L(\tau)\psi + L(\tau)\alpha \cdot \varphi + N(\tau, \alpha \cdot \varphi + \psi), T^*(g)\varphi_i^* \rangle \\ &= \langle T(g)L(\tau)\{\alpha \cdot \varphi + \psi\} + T(g)N(\tau, \alpha \cdot \varphi + \psi), \varphi_i^* \rangle \\ &= \langle L(\tau)\{T(g)\alpha \cdot \varphi + T(g)\psi\} \\ &\quad + N(\tau, (T(g)\alpha) \cdot \varphi + T(g)\psi), \varphi_i^* \rangle \\ &= F_i(T(g)\alpha, \tau). \end{aligned}$$

Finally, since $F(\alpha, \tau)$ is covariant under $T(g)$ independently of τ, each term $F_{ij}(\alpha)$ in (3.6) is covariant under $T(g)$. In particular, equations (3.8) are covariant under the symmetry group $\mathcal{G}$.

5. Tensor character of the bifurcation equations. In § 3 we saw how to reduce a bifurcation problem to one of solving an equation of the form

$$A\xi + M(\xi) = 0, \tag{5.1}$$

where M is a homogeneous operator of degree k mapping a finite-dimensional complex vector space V into itself, and A is a linear transformation on V. In this section we discuss the tensor character of (5.1)—in particular, the tensor character of the transformation M.

First, M may be derived from a k linear operator B, where B is given by

$$B(\mathbf{v}_1, \cdots, \mathbf{v}_k) = \frac{\partial^k}{\partial t_1 \partial t_2 \cdots \partial t_k} M(t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_k\mathbf{v}_k)\Big|_{t_i=0}$$

(see [14, p. 67]).

If M is covariant under a group $\mathcal{G}$, then B possesses the covariance property

$$T(g)B(\mathbf{v}_1, \cdots, \mathbf{v}_k) = B(T(g)\mathbf{v}_1, \cdots, T(g)\mathbf{v}_k). \tag{5.2}$$

Furthermore, B is linear in each variable; that is, B is a k-linear operator on $V \times V \times \cdots \times V$ into V. We wish to reinterpret B as a linear transformation from $V \otimes \cdots \otimes V$ into V—that is, as a linear transformation from the tensor product space $V^{\otimes k}$ into V.

Let V, W be finite-dimensional vector spaces with bases $\mathbf{v}_i$, $\mathbf{w}_j$, $i = 1, \cdots, n$, $j = 1, \cdots, m$, respectively. We consider the formal tensor product space $V \otimes W$ as the nm-dimensional vector space with basis elements $\{v_i \otimes w_j\}$, $1 \leqq i \leqq n$, $1 \leqq j \leqq m$. Thus any $\mathbf{x}$ in $V \otimes W$ can be written uniquely as

$$\mathbf{x} = \sum_{i,j} \alpha_{ij} \mathbf{v}_i \otimes \mathbf{w}_j.$$

The tensor product operation is to obey the following rules:

$$\alpha(\mathbf{v} \otimes \mathbf{w}) = \alpha\mathbf{v} \otimes \mathbf{w} = \mathbf{v} \otimes \alpha\mathbf{w},$$

$$(\mathbf{u} + \mathbf{v}) \otimes \mathbf{w} = \mathbf{u} \otimes \mathbf{w} + \mathbf{v} \otimes \mathbf{w},$$

$$\mathbf{v} \otimes (\mathbf{u} + \mathbf{w}) = \mathbf{v} \otimes \mathbf{u} + \mathbf{v} \otimes \mathbf{w}.$$

A tensor B of type (0, 2) on V is an element of $(V \otimes V)^*$—that is, it is a linear functional on $V \otimes V$. Let $\varphi_1, \cdots, \varphi_n$ be a basis for V and $\varphi_1^*, \cdots, \varphi_n^*$ be the dual basis for V^*. Consider the functionals on $(V \otimes V)^*$ given by

$$(\varphi_i^* \otimes \varphi_j^*)(\mathbf{v} \otimes \mathbf{w}) = \varphi_i(\mathbf{v})\varphi_j(\mathbf{w}) \tag{5.3}$$

We extend these operations to all of $V \otimes W$ by linearity. That is,

$$(\varphi_i^* \otimes \varphi_j^*)\Big(\sum_{r,s} \alpha_{rs}\varphi_r \otimes \varphi_s\Big) = \sum_{r,s} \alpha_{rs}(\varphi_i^* \otimes \varphi_j^*)(\varphi_r \otimes \varphi_s)$$

$$= \sum_{r,s} \alpha_{rs}\delta_{ir}\delta_{js} = \alpha_{ij}.$$

Then $\varphi_i^* \otimes \varphi_j^*$ is an element of $(V \otimes V)^*$. A complete basis for $(V \otimes V)^*$ is $\{\varphi_i^* \otimes \varphi_j^*\}$, $1 \leqq i, j \leqq n$, and so

$$(V \otimes V)^* \cong V^* \otimes V^*.$$

Note that the right side of (5.3) is bilinear in $\mathbf{v}$ and $\mathbf{w}$. Thus the functional $B_{ij}(\mathbf{v}, \mathbf{w}) = (\varphi_i^* \otimes \varphi_j^*)(\mathbf{v} \otimes \mathbf{w})$ is bilinear in $\mathbf{v}$ and $\mathbf{w}$ and linear in $\mathbf{v} \otimes \mathbf{w}$.

A general tensor B of type (0, 2) on V can be written in the form

$$B = \sum_{i,j} B_{ij}\varphi_i^* \otimes \varphi_j^*.$$

Since each term in the sum is a bilinear functional on $V\times V$, B itself is a bilinear functional on $V\times V$. Conversely, suppose $F(\mathbf{u}, \mathbf{v})$ is any bilinear functional on $V\times V$. Define a tensor B in $V^*\times V^*$ by

$$B(\varphi_i\otimes\varphi_j) = F(\varphi_i, \varphi_j)$$

and by extending B to the entire space $(V\otimes V)$ by linearity. Then

$$\begin{aligned} F(\mathbf{u}, \mathbf{v}) &= F\Big(\sum_i u_i\varphi_i, \sum_j v_j\varphi_j\Big) = \sum_{i,j} u_i v_j F(\varphi_i, \varphi_j) \\ &= \sum_{i,j} u_i v_j B(\varphi_i\otimes\varphi_j) = B\Big(\sum_{i,j} u_i\varphi_i\otimes v_j\varphi_j\Big) \\ &= B\Big(\sum_i u_i\varphi_i\otimes\sum_j v_j\varphi_j\Big) = B(\mathbf{u}\otimes\mathbf{v}). \end{aligned}$$

Thus every bilinear functional may be represented as a tensor in $(V\otimes V)^*$—that is, as a linear functional on $V\otimes V$.

The above idea carries over immediately to k-linear functionals in an obvious way: every k-linear functional is a linear functional on $V^{\otimes k}$.

Now suppose B is a k-linear operator from $V\times V\times\cdots\times V$ to V. For convenience, take $k = 2$. Then

$$B(\mathbf{u}, \mathbf{v}) = B_i(\mathbf{u}, \mathbf{v})\varphi_i$$

where the bilinear functionals B_i are given by

$$B_i(\mathbf{u}, \mathbf{v}) = \langle B(\mathbf{u}, \mathbf{v}), \varphi_i^*\rangle.$$

As we have seen above, $(V\otimes V)^*\cong V^*\otimes V^*$, so the functionals B_i are elements of $V^*\otimes V^*$ and may be written in the form

$$B_i = B_{jki}\varphi_i^*\otimes\varphi_k^*.$$

The mapping B can then be represented in the form

$$B = B_{jki}\varphi_j^*\otimes\varphi_k^*\otimes\varphi_i,$$

with it understood that

$$B(\mathbf{u}, \mathbf{v}) = B_{jki}\langle\mathbf{u}, \varphi_j^*\rangle\langle\mathbf{v}, \varphi_k^*\rangle\varphi_i.$$

If we further identify V^* with V as in § 4, then B may be written

$$B = B_{jki}\varphi_j\otimes\varphi_k\otimes\varphi_i,$$

where the components B_{jki} are given by

$$B_{jki} = (B(\varphi_j, \varphi_k), \varphi_i);$$

here $(\cdot, \cdot)$ denotes the inner product on V. Thus

$$B(\mathbf{u}, \mathbf{v}) = B_{jki}(\mathbf{u}, \varphi_j)(\mathbf{v}, \varphi_k)\varphi_i.$$

The invariance property carried by B is then expressed as

$$B_{ijk}(\varphi_j\otimes\varphi_k)T(g)\varphi_i = B_{ijk}(T(g)(\varphi_j\otimes\varphi_k))\varphi_i.$$

The bifurcation equations

$$\mathbf{v}+B(\mathbf{v},\mathbf{v})=0$$

may be written as

$$v_i+B_{ijk}v_jv_k=0,$$

where $v_i=(\mathbf{v},\varphi_i)$. Corresponding to the term $B_{ijk}v_jv_k$, therefore, is the tensor

$$B=B_{ijk}(\varphi_j\otimes\varphi_k)\varphi_i.$$

6. Group representations and their tensor products. Let T_g be a representation of the group $\mathcal{G}$ on a finite-dimensional vector space V. If $\varphi_i,\cdots,\varphi_n$ is a basis for V we obtain automatically a matrix representation $T_{ij}(g)$ defined by (4.3).

Two matrix representations T and T' are said to be equivalent if there exists a nonsingular matrix S such that

$$T'(g)=ST(g)S^{-1}.$$

A subspace W is invariant under T if $T(g)w\in W$ for every $g\in\mathcal{G}$ and every $w\in W$. The representation T is *reducible* if there is a proper subspace of V which is invariant under T. In finite dimensions, every representation can be assumed to be unitary, and the vector space V decomposes into a direct sum of mutually orthogonal irreducible invariant subspaces. We write

$$V=\sum_{k=1}^{l}\oplus V_k,\qquad T\cong\sum_{k=1}^{l}\oplus T^{(k)},\tag{6.1}$$

where the representations $T^{(k)}$ are irreducible unitary representations on the subspaces V_k.

Given a finite group $\mathcal{G}$, there are a finite number of irreducible representations $T^{(\mu)}$ of $\mathcal{G}$. The dimension of each representation is denoted by n_μ, while the number α of nonequivalent irreducible representations is equal to the number of conjugacy classes in $\mathcal{G}$.

Given an arbitrary representation T on a vector space V, we may write

$$T\cong\sum_{\mu=1}^{\alpha}\oplus a_\mu T^{(\mu)}$$

where the integer a_μ is the multiplicity of the irreducible representation $T^{(\mu)}$ in T.

The character of a representation is the function on $\mathcal{G}$ given by

$$\chi(g)=\operatorname{tr}T(g).$$

The character is basis independent since the trace of a matrix is invariant under a change of basis. Moreover, *the character χ is constant on conjugacy classes.*

An inner product is defined on characters by

$$\langle\chi,\tilde{\chi}\rangle=\frac{1}{N}\sum_{g\in\mathcal{G}}\chi(g)\tilde{\chi}(g^{-1})$$

where N is the order of $\mathcal{G}$. The multiplicities a_μ are then given by

$$a_\mu=\langle\chi,\chi^{(\mu)}\rangle.$$

If V and W are vector spaces with representations T, T' on V and W, respectively, then the *tensor product* representation $T\otimes T'$ on $V\otimes W$ is defined by

$$(T(g)\otimes T'(g))(\mathbf{v}\otimes\mathbf{w}) = T(g)\mathbf{v}\otimes T'(g)\mathbf{w}$$

and extending $T\otimes T'$ by linearity to the entire space $V\otimes W$. One can easily show that the character of $T\otimes T'$ is

$$(\chi\otimes\chi')(g) = \chi(g)\chi'(g).$$

If $\{T^{(\mu)}\}$, $1\leqq\mu\leqq\alpha$, are a complete set of nonequivalent irreducible representations, then we can expand

$$T^{(\mu)}\otimes T^{(\upsilon)} = \sum_{\xi=1}^{\alpha}\oplus a_\xi T^{(\xi)}. \tag{6.2}$$

That is, $V\otimes V$ splits into a direct sum of irreducible invariant subspaces under $T^{(\mu)}\otimes T^{(\upsilon)}$, and this representation may be decomposed accordingly into a direct sum of irreducible representations. The multiplicities a_ξ are given by

$$a_\xi = \langle\chi^{(\mu)}\chi^{(\upsilon)}, \chi^{(\xi)}\rangle. \tag{6.3}$$

An exposition of the above theory may be found in [8, Chap. 3].

Now suppose B is a bilinear mapping from $V\times V$ to V possessing the covariance property

$$T(g)B(\mathbf{u}, \mathbf{v}) = B(T(g)\mathbf{u}, T(g)\mathbf{v}). \tag{6.4}$$

Let the tensor F in $V^{\otimes 3}$ be given by

$$F(\mathbf{u}, \mathbf{v}, \mathbf{w}) = (B(\mathbf{u}, \mathbf{v}), \mathbf{w}). \tag{6.5}$$

Then F is invariant under $T^{\otimes 3}$ if T is unitary with respect to the inner product $(\cdot, \cdot)$. In fact,

$$\begin{aligned} T^{\otimes 3}F(\mathbf{u}, \mathbf{v}, \mathbf{w}) &= F(T\mathbf{u}, T\mathbf{v}, T\mathbf{w}) = (B(T\mathbf{u}, T\mathbf{v}), T\mathbf{w}) \\ &= (TB(\mathbf{u}, \mathbf{v}), T\mathbf{w}) = (B(\mathbf{u}, \mathbf{v}), \mathbf{w}) \\ &= F(\mathbf{u}, \mathbf{v}, \mathbf{w}). \end{aligned}$$

We have shown that

$$T^{\otimes 3}F = F. \tag{6.6}$$

The set of all tensors F satisfying (6.6) constitutes a subspace of $V^{\otimes 3}$. If the tensor $F = F_{ijk}\varphi_i\otimes\varphi_j\otimes\varphi_k$ has the property (6.6), then the mapping $B(\mathbf{u}, \mathbf{v}) = F_{ijk}u_iv_j\varphi_k$ has the covariance property (6.4). This gives us a convenient way to compute such covariant bilinear mappings B.

7. Subspaces of tensors invariant under a symmetry group. As a first application of group representation theory, let us calculate the dimension of the subspace of tensors (6.4) which are invariant under a symmetry group. The dimension of the tensor product space $V^{\otimes 3}$ in which F lies is n^3, where $n = \dim V$. The dimension of the subspace of tensors F invariant under $T^{\otimes 3}$, or equivalently, for which

$$T_gB(\mathbf{u}, \mathbf{v}) = B(T_g\mathbf{u}, T_g\mathbf{v}), \tag{7.1}$$

is in general smaller, depending on the symmetry group $\mathcal{G}$. A further reduction in the dimension of the subspace of invariant tensors is achieved because the tensor F may be assumed to be symmetric in the first two variables. In fact, from the bifurcation equation (3.8), we see that the only term which appears is the symmetric expression $B(\mathbf{u}, \cdots, \mathbf{u})$. In the case $k = 2$, we replace $B(\mathbf{u}, \mathbf{v})$ by

$$\tfrac{1}{2}(B(\mathbf{u}, \mathbf{v}) + B(\mathbf{v}, \mathbf{u}));$$

then the bifurcation equation $\xi + B(\xi, \xi) = 0$ is unchanged, and F is symmetric in its first two variables.

Let us calculate the dimension of the invariant subspace when the symmetry group $\mathcal{G}$ is D_3 or D_4. These are the symmetry groups of the equilateral triangle and square, respectively. The group D_3, of order six, is generated by two elements g, h, with $g^3 = h^2 = e$ and $hgh = g^{-1}$. The element g represents a counterclockwise rotation through 120° and h represents a reflection across a median. In terms of the diagram in Fig. 7.1, g is equivalent to the permutation (1 2 3), while h is the permutation (1 2). The conjugacy classes of D_3 are $\{e\}$, $\{g, g^2\}$, and $\{h, gh, g^2h\}$ [8, *p.* 87].

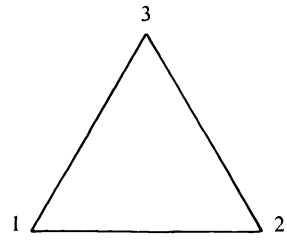


FIG. 7.1

There are three irreducible representations of dimensions n_1, n_2, n_3 respectively, with $n_1 = 1$, $n_2 = 1$, $n_3 = 2$. The character table for group D_3 is Table 7.1. (See [8, p. 87].)

TABLE 7.1

D_3	$\{e\}$	$\{g, g^2\}$	$\{h, gh, g^2h\}$
$\chi^{(1)}$	1	1	1
$\chi^{(2)}$	1	1	−1
$\chi^{(3)}$	2	−1	0

Suppose the dimension of the nullspace V is four and that V decomposes into the irreducible invariant subspaces

$$V = V^{(1)} \oplus V^{(2)} \oplus V^{(3)},$$

where $\dim V^{(\mu)} = n_\mu$, $\mu = 1, 2, 3$. The representation T then can be written

$$T = T^{(1)} \oplus T^{(2)} \oplus T^{(3)}. \tag{7.2}$$

The character χ of T is

$$\chi(g) = \operatorname{tr} T(g) = \chi^{(1)}(g) + \chi^{(2)}(g) + \chi^{(3)}(g).$$

Thus, for χ, we have Table 7.2.

TABLE 7.2

D_3	$\mathscr{E}$	$2\mathscr{C}_3$	$3\mathscr{C}_2$
χ	4	1	0
χ^2	16	1	0
χ^s	10	1	2

The character of $T \otimes T$ is $\chi^2(g)$, but the character of $T \otimes T$ restricted to the subspace of symmetric tensors on $V \otimes V$ is given instead by

$$\chi^s(\sigma) = \tfrac{1}{2}(\chi^2(\sigma) + \chi(\sigma^2)).$$

When $\sigma = e$, $\chi(\sigma) = 4$, $\chi^2 = 16$ and $\chi(\sigma^2) = \chi(e) = 4$. So $\chi^s(e) = 10$. When $\sigma = g$, $\chi^2(g) = 1$ but $\chi(g^2) = 1$, so $\chi^s(g) = 1$. When $\sigma = h$, $h^2 = e$ so $\chi^s(h) = \frac{1}{2}(\chi^2(h) + \chi(e)) = \frac{1}{2}(0+4) = 2$. Therefore the character of $(T \otimes T)^s \otimes T$ is Table 7.3.

TABLE 7.3

D_3	$\mathscr{E}$	$2\mathscr{C}_3$	$3\mathscr{C}_2$
$\chi^s\chi$	40	1	0

Now we write

$$(T \otimes T)^s \otimes T = a_1 T^{(1)} + a_2 T^{(2)} + a_3 T^{(3)}$$

where a_i is the multiplicity of the representation $T^{(i)}$. Since $T^{(1)}$ is the identity representation, that is,

$$T^{(1)}\mathbf{a} = \mathbf{a} \qquad \text{for all } \sigma \in D_3,$$

we want to know the number a_1. This number is the dimension of the subspace of tensors which are invariant under $(T \otimes T)^s \otimes T$.

From (6.4),

$$a^{(1)} = \langle \chi^s\chi, \chi^{(1)} \rangle = \tfrac{1}{6}(1 \cdot \chi^s\chi(e) + 2\chi^s\chi(g) + 3\chi^s\chi(h))$$
$$= \tfrac{1}{6}(1 \cdot 40 + 2 \cdot 1 + 3 \cdot 0) = 7.$$

In this case, the dimension of the given subspace is 7; there are 7 linearly independent tensors B such that

$$T_g B(\mathbf{u}, \mathbf{v}) = B(T_g\mathbf{u}, T_g\mathbf{v}) \quad \text{and} \quad B(\mathbf{u}, \mathbf{v}) = B(\mathbf{v}, \mathbf{u})$$

when the group G is D_3. There are therefore 7 independent tensor quantities

$$B_{jki} = (B(\boldsymbol{\varphi}_j, \boldsymbol{\varphi}_k), \boldsymbol{\varphi}_i)$$

to be evaluated. We shall compute the precise form of the general tensor in § 8.

8. Computation of invariant bifurcation equations: An example. In this section we compute the general bifurcation equations (1.2) when B is invariant under the symmetry group D_3 and $\mathbf{v}$ lies in a four-dimensional vector space V. We assume that the representation T on V is given by $T = T^{(1)} \oplus T^{(2)} \oplus T^{(3)}$ as in (7.2). We emphasize that this is only an assumption for the sake of example. In practice, the representation T must be computed. This is done explicitly for a problem in fluid mechanics in [16]. (See Theorem 6.1.)

The bifurcation equations $\mathbf{v} + B(\mathbf{v}, \mathbf{v}) = 0$ may be written in the form

$$v_i + B_{jki} v_j v_k = 0, \qquad i, j, k = 1, \cdots, 4.$$

The components B_{jki} comprise a set of 64 quantities. However, as we saw in the previous section, only seven of these quantities are independent. In fact, we shall see that these equations may be reduced to the following special form:

$$\begin{aligned} &v_1 + a v_1^2 + b v_2^2 + c(v_3^2 + v_4^2) &&= 0, \\ &v_2 + 2d v_1 v_2 &&= 0, \\ &v_3 + 2e v_1 v_3 + 2f v_2 v_4 + 2g v_3 v_4 &&= 0, \\ &v_4 + 2e v_1 v_4 - 2f v_2 v_3 + g(v_3^2 - v_4^2) &&= 0, \end{aligned} \tag{8.1}$$

where the seven parameters $a, \cdots, g$ are given by

$$\begin{aligned} a &= B_{111} = (B(\varphi_1, \varphi_1), \varphi_1), \\ b &= B_{221} = (B(\varphi_2, \varphi_2), \varphi_1), \\ c &= B_{331} = (B(\varphi_3, \varphi_3), \varphi_1), \\ d &= B_{122} = (B(\varphi_1, \varphi_2), \varphi_2), \\ e &= B_{133} = (B(\varphi_1, \varphi_3), \varphi_3), \\ f &= B_{243} = (B(\varphi_2, \varphi_4), \varphi_3), \\ g &= B_{343} = (B(\varphi_3, \varphi_4), \varphi_3). \end{aligned}$$

Step 1. Choose a standard matrix representation for the two-dimensional representation $T^{(3)}$. We take

$$g \sim \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \qquad h \sim \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \tag{8.2}$$

[8, p. 87]. This is the matrix representation obtained by considering the action of D_3 as the group of rotations and reflections which preserve an equilateral triangle in the plane. Whenever we choose a basis $\{\mathbf{w}_1, \mathbf{w}_2\}$ for a two-dimensional subspace

which transforms according to the rep $T^{(3)}$, we choose $\mathbf{w}_1$ and $\mathbf{w}_2$ so that they transform according to the above matrices. That is,

$$
\begin{aligned}
T(g)\mathbf{w}_1 &= -\frac{1}{2}\mathbf{w}_1 - \frac{\sqrt{3}}{2}\mathbf{w}_2, \\
T(g)\mathbf{w}_2 &= \frac{\sqrt{3}}{2}\mathbf{w}_1 - \frac{1}{2}\mathbf{w}_2, \\
T(h)\mathbf{w}_1 &= -\mathbf{w}_1, \qquad T(h)\mathbf{w}_2 = \mathbf{w}_2.
\end{aligned} \tag{8.3}
$$

Such a basis will be called a *standard* basis.

Step 2. Construct a standard basis for V.

In this case, we denote a standard basis for V by $\{\varphi_1, \varphi_2, \varphi_3, \varphi_4\}$, where $T(g)\varphi_1 = \chi^{(1)}(g)\varphi_1$, $T(g)\varphi_2 = \chi^{(2)}(g)\varphi_2$ and $\{\varphi_3, \varphi_4\}$ transform according to the standard matrix representation $T^{(3)}$ above.

Step 3. Construct a table of tensor products of irreducible reps $T^{(\mu)} \otimes T^{(\nu)}$ (Table 8.1). Such a table is readily constructed by using the character table for D_3 and the formula (6.3). Some simplifications are obtained, for example, by noting that

$$
\begin{aligned}
\chi^{(2)}\chi^{(2)} = \chi^{(1)} &\Rightarrow T^{(2)} \otimes T^{(2)} = T^{(1)}, \\
\chi^{(2)}\chi^{(3)} = \chi^{(3)} &\Rightarrow T^{(2)} \otimes T^{(3)} = T^{(3)},
\end{aligned}
$$

etc.

TABLE 8.1
Tensor product table

	$T^{(1)}$	$T^{(2)}$	$T^{(3)}$
$T^{(1)}$	$T^{(1)}$	$T^{(2)}$	$T^{(3)}$
$T^{(2)}$	$T^{(2)}$	$T^{(1)}$	$T^{(3)}$
$T^{(3)}$	$T^{(3)}$	$T^{(3)}$	$T^{(1)} \oplus T^{(2)} \oplus T^{(3)}$

Step 4. Compute $T \otimes T$ from Table 8.1 in Step 2.

$$T \otimes T = 3T^{(1)} \oplus 3T^{(2)} \oplus 5T^{(3)}.$$

This expansion shows that the sixteen-dimensional space $V \otimes V$ decomposes into 3 one-dimensional subspaces which transform according to $T^{(1)}$, 3 one-dimensional subspaces which transform according to $T^{(2)}$, and 5 two-dimensional subspaces which transform according to $T^{(3)}$.

Step 5. Construct a standard basis for $V \otimes V$.

First we find the 3 one-dimensional subspaces which transform according to T_1. From the tensor product table (Table 8.1), we see that $T^{(1)}$ arises in $T^{(1)} \otimes T^{(1)}$, $T^{(2)} \otimes T^{(2)}$ and $T^{(3)} \otimes T^{(3)}$. The first two subspaces are therefore given by $\varphi_1 \otimes \varphi_1$ and $\varphi_2 \otimes \varphi_2$, while the third one is a linear combination of φ_3 and φ_4:

$$\sum_{i,j=3,4} a_{ij}\varphi_i \otimes \varphi_j.$$

The vector must transform according to the identity representation, so we must have

$$\sum_{i,j=3,4} a_{ij} T^{(3)}(\sigma)\varphi_i \otimes T^{(3)}(\sigma)\varphi_j = \sum_{i,j=3,4} a_{ij}\varphi_i \otimes \varphi_j$$

for all $\sigma \in D_3$.

By choosing $\sigma = g, h$, the generators of D_3, we may determine the coefficients a_{ij} up to a constant factor. The result is $\varphi_3 \otimes \varphi_3 + \varphi_4 \otimes \varphi_4$.

Next we compute the one-dimensional subspaces which transform according to $T^{(2)}$, and then the 5 two-dimensional subspaces which transform according to $T^{(3)}$. We leave the first case to the reader and take up the second. From the Table 8.1 we see which tensor products contain $T^{(3)}$. From the fact that $T^{(3)}$ is contained in $T^{(1)} \otimes T^{(3)}$, $T^{(2)} \otimes T^{(3)}$, $T^{(3)} \otimes T^{(1)}$ and $T^{(3)} \otimes T^{(2)}$, we get the subspaces

$$\begin{aligned} T^{(1)} \otimes T^{(3)} &\quad \text{yields} \quad \{\varphi_1 \otimes \varphi_3, \varphi_1 \otimes \varphi_4\}, \\ T^{(3)} \otimes T^{(1)} &\quad \text{yields} \quad \{\varphi_3 \otimes \varphi_1, \varphi_4 \otimes \varphi_1\}, \\ T^{(2)} \otimes T^{(3)} &\quad \text{yields} \quad \{\varphi_2 \otimes \varphi_4, \varphi_2 \otimes \varphi_3\}, \\ T^{(3)} \otimes T^{(3)} &\quad \text{yields} \quad \{\varphi_4 \otimes \varphi_2, \varphi_3 \otimes \varphi_2\}. \end{aligned}$$

The remaining one is more difficult. Since $T^{(3)}$ is contained in $T^{(3)} \otimes T^{(3)}$, we get a two-dimensional subspace of the form

$$\mathbf{w}_1 = \sum_{i,j=3,4} a_{ij}\varphi_i \otimes \varphi_j, \qquad \mathbf{w}_2 = \sum_{i,j=3,4} b_{ij}\varphi_i \otimes \varphi_j.$$

The coefficients a_{ij} and b_{ij} are to be determined so that $\mathbf{w}_1$ and $\mathbf{w}_2$ transform according to (9.4). The result is listed in Table 8.2 below.

We have now got all the invariant subspaces of $V \otimes V$. It remains to compute the one-dimensional invariant subspaces of $V^{\otimes 3}$ which transform according to the identity representation.

Step 6. Computation of the tensors F such that $T^{\otimes 3}F = F$.

We begin by proving

LEMMA 8.1. *Let M and N be complex finite-dimensional inner product spaces and let R and S be irreducible unitary representations on M and N, respectively. Let $\bar{R}$ denote a complex conjugate representation for R* (defined below). *Then a necessary and sufficient condition for $\bar{R} \otimes S$ to contain the identity representation is that R and S be unitarily equivalent. In that case, the identity representation is contained precisely once. If R is a real representation then $R \otimes S$ contains the identity representation once if and only if R and S are unitarily equivalent.*

Proof. A quick proof may be had in the case of finite groups (or compact Lie groups) by the use of characters. If χ_R and χ_S are the characters of R and S, then the character of $\bar{R} \otimes S$ is $\bar{\chi}_R \chi_S$. The number of times $\bar{R} \otimes S$ contains the identity representation is given by (see (6.3))

$$a_1 = \langle \bar{\chi}_R \chi_S, 1 \rangle = \langle \chi_S, \chi_R \rangle.$$

This number is either zero or one since R and S are irreducible representations. It is one if and only if R and S are unitarily equivalent. The proof may be extended immediately to compact Lie groups.

The proof below makes no assumption about the group. It is based on a suggestion made to the author by O. E. Lanford III.

The complex conjugate of a representation is defined as follows. Choose an orthonormal basis $\{\varphi_1, \cdots, \varphi_n\}$ for M and let $R_{ij}(g)$ be the matrix of $R(g)$ relative to $\{\varphi_1, \cdots, \varphi_n\}$. Then $\overline{R_{ij}(g)}$ defines a new matrix representation of the group. The collection of such matrix representations obtained by choosing all possible orthonormal bases for M defines a class of unitarily equivalent matrix representations. In this way, we obtain the conjugate representation.

Let M^* be the space of linear functionals on M and denote by $\langle \cdot, \cdot \rangle$ the bilinear functional on $M \times M^*$ given by

$$\langle u, v^* \rangle = v^*(u).$$

By contrast, the inner product $(\cdot, \cdot)$ is anti-linear in the second variable. Given an operator R, denote by R^* and R^+, respectively, the adjoints of R relative to the bilinear form $\langle \cdot, \cdot \rangle$ and the inner product $(\cdot, \cdot)$. If $R_{ij}(g)$ is the matrix of R relative to $\{\varphi_1, \cdots, \varphi_n\}$, then R^* has the matrix $R_{ij}(g)$ and R^+ has the matrix $\overline{R_{ij}(g)}$. The matrix of R^* is that one which is obtained by choosing the dual basis $\{\varphi_1^*, \cdots, \varphi_n^*\}$ for M^* $\cdot$ $(\varphi_j^*(\varphi_i) = \langle \varphi_i, \varphi_j^* \rangle = \delta_{ij}.)$ Let $\tilde{R}(g) = R^*(g^{-1})$. The matrix of $\tilde{R}(g)$ is

$$\tilde{R}_{ij}(g) = (R^*(g^{-1}))_{ij} = R_{ji}(g^{-1}) = \overline{R_{ij}(g)},$$

provided R is unitary. Thus the contragradient representation is unitarily equivalent to the conjugate representation if R is unitary. It therefore suffices to prove the lemma for the product representation $\tilde{R} \otimes S$.

We identify $M^* \otimes N$ with $\mathscr{L}(M, N)$ by the isomorphism

$$\sum a_{ij} \varphi_i^* \otimes \psi_j \leftrightarrow \sum a_{ij} \langle \cdot, \varphi_i^* \rangle \psi_j,$$

where the expression on the right denotes the linear mapping

$$\varphi \to \sum a_{ij} \langle \varphi, \varphi_i^* \rangle \psi_j$$

from M to N. Here the $\{\varphi_i^*\}$ form a basis for M^*, while the $\{\psi_j\}$ form a basis for N. Define a representation U on $\mathscr{L}(M, N)$ by $UA = SAR^{-1}$. This representation is equivalent to the tensor product representation $\tilde{R} \otimes S$ on $M^* \otimes N$. In fact,

$$\begin{aligned} \sum a_{ij} R^*(g^{-1}) \varphi_i^* \otimes S\psi_j &\leftrightarrow \sum a_{ij} \langle \cdot, R^*(g^{-1}) \varphi_i^* \rangle S\psi_j \\ &= \sum a_{ij} \langle R^{-1}(g) \cdot, \varphi_i^* \rangle S\psi_j. \end{aligned}$$

Therefore $\tilde{R} \otimes S$ contains the identity representation precisely as many times as does U. But U contains the identity representation if and only if there exist transformations A in $\mathscr{L}(M, N)$ such that $UA = A$, i.e., $SA = AR$. By Schur's lemma [8, Thm. 3.4, p. 64] this happens if and only if R and S are unitarily equivalent. In that case, M and N are isomorphic since R and S are assumed to be irreducible. We can therefore put $R = S$ and $M = N$ and apply Schur's second theorem [8, Thm. 3.5] to conclude that $A = \lambda E$, where E is the identity transformation. There is thus a one-dimensional subspace of $\mathscr{L}(M, N)$ which is invariant under U, so U contains the identity representation only once.

In the present application, R is a real representation.

We are now ready to complete our task. Lemma 8.1 is to be applied to the case where M is an irreducible subspace of $V\otimes V$ and N is an irreducible subspace of V. From Step 4, we saw that $T\otimes T$ could be written as a direct sum of irreducible representations on $V\otimes V$, and we take one of these irreducible representations for R. Similarly, the representation S of Lemma 8.1 is taken to be one of the irreducible representations in the composition of T. By the lemma, these must be equivalent if their tensor product is to contain the identity representation. Thus, from $\varphi_1\otimes\varphi_1$ and φ_1, we get the subspace $\varphi_1\otimes\varphi_1\otimes\varphi_1$, etc. We take all three subspaces of $V\otimes V$ which transform according to $T^{(1)}$, tensor these with φ_1, and get three subspaces of $V^{\otimes 3}$ which transform according to $T^{(1)}$. Then we take all subspaces of $V\otimes V$ which transform according to $T^{(2)}$, tensor these with φ_2, and obtain three additional invariant subspaces of $V^{\otimes 3}$. Finally we come to the two-dimensional subspaces of $V\otimes V$. These must be tensored with the two-dimensional subspaces of V in such a way as to obtain the invariant subspaces of $V^{\otimes 3}$.

Exactly how this is to be done is already given to us in our calculation of the invariant subspaces of $V\otimes V$, where we saw that $\varphi_3\otimes\varphi_3+\varphi_4\otimes\varphi_4$ transforms according to $T^{(1)}$. More generally, we have

LEMMA 8.2. *Let V, W each be two-dimensional subspaces which transform according to $T^{(3)}$. Let $\{\mathbf{w}_1, \mathbf{w}_2\}$ and $\{\psi_1, \psi_2\}$ be standard bases for V and W. Then the invariant subspace of $V\otimes W$ which transforms according to the identity representation is*

$$\mathbf{w}_1\otimes\psi_1+\mathbf{w}_2\otimes\psi_2.$$

Accordingly, the subspace $\{\varphi_1\otimes\varphi_3, \varphi_1\otimes\varphi_4\}$ must be tensored with $\{\varphi_3, \varphi_4\}$ in the following way:

$$(\varphi_1\otimes\varphi_3)\otimes\varphi_3+(\varphi_1\otimes\varphi_4)\otimes\varphi_4.$$

A complete list of all subspaces of $V^{\otimes 3}$ which transform according to $T^{(1)}$ are given in the third column of Table 8.2. It is then an easy matter to write down the bifurcation equations (8.1).

TABLE 8.2

V	$V\otimes V$	$V\otimes V\otimes V$
$T^{(1)}$	$\varphi_1\otimes\varphi_1$	$\varphi_1\otimes\varphi_1\otimes\varphi_1$
	$\varphi_2\otimes\varphi_2$	$\varphi_2\otimes\varphi_2\otimes\varphi_1$
φ_1	$\varphi_3\otimes\varphi_3+\varphi_4\otimes\varphi_4$	$(\varphi_3\otimes\varphi_3+\varphi_4\otimes\varphi_4)\otimes\varphi_1$
$T^{(2)}$	$\varphi_1\otimes\varphi_2$	$\varphi_1\otimes\varphi_2\otimes\varphi_2$
φ_2	$\varphi_2\otimes\varphi_1$	$\varphi_2\otimes\varphi_1\otimes\varphi_2$
	$\varphi_3\otimes\varphi_4-\varphi_4\otimes\varphi_3$	$(\varphi_3\otimes\varphi_4-\varphi_4\otimes\varphi_3)\otimes\varphi_2$
$T^{(3)}$	$\{\varphi_1\otimes\varphi_3, \varphi_1\otimes\varphi_4\}$	$\varphi_1\otimes\varphi_3\otimes\varphi_3+\varphi_1\otimes\varphi_4\otimes\varphi_4$
$\{\varphi_3, \varphi_4\}$	$\{\varphi_3\otimes\varphi_1, \varphi_4\otimes\varphi_1\}$	$\varphi_3\otimes\varphi_1\otimes\varphi_3+\varphi_4\otimes\varphi_1\otimes\varphi_4$
	$\{\varphi_2\otimes\varphi_4, -\varphi_2\times\varphi_3\}$	$\varphi_2\otimes\varphi_4\otimes\varphi_3-\varphi_2\otimes\varphi_3\otimes\varphi_4$
	$\{\varphi_4\otimes\varphi_2, -\varphi_3\otimes\varphi_2\}$	$\varphi_4\otimes\varphi_2\otimes\varphi_3-\varphi_3\otimes\varphi_2\otimes\varphi_4$
	$\{\varphi_3\otimes\varphi_4+\varphi_4\otimes\varphi_3,$	$(\varphi_3\otimes\varphi_4+\varphi_4\otimes\varphi_3)\otimes\varphi_3$
	$\varphi_3\otimes\varphi_3-\varphi_4\otimes\varphi_4\}$	$+(\varphi_3\otimes\varphi_3-\varphi_4\otimes\varphi_4)\otimes\varphi_4$

Acknowledgments. We wish to thank Professors L. E. Scriven and W. Miller for helpful discussions of some of the ideas presented in this paper.

REFERENCES

[1] J. I. GIMITRO, *Concentration patterns generated by reaction and diffusion*, Ph.D. thesis, Univ. of Minnesota, Minneapolis, 1969.

[2] L. GRAVES, *Remarks on singular points of functional equations*, Trans. Amer. Math. Soc., 79 (1955), pp. 150–157.

[3] M. HAMERMESH, *Group Theory and its Application to Physical Problems*, Addison-Wesley, Reading, Mass., 1962.

[4] K. KIRCHGÄSSNER, *Multiple eigenvalue bifurcation for holomorphic mappings*, Contributions to Nonlinear Functional Analysis, Academic Press, New York, 1971, pp. 69–97.

[5] H. KIELHÖFER AND K. K.IRCHGÄSSNER, *Stability and bifurcation in fluid mechanics*, Rocky Mountain J. Math., 3 (1973), pp. 275–318.

[6] G. YA LYUBARSKII, *The Application of Group Theory in Physics*, Pergamon Press, New York, 1960.

[7] J. B. MCLEOD AND D. H. SATTINGER. *Loss of stability and bifurcation at double eigenvalues*, J. Functional Anal., 14 (1973), pp. 62–84.

[8] W. MILLER, *Symmetry Groups and Their Applications*, Academic Press, New York, 1972.

[9] H. G. OTHMER AND L. E. SCRIVEN, *Nonlinear aspects of dynamic pattern in cellular networks*, J. Theor. Biol., 43 (1974), pp. 83–112.

[10] ———, *Instability and dynamic pattern in cellular networks*, Ibid., 32 (1971), pp. 507–537.

[11] D. RUELLE, *Bifurcations in the presence of a symmetry group*, Arch. Rational Mech. Anal., 51 (1973), pp. 136–152.

[12] D. SATHER, *Branching of solutions of an equation in Hilbert space*, Ibid., 36 (1970), pp. 47–64.

[13] ———, *Branching of solutions of nonlinear equations*, Rocky Mountain J. Math., 3 (1973), pp. 203–250.

[14] D. H. SATTINGER, *Topics in Stability and Bifurcation Theory*, Springer Mathematics Lecture Notes, vol. 309, Springer-Verlag, New York, 1973.

[15] ———, *Transformation groups and bifurcation at multiple eigenvalues*, Bull. Amer. Math. Soc., 79 (1973), pp. 709–711.

[16] ———, *Group representation theory, bifurcation theory and pattern formation*, preprint.

[17] T. H. SCHWAB, *Dynamic concentration patterns generated by reaction and diffusion in open planar systems*, Ph.D. dissertation, Univ. of Minnesota, Minneapolis, 1975.

SIAM J. MATH. ANAL.
Vol. 8, No. 2, April 1977

ON THE ZEROS OF CERTAIN JACOBI POLYNOMIAL SUMS*

J. BUSTOZ†

Abstract. Let $P_k^{(\alpha,\beta)}(x)$ be the Jacobi polynomial of degree k with parameters α, β. It is known that

$$\sum_{k=0}^{n} \frac{P_k^{(\beta,\beta)}(\cos\theta)}{P_k^{(\beta,\beta)}(1)} \geqq 0 \quad \text{for } \beta \geqq 0.$$

It has been conjectured that

$$\sum_{k=0}^{n} \frac{P_k^{(\beta,\beta)}(\cos\theta)}{P_k^{(\beta,\beta)}(1)} z^k \neq 0 \quad \text{if } |z|<1, \quad \beta>0.$$

This conjecture has been verified for $\beta = 0$ and $\beta = \frac{1}{2}$. Here we prove the conjecture for $\beta = 1$ and $\beta = 2$ and give a more general inequality valid for $\beta = 0, \frac{1}{2}, 1, 2$.

1. Introduction. Let $P_k^{(\alpha,\beta)}(x)$ denote the Jacobi polynomial with parameters α, β and set

$$R_k^{(\beta)}(\cos\theta) = P_k^{(\beta,\beta)}(\cos\theta)/P_k^{(\beta,\beta)}(1).$$

It is known that if $\beta \geqq 0$ and $0 \leqq \theta \leqq \pi$, then

$$\sum_{k=0}^{n} R_k^{(\beta)}(\cos\theta) \geqq 0. \tag{1.1}$$

When $\beta = \frac{1}{2}$, inequality (1.1) reduces to the well-known Fejer–Jackson inequality

$$\sum_{k=0}^{n} \frac{\sin(k+1)\theta}{k+1} \geqq 0. \tag{1.2}$$

R. Askey has conjectured [1] that the inequality (1.1) can be generalized to the following: If $\beta \geqq 0$, $0 \leqq \theta \leqq \pi$ and $|z|<1$, then

$$P_n(z;\theta,\beta) = \sum_{k=0}^{n} R_k^{(\beta)}(\cos\theta) z^k \neq 0. \tag{1.3}$$

Inequality (1.3) is known to be true for $\beta = 0$ and $\beta = \frac{1}{2}$. When $\beta = 0$, the polynomials $R_k^{(\beta)}(\cos\theta)$ are the Legendre polynomials. Szegö proved the inequality (1.3) when $\beta = 0$ [4], and the case $\beta = \frac{1}{2}$ is proved in [2]. In this note, we will prove (1.3) for $\beta = 1$ and $\beta = 2$ and will show that a more general inequality is valid for $\beta = 0, \frac{1}{2}, 1, 2$, and is also true in the limit as $\beta \to \infty$. (This limiting case was observed by R. Askey.)

2. Development. Define the functions $G_\beta(z;\theta)$ by

$$G_\beta(z;\theta) = \sum_{k=0}^{\infty} R_k^{(\beta)}(\cos\theta) z^{k+1}. \tag{2.1}$$

* Received by the editors August 1, 1975.

† Department of Mathematical Sciences, University of Cincinnati, Cincinnati, Ohio. Now at Department of Mathematics, Arizona State University, Tempe, Arizona 85281.

We will show that the partial sums $S_n(z;\theta,\beta)$ of (2.1) vanish in $|z|<1$ only when $z=0$. This will prove (1.3) for these values of β since $S_n(z;\theta,\beta)=zP_n(z;\theta,\beta)$. By using Gegenbauer's integral

$$R_k^{(\beta)}(\cos\theta)=\frac{\Gamma(\beta+1)}{\Gamma(\beta+\frac{1}{2})\Gamma(\frac{1}{2})}\int_0^\pi [\cos\theta+i\sin\theta\cos t]^k\sin^{2\beta}t\,dt,$$

we can prove that the functions G_β for $\beta>-\frac{1}{2}$ satisfy the recursion relation

$$G_{\beta+1}=\frac{2(\beta+1)}{2\beta+1}\frac{(1-2z\cos\theta+z^2)G_\beta-z(1-z\cos\theta)}{z^2\sin^2\theta}. \tag{2.2}$$

R. Askey pointed out to the author that $G_\beta(z;\theta)$ can be written in a hypergeometric form by using Gegenbauer's integral. The formula thus obtained is

$$G_\beta(z;\theta)=\frac{z}{1-z\cos\theta}\,{}_2F_1\left(\tfrac{1}{2},1;\beta+1;-\left(\frac{z\sin\theta}{1-z\cos\theta}\right)^2\right).$$

The recursion relation (2.2) can then be obtained by applying one of the Gauss contiguous relations satisfied by ${}_2F_1$'s. The fact that $G_\beta(z;\theta)$ is a multiple of a ${}_2F_1$ might be of use in proving a conjecture about $G_\beta(z;\theta)$ that we will make presently. When $\beta=0$, (2.1) is the generating relation for the Legendre polynomials so that

$$G_0(z;\theta)=z(1-2z\cos\theta+z^2)^{-1/2},$$

and thus by applying (2.2), we can sum the series (2.1) when β is a natural number. It is convenient for the moment to write $A=(1-2z\cos\theta+z^2)^{1/2}$ and $B=1-z\cos\theta$. Then simple computations give

$$G_1(z;\theta)=\frac{2z}{A+B},\qquad G_2(z;\theta)=\frac{4z(2A+B)}{3(A+B)^2}.$$

It is possible to establish a general formula for G_n, n a natural number, of the form

$$G_n=z\phi_n(A,B)\cdot(A+B)^{-n},$$

where $\phi_n(A,B)$ is a polynomial in A and B of degree $n-1$.

The proof of (1.3) for $\beta=\frac{1}{2}$ used the notion of univalence, and this will also play a basic role here. A function $f(z)$ analytic in $|z|<1$ is said to be starlike of order $\frac{1}{2}$ if

$$\operatorname{Re}\frac{zf'(z)}{f(z)}\geqq\frac{1}{2};\qquad |z|<1. \tag{2.3}$$

This class of functions will be denoted by $\mathrm{St}(\frac{1}{2})$. If $f\in\mathrm{St}(\frac{1}{2})$, then f is univalent in $|z|<1$. St. Ruscheweyh and T. Sheil-Small have proved [3] that if $f\in\mathrm{St}(\frac{1}{2})$, $f=z+a_2z^2+\cdots$, and if $S_n(z)$ is the partial sum of f, then

$$\operatorname{Re}\frac{f(z)}{S_n(z)}\geqq\tfrac{1}{2},\qquad |z|<1. \tag{2.4}$$

Since f is univalent and $f(0)=0$, (2.4) implies that $S_n(z)/z$ does not vanish in $|z|<1$. Thus one way to prove (1.3) is to prove that $G_\beta \in \mathrm{St}(\frac{1}{2})$. We conjecture that this is so for $\beta \geqq 0$.

THEOREM 1. *$G_\beta \in \mathrm{St}(\frac{1}{2})$ if $\beta = 0, \frac{1}{2}, 1, 2$.*

Proof. We will omit some computations. For convenience we set $H_\beta(z) = zG'_\beta/G_\beta$.

We find easily that $H_0(z) = (1-z\cos\theta)/(1-2z\cos\theta+z^2)$ and that $\operatorname{Re} H_0 \geqq \frac{1}{2}$. Hence $G_0 \in \mathrm{St}(\frac{1}{2})$. Computing H_1, we find $H_1(z) = G_0(z;\theta)/z$, and we have $\operatorname{Re} H_1 \geqq \frac{1}{2}$ by (2.4) with $n=1$ and $f = G_0$, so $G_1 \in \mathrm{St}(\frac{1}{2})$. For G_2 we get

$$H_2(z) = \frac{3}{2(1-2z\cos\theta+z^2)^{1/2}+1-z\cos\theta}.$$

Now $\operatorname{Re} H_2 \geqq \frac{1}{2}$ for $|z|<1$ if and only if $|(1/H_2)-1|<1$, $|z|<1$. This last inequality is equivalent to

$$|2(1-2z\cos\theta+z^2)^{1/2}-2-z\cos\theta|<3, \qquad |z|<1. \tag{2.5}$$

Since $H_1 \in \mathrm{St}(\frac{1}{2})$, we have that $|(1-2z\cos\theta+z^2)^{1/2}-1|<1$, $|z|<1$, and (2.5) follows by the triangle inequality. Thus $G_1 \in \mathrm{St}(\frac{1}{2})$. A function f analytic in $|z|<1$ is convex univalent if $\operatorname{Re} zf''/f' \geqq -1$ in $|z|<1$. Convex functions are starlike of order $\frac{1}{2}$, so to show $G_{1/2} \in \mathrm{St}(\frac{1}{2})$, we can prove convexity. Computing $G'_{1/2}$, we get

$$G'_{1/2} = (1-2z\cos\theta+z^2)^{-1},$$

and then we get

$$\frac{zG''_{1/2}}{G'_{1/2}} + 1 = \frac{1-z^2}{1-2z\cos\theta+z^2},$$

and we easily find that this last quantity has a nonnegative real part in $|z|<1$. This completes the proof.

As $\beta \to \infty$, the function $G_\beta(z;\theta)$ converges to $G_\infty(z;\theta) = z(1-z\cos\theta)^{-1}$. An easy computation shows that $\operatorname{Re} zG'_\infty/G_\infty \geqq \frac{1}{2}$, and thus our conjecture that $G_\beta(z;\theta) \in \mathrm{St}(\frac{1}{2})$ is true in the limit. We remark that the inequality $\operatorname{Re} G_\beta(z;\theta)/S_n(z;\theta,\beta) \geqq \frac{1}{2}$ for $|z|<1$ is more than we need to conclude inequality (1.3). It would suffice to show that $\operatorname{Re} G_\beta(z;\theta)/S_n(z;\theta,\beta) \geqq 0$. As a matter of fact, this is exactly what Szegö did in order to prove (1.3) for $\beta = 0$.

Although it is possible to find a general formula for G_n when n is a natural number, we have been unable to prove Theorem 1 for all natural numbers n. The sums for H_n become very hard to handle for $n > 2$.

As a consequence of Theorem 1 and inequality (2.4), we may conclude that (1.3) holds for $\beta = 0, \frac{1}{2}, 1, 2$. The fact that (2.4) implies

$$\left|\frac{S_n(z)}{f(z)} - 1\right| < 1, \qquad |z|<1, \tag{2.6}$$

gives a general inequality that includes (1.3). To see this, we observe that (2.6) implies that $S_n(z)/f(z)$ lies in a convex region excluding zero. Hence if

$\lambda_0, \lambda_1, \cdots, \lambda_n$ are nonnegative numbers with at least one of them positive and $S_n(z) = zP_n(z)$, we can conclude that

$$\sum_{k=0}^{n} \lambda_k P_k(z) \neq 0, \qquad |z| < 1. \tag{2.7}$$

In particular, (2.7) holds when we take for $P_k(z)$ the polynomials (1.3). This can be restated as

THEOREM 2. *If $\lambda_k \geqq 0$ is a nonincreasing sequence with at least one $\lambda_k \neq 0$, then for every $n = 0, 1, 2, \cdots$,*

$$\sum_{k=0}^{n} \lambda_k R_k^{(\beta)}(\cos\theta) z^k \neq 0, \qquad |z| < 1, \quad \beta = 0, \tfrac{1}{2}, 1, 2.$$

We conjecture that Theorem 2 holds for $\beta \geqq 0$, as would be true of course if Theorem 1 holds for $\beta \geqq 0$.

In [2] we proved that the polynomials

$$\sum_{k=0}^{n} \binom{n-k+\alpha}{n-k} R_k^{(\beta)}(\cos\theta) z^k$$

do not vanish in $|z| < 1$ if $\alpha \geqq 1$ and $\beta \geqq \frac{1}{2}$, and conjectured that this remains true if $\alpha \geqq 0$, $\beta \geqq 0$. It appears that more than this is true because as a special case of Theorem 2 we have the following:

COROLLARY. *If $\alpha \geqq 0$, $\beta = 0, \frac{1}{2}, 1, 2$ and $m \leqq n$, then*

$$\sum_{k=0}^{m} \binom{n-k+\alpha}{n-k} R_k^{(\beta)}(\cos\theta) z^k \neq 0, \qquad |z| < 1.$$

REFERENCES

[1] R. ASKEY, *Positive Jacobi polynomial sums*, Tôhoku Math. J., 24 (1972), pp. 109–119.

[2] J. BUSTOZ, *Jacobi polynomial sums and univalent Cesaro means*, Proc. Amer. Math. Soc., 50 (1975), pp. 259–264.

[3] ST. RUSCHEWEYH AND T. SHEIL-SMALL, *Hadamard products of schlicht functions and the Polya–Schoenberg conjecture*, Comment. Math. Helv., 48 (1973), pp. 119–136.

[4] G. SZEGÖ, *Zur theorie die Legendreschen Polynome*, Jber. Deutsch. Math. Verein, 40 (1931), pp. 163–166.

SIAM J. MATH. ANAL.
Vol. 8, No. 2, April 1977

MOMENT THEORY FOR WEAK CHEBYSHEV SYSTEMS WITH APPLICATIONS TO MONOSPLINES, QUADRATURE FORMULAE AND BEST ONE-SIDED L^1-APPROXIMATION BY SPLINE FUNCTIONS WITH FIXED KNOTS*

C. A. MICCHELLI† AND ALLAN PINKUS‡

Abstract. The chief purpose of this paper is to present an alternative approach to results concerning the existence and uniqueness of monosplines which have a maximum number of zeros (the fundamental theorem of algebra for monosplines). In addition, we discuss the related problems of "double precision" quadrature formulae and one-sided L^1-approximation by spline functions with fixed knots.

1. Introduction. The chief purpose of this paper is to present an alternative approach to the results of S. Karlin and L. Schumaker [6] and S. Karlin and C. A. Micchelli [3] concerning the existence and uniqueness of monosplines which have a maximum number of zeros (the fundamental theorem of algebra for monosplines). In addition, we discuss the related problems of "double precision" quadrature formulae and one-sided L^1-approximation by spline functions with fixed knots.

Our approach to these problems is based on moment theory. The relationship of the above problems to moment theory is not surprising. In fact, I. J. Schoenberg originally suggested this relationship in [13] and S. Karlin and W. J. Studden [7] discuss a special case of the fundamental theorem of algebra for monosplines by means of moment theory.

However, the method used in [6] (and later in [3], [4] and [10]) to prove Schoenberg's conjecture [13] does not use moment theory and is needlessly complicated. Our proof uses Theorem 2.1; see Theorem 5.1 and Corollary 5.1. Nevertheless, the methods of [6] are indeed valuable when the simplicity of moment theory is not applicable, as in [4] and [10].

A thorough treatment, with improvements, of M. G. Krein's work [8] on moment theory for Chebyshev systems is contained in [7]. The basic difficulty that we face here is to provide a suitable version of these results for weak Chebyshev systems. In his thesis [1], H. Burchard studied the problem of interpolation of data by generalized convex functions and was also led to the problem to extending moment theory to weak Chebyshev systems. His extension, however, is too restrictive for the application we have in mind. For the related problem of determining the "envelope" of smooth functions "pinned down" on some partition, see Micchelli and Miranker [14].

In § 2 we present an extension of moment theory to weak Chebyshev systems, which improves on Burchard's result. We also discuss the related problem of one-sided approximation for weak Chebyshev systems. In § 3 we apply the general theory to certain classes of spline functions to obtain "double precision" quadrature formulae. Section 4 contains our version of the fundamental theorem of algebra for monosplines satisfying mixed boundary conditions, which subsumes [3] and [6].

* Received by the editors September 4, 1975, and in revised form December 22, 1975.

† IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598.

‡ Mathematics Research Center, University of Wisconsin—Madison, Madison, Wisconsin 53706.

2. Moment theory for weak Chebyshev systems. We begin by recalling the main results of moment theory for Chebyshev systems. Let M be a linear subspace of $C[0, 1]$ of dimension n spanned by the functions $u_1(t), \cdots, u_n(t)$ and let $d\alpha(t)$ be a nonnegative finite measure on $[0, 1]$. If $d\alpha(t)$ is a discrete measure

$$\int_0^1 f(r)\, d\alpha(t) = \sum_{j=1}^{N} \lambda_j f(t_j), \qquad f \in C[0, 1],$$

$\lambda_1 > 0, \cdots, \lambda_N > 0$, $0 \leqq t_1 < \cdots < t_N \leqq 1$, then the index of $d\alpha$, denoted by $I(\alpha)$, is defined to be $\sum_{j=1}^{N} \omega(t_j)$ where $\omega(t) = \frac{1}{2}$ when $t \in \{0, 1\}$, and $\omega(t) = 1$ when $t \in (0, 1)$. Following [8], we call $d\alpha(t)$ a positive measure (relative to M) provided that $\int_0^1 u(t)\, d\alpha(t) > 0$ whenever u is a nontrivial nonnegative function in M. A positive measure corresponds to an interior point of the moment space determined by the set of functions $\{u_1(t), \cdots, u_n(t)\}$ [7]. The measure $d\alpha_0$ is said to be a principal representation of $d\alpha$ provided that $\int_0^1 u(t)\, d\alpha_0(t) = \int_0^1 u(t)\, d\alpha(t)$ for all $u \in M$ and $I(\alpha_0) = n/2$. If $d\alpha_0$ has mass at one, it is referred to as an upper principal representation; otherwise, it is called a lower principal representation. The set of functions $\{u_1(t), \cdots, u_n(t)\}$ is called a Chebyshev system on $[0, 1]$ (M a Chebyshev subspace) provided that

$$\begin{vmatrix} u_1(t_1) & \cdots & u_1(t_n) \\ \vdots & & \vdots \\ u_n(t_1) & \cdots & u_n(t_n) \end{vmatrix} > 0$$

for all $0 \leqq t_1 < t_2 < \cdots < t_n \leqq 1$.

The following result of Krein is proven in [7]. If M is a Chebyshev subspace then every positive measure $d\alpha(t)$ has exactly two principal representations,

$$\int_0^1 u(t)\, d\alpha(t) = \int_0^1 u(t)\, d\underline{\alpha}(t) = \int_0^1 u(t)\, d\bar{\alpha}(t), \qquad u \in M,$$

where $d\underline{\alpha}$ is a lower principal representation and $d\bar{\alpha}$ is an upper principal representation.

Let us denote by K the convexity cone generated by M. Thus every $f \in K$ is a function defined on $(0, 1)$ which satisfies the inequality

$$\begin{vmatrix} u_1(t_1) & \cdots & u_1(t_{n+1}) \\ u_2(t_1) & \cdots & u_2(t_{n+1}) \\ \vdots & & \\ u_n(t_1) & \cdots & u_n(t_{n+1}) \\ f(t_1) & \cdots & f(t_{n+1}) \end{vmatrix} \geqq 0,$$

for all $0 < t_1 < \cdots < t_{n+1} < 1$. The principal representations corresponding to a positive measure have the additional property that for $f \in K_0 = C[0, 1] \cap K$,

$$\int_0^1 f(t)\, d\underline{\alpha}(t) \leqq \int_0^1 f(t)\, d\alpha(t) \leqq \int_0^1 f(t)\, d\bar{\alpha}(t). \tag{2.1}$$

This inequality is called the Markoff–Krein inequality.

We shall now discuss the extension of the above results to a weak Chebyshev system which contains a positive function.

A set of linearly independent continuous functions $\{u_i(t)\}_{i=1}^n$ form a weak Chebyshev system (M a weak Chebyshev subspace) on $[0, 1]$ provided that for all points $0 \leqq t_1 < \cdots < t_n \leqq 1$,

$$\begin{vmatrix} u_1(t_1) & \cdots & u_1(t_n) \\ u_2(t_1) & \cdots & u_2(t_n) \\ \vdots & & \vdots \\ u_n(t_1) & \cdots & u_n(t_n) \end{vmatrix} \geqq 0.$$

THEOREM 2.1. *Let M be a weak Chebyshev subspace of dimension $2l$ on $[0, 1]$, which contains a strictly positive function on $(0, 1)$. Then every positive measure relative to M has a lower principal representation.*

Proof. The basic idea of the proof is to "smooth" the weak Chebyshev system $\{u_i(t)\}_{i=1}^{2l}$ into a Chebyshev system and then apply the previous results for Chebyshev systems. Specifically, let $\delta > 0$ and define

$$u_i(t; \delta) = \frac{1}{\sqrt{2\pi\delta}} \int_0^1 e^{-(x-t)^2/2\delta} u_i(x)\, dx.$$

Then $\{u_i(t; \delta)\}_{i=1}^{2l}$ is a Chebyshev system [7], and $\lim_{\delta \to 0^+} u_i(t; \delta) = u_i(t)$ uniformly in any closed subinterval of $(0, 1)$ and $\lim_{\delta \to 0^+} u_i(t; \delta) = u_i(t)/2$, for $t \in \{0, 1\}$, $i = 1, \cdots, 2l$. Thus for every positive measure $d\alpha$, there exists a lower principal representation $d\underline{\alpha}_\delta$ such that for $i = 1, \cdots, 2l$,

$$\int_0^1 u_i(t; \delta)\, d\alpha(t) = \int_0^1 u_i(t; \delta)\, d\underline{\alpha}_\delta(t) = \sum_{j=1}^{l} \lambda_j(\delta) u_i(t_j(\delta); \delta), \tag{2.2}$$

where $\lambda_j(\delta) > 0, j = 1, \cdots, l$, and $0 < t_1(\delta) < \cdots < t_l(\delta) < 1$. (We may assume that

$$\lim_{\delta \to 0^+} \int_0^1 u_i(t; \delta)\, d\underline{\alpha}_\delta(t) = \int_0^1 u_i(t)\, d\alpha(t), \qquad i = 1, \cdots, 2l. \tag{2.3}$$

Otherwise, we construct another measure $d\hat{\alpha}(t)$, positive relative to M, such that

$$\lim_{\delta \to 0^+} \int_0^1 u_i(t; \delta)\, d\hat{\alpha}(t) = \int_0^1 u_i(t)\, d\alpha(t), \qquad i = 1, \cdots, 2l,$$

and let $d\underline{\hat{\alpha}}_\delta(t)$ be the lower principal representation of $d\hat{\alpha}(t)$. Then clearly (2.3) is valid.) Construct the "polynomial" $u(t; \delta) = \sum_{i=1}^{2l} a_i(\delta) u_i(t; \delta)$ which satisfies $u(t_i(\delta); \delta) = 0$, $i = 1, \cdots, l$, $u(t; \delta) \geqq 0$ for $t \geqq t_1(\delta)$, and $\sum_{i=1}^{2l} [a_i(\delta)]^2 = 1$. Since $\{u_i(t; \delta)\}_{i=1}^{2l}$ is a Chebyshev system on $[0, 1]$, the above conditions uniquely determine $u(t; \delta)$. Hence from (2.2), we conclude that $\int_0^1 u(t; \delta)\, d\alpha(t) = 0$ for all $\delta > 0$. Since $0 < t_1(\delta) < \cdots < t_l(\delta) < 1$, there exists a subsequence $\{\delta_k\}_{k=1}^{\infty}$, $\delta_k \downarrow 0$, such that $t_i(\delta_k) \to t_i$, $i = 1, \cdots, l$, where $0 \leqq t_1 \leqq \cdots \leqq t_l \leqq 1$. Choosing a subsequence of $\{\delta_k\}_{k=1}^{\infty}$, if necessary, it follows that $u(t; \delta_k) \to u(t)$ as $k \uparrow \infty$, uniformly in any closed subinterval of $(0, 1)$. Thus $\int_0^1 u(t)\, d\alpha(t) = 0$, and $u(t_i) = 0$, $i = 1, \cdots, l$,

while $u(t) \geqq 0$ for $t \geqq t_1$, and $u(t) \not\equiv 0$. Since $d\alpha$ is a positive measure relative to M, $t_1 > 0$. In a similar manner, one proves that $t_l < 1$. Furthermore, using the fact that M contains a function which positive on (0, 1), it follows that $\lambda_1(\delta), \cdots, \lambda_l(\delta)$ are uniformly bounded in δ. Now we may easily pass to the limit in (2.2), perhaps through a subsequence, and obtain a limit $d\underline{\alpha}(t)$ for $d\underline{\alpha}_\delta(t)$, where

$$\int_0^1 u_i(t)\, d\alpha(t) = \int_0^1 u_i(t)\, d\underline{\alpha}(t), \qquad i = 1, \cdots, 2l.$$

The proof of Theorem 2.1 will be complete if we can show that $I(d\underline{\alpha}) = l$. From the above analysis, $I(d\underline{\alpha}) \leqq l$. If strict inequality holds, then we construct by smoothing, a nonnegative nontrivial polynomial in $\{u_i(t)\}_{i=1}^{2l}$ which vanishes at the points of increase of $d\underline{\alpha}$. This contradicts the positivity of the measure $d\alpha$. Thus $I(d\underline{\alpha}) = l$, and Theorem 2.1 is proven.

Under the stronger assumption that M contains a strictly positive function on the *closed* interval [0, 1], we may prove the following result.

THEOREM 2.2. *Let M be a weak Chebyshev subspace on* [0, 1] *which contains a function which is strictly positive in* [0, 1]. *Then every measure which is positive relative to M has an upper and lower representation.*

Remark 2.1. Theorem 2.2 was proven in [1] under the stronger hypothesis that M is a weak Chebyshev subspace on some closed interval strictly containing [0, 1].

Remark 2.2. Theorems 2.1 and 2.2 are not valid if there does not exist a positive function within the weak Chebyshev subspace. For example, consider the two-dimensional weak Chebyshev subspace composed of cubic polynomials which have a double zero at $\frac{1}{2}$. On [0, 1] the positive measure dt has no lower principal representation relative to this subspace.

LEMMA 2.1. *Let M be a weak Chebyshev subspace of dimension n on* [0, 1] *and let $f \in K_0 - M$. If $d\underline{\alpha}$ is a lower principal representation of some measure which is positive relative to M, then there exists a nontrivial nonnegative function $g(t) = c_0 f(t) + \sum_{i=1}^{n} c_i u_i(t)$ such that $\int_0^1 g(t)\, d\underline{\alpha}(t) = 0$, and for any $g(t)$ of the above form, $c_0 > 0$. If we replace $d\underline{\alpha}(t)$ by an upper principal representation $d\bar{\alpha}(t)$, then the above holds with $c_0 < 0$.*

Proof. If M is a Chebyshev subspace, then the proof of Lemma 2.1 may be found in [7]. For M as above, the existence of a $g(t)$ as indicated in the statement of the lemma follows by smoothing as in the proof of Theorem 2.1. Thus for $d\underline{\alpha}(t)$, $c_0 \geqq 0$, and for $d\bar{\alpha}(t)$, $c_0 \leqq 0$. However, if in either case $c_0 = 0$, then we contradict the positivity of $d\underline{\alpha}$, or $d\bar{\alpha}$ relative to M.

COROLLARY 2.1. *Let M be a weak Chebyshev subspace of dimension n on* [0, 1]. *If $d\underline{\alpha}(t)$ and $d\bar{\alpha}(t)$ are lower and upper principal representations of a positive measure $d\alpha(t)$ relative to M, then*

$$\int_0^1 f(t)\, d\underline{\alpha}(t) \leqq \int_0^1 f(t)\, d\alpha(t) \leqq \int_0^1 f(t)\, d\bar{\alpha}(t),$$

for any $f \in K_0$.

Proof. We may assume without loss of generality that $f \in K_0 - M$. Then from Lemma 2.1 we conclude that there exists a nontrivial nonnegative function

$g(t) = c_0 f(t) + \sum_{j=1}^{n} c_j u_j(t)$ with $c_0 > 0$ such that $\int_0^1 g(t)\, d\underline{\alpha}(t) = 0$. Hence

$$0 \leqq \int_0^1 g(t)\, d\alpha(t) = \int_0^1 g(t)\, d\alpha(t) - \int_0^1 g(t)\, d\underline{\alpha}(t)$$
$$= c_0\left(\int_0^1 f(t)\, d\alpha(t) - \int_0^1 f(t)\, d\underline{\alpha}(t)\right).$$

This proves the lower inequality. The upper inequality is similarly proven.

Remark 2.3. When $d\alpha(t)$ satisfies the hypothesis of Corollary 2.1 and is also a positive measure relative to the subspace M_f which is spanned by the functions $\{f, u_1, \cdots, u_n\}$, then strict inequality holds in the above inequality whenever $f \in K_0 - M$.

We will denote the smallest linear subspace containing K_0 by $[K_0]$.

The following useful corollary appears in [1] in a weaker form.

COROLLARY 2.2. *Let M be a weak Chebyshev subspace of dimension n. If $[K_0]$ contains an n-dimensional Chebyshev system on $[0, 1]$, then every positive measure relative to M has at most one upper and one lower principal representation.*

Proof. Suppose $d\bar{\alpha}_1$ and $d\bar{\alpha}_2$ are two upper principal representations for $d\alpha(t)$. Then according to Corollary 2.1, $\int_0^1 f(t)\, d\bar{\alpha}_1(t) = \int_0^1 f(t)\, d\bar{\alpha}_2(t)$ for all $f \in K_0$. Since $[K_0]$ contains a Chebyshev system of dimension n, and $I(d\bar{\alpha}_1) = I(d\bar{\alpha}_2) = n/2$, we conclude that $d\bar{\alpha}_1 = d\bar{\alpha}_2$.

Chebyshev systems have the property that for any points $0 \leqq x_1 < \cdots < x_n \leqq 1$ and any data $y_1, y_2, \cdots, y_n$ there exists a unique $u \in M$ satisfying

$$u(x_i) = y_i, \qquad i = 1, 2, \cdots, n. \tag{2.4}$$

For a weak Chebyshev system, the determinant of the linear system (2.4) may be zero. However there does exist at least one set of points in $[0, 1]$ for which (2.4) has a unique solution. We will show that the support of the principal representations of positive measures has this property under the assumptions of Corollary 2.2. To explain this further let us suppose for the moment that $n = 2l$, $M \subseteq C^1[0, 1]$, and $d\alpha$ is a positive measure with lower principal representation $d\underline{\alpha}$. Then

$$\int_0^1 f(t)\, d\alpha(t) = \sum_{j=1}^{l} \lambda_j f(t_j), \qquad f \in M,$$

where $\lambda_1 > 0, \lambda_2 > 0, \cdots, \lambda_l > 0, 0 < t_1 < \cdots < t_l < 1$.

We associate with $d\underline{\alpha}$ the interpolation problem

$$u(t_i) = y_i^0, \qquad i = 1, 2, \cdots, l,$$
$$u'(t_i) = y_i^1, \qquad i = 1, 2, \cdots, l.$$

This set of equations has a unique solution for all real data $\{y_i^j\}$, $i = 1, \cdots, l$, $j = 0, 1$, provided that the homogeneous set of equations

$$u(t_i) = 0, \qquad i = 1, \cdots, l,$$
$$u'(t_i) = 0, \qquad i = 1, \cdots, l, \tag{2.5}$$

has only the trivial solution in M. We will denote (2.5) simply by $u(d\underline{\alpha}) = 0$. In general, for any discrete measure, we interpret $u(d\beta) = 0$ as the interpolation

problem which sets $u(t)$ and its first derivative equal to zero at an interior point of the support of $d\beta$ while at an endpoint of $[0, 1]$ only the value of $u(t)$ is set equal to zero.

COROLLARY 2.3. *Suppose M is a weak Chebyshev system contained in $C^1[0, 1]$, and $[K_0] \cap C^1[0, 1]$ contains an n-dimensional Chebyshev system on $[0, 1]$. If for all $f \in K_0$, $d\alpha$ is a positive measure for the subspace M_f, then $u(d\beta) = 0$ has only the zero solution in M, if $d\beta$ is a principal representation for $d\alpha$.*

Proof. Let us again restrict ourselves to the case when $n = 2l$ and to the interpolation problem corresponding to the lower principal representation. Thus we are required to show that the only solution to the system of equations

$$\begin{aligned} &\sum_{i=1}^{2l} a_i u_i(t_j) = 0, && j = 1, 2, \cdots, l, \\ &\sum_{i=1}^{2l} a_i u_i'(t_j) = 0, && j = 1, 2, \cdots, l, \end{aligned} \tag{2.6}$$

is the zero solution. Suppose to the contrary that (2.6) has a nontrivial solution. Then we conclude that there exist constants c_i^0, c_i^1, $i = 1, 2, \cdots, l$, not all zero, such that

$$F(u) \equiv \sum_{i=1}^{l} c_i^0 u(t_i) + \sum_{i=1}^{l} c_i^1 u'(t_i) = 0,$$

for all $u \in M$. From our hypothesis, there exists an $f_0 \in K \cap C^1[0, 1]$ such that $F(f_0) \neq 0$. Choose a constant c such that

$$\int_0^1 v \, d\alpha = \int_0^1 v \, d\underline{\alpha} + cF(v), \qquad v \in M_{f_0}.$$

We arrive at a contradiction, as before, by constructing a nontrivial nonnegative $v_0 \in M_{f_0}$ which vanishes on the support of $d\underline{\alpha}$ and necessarily has the property that $F(v_0) = 0$. Thus we have contradicted the fact that $d\alpha$ is a positive measure for M_{f_0}.

Remark 2.4. Let $w(t) > 0$, $t \in [0, 1]$. Then the measure $d\alpha(t) = w(t)\, dt$ is a positive measure for all subspaces M_f, $f \in K_0$.

Corollary 2.3 enables us to treat the question of one-sided approximation by weak Chebyshev systems. Let us consider the minimum problem

$$\min_{\substack{u \leqq f \\ u \in M}} \int_0^1 (f(t) - u(t))\, d\alpha(t). \tag{2.7}$$

COROLLARY 2.4. *Let the hypothesis of Corollary* 2.3 *hold and suppose $d\underline{\alpha}$ is a lower principal representation for $d\alpha$. Then every $f \in K \cap C^1[0, 1]$ has a unique best one-sided approximation from below. The best approximation u_0 to f is determined uniquely by the interpolation conditions $(u_0 - f)(d\underline{\alpha}) = 0$.*

Proof. Let u_0 be uniquely determined by the conditions $(u_0-f)(d\underline{\alpha})=0$. From Lemma 2.1 and Corollary 2.3, we conclude that $u_0 \leqq f$. Now let u be any element in M such that $u \leqq f$. Then

$$\begin{aligned}\int_0^1 u(t)\,d\alpha(t) &= \int_0^1 u(t)\,d\underline{\alpha}(t) \leqq \int_0^1 f(t)\,d\underline{\alpha}(t) \\ &= \int_0^1 u_0(t)\,d\underline{\alpha}(t) = \int_0^1 u_0(t)\,d\alpha(t).\end{aligned} \tag{2.8}$$

Thus u_0 is a best one-sided approximation to f from below. Furthermore, if u^0 is any other best one-sided approximation to f from below, then according to (2.8) we have $\int_0^1 (f(t)-u^0(t))\,d\underline{\alpha}(t)=0$. Thus $(f-u^0)(d\underline{\alpha})=0$, and from Corollary 2.3 we conclude that $u^0 \equiv u_0$.

Remark 2.5. The unique one-sided approximation from above for $f \in K \cap C^1[0,1]$ is determined by the interpolation conditions $(f-u_0)(d\bar{\alpha})=0$, if $d\bar{\alpha}$ exists.

We end this section with some remarks concerning weak Chebyshev systems which satisfy linear constraints. This will enable us to conveniently apply the above results to certain classes of spline functions.

Given linear functionals $L_1(u), \cdots, L_k(u)$ defined on the linear subspace M spanned by the functions $\{u_i\}_{i=1}^{n+r}$, we denote by $M(L)$ the subspace of functions in M which satisfy the linear constraints $L_i(u)=0$, $i=1,2,\cdots,k$. We may construct a basis for $M(L)$ in the following way. We define the $(k+s)$th order determinants

$$U\begin{pmatrix} i_1, \cdots, i_k, i_{k+1}, \cdots, i_{k+s} \\ L_1, \cdots, L_k, x_1, \cdots, x_s \end{pmatrix}$$

$$= \begin{vmatrix} L_1(u_{i_1}) & \cdots & L_k(u_{i_1}) & u_{i_1}(x_1) & \cdots & u_{i_1}(x_s) \\ \vdots & & \vdots & \vdots & & \vdots \\ L_1(u_{i_{k+s}}) & \cdots & L_k(u_{i_{k+s}}) & u_{i_{k+s}}(x_1) & \cdots & u_{i_{k+s}}(x_s) \end{vmatrix}$$

for $1 \leqq i_1 < \cdots < i_{k+s} \leqq n+r$ and $0 \leqq x_1 < \cdots < x_s \leqq 1$. If the set of linear functionals $\{L_i\}_{i=1}^k$ is independent over M, that is, rank $\|L_i(u_j)\|_{i=1\ j=1}^{k\ \ n+r} = k$, then there exists exist $i_1 < \cdots < i_k$ such that

$$d = U\begin{pmatrix} i_1, \cdots, i_k \\ L_1, \cdots, L_k \end{pmatrix} \neq 0,$$

and the functions

$$v_{i_l'}(t) = U\begin{pmatrix} i_1, \cdots, i_k i_l' \\ L_1, \cdots, L_k, t \end{pmatrix}, \qquad l = 1, 2, \cdots, n+r-k,$$

where $\{i_1', \cdots, i_{n+r-k}'\}$ are the set of complementary ordered indices to $\{i_1, \cdots, i_k\}$ in $\{1, 2, \cdots, n+r\}$, form a basis for $M(L)$. Furthermore, employing

Sylvester's determinant identity, see Karlin [2], we have for some $\sigma = \pm 1$

$$\begin{vmatrix} v_{i_1'}(x_1) & \cdots & v_{i_1'}(x_{n+r-k}) \\ v_{i_2'}(x_1) & \cdots & v_{i_2'}(x_{n+r-k}) \\ \vdots & & \vdots \\ v_{i_{n+r-k}'}(x_1) & \cdots & v_{i_{n+r-k}'}(x_{n+r-k}) \end{vmatrix} = \sigma d^{n+r-k-1} U\begin{pmatrix} 1, 2, \cdots, k, k+1, \cdots, n+r \\ L_1, \cdots, L_k, x_1, \cdots, x_{n+r-k} \end{pmatrix}.$$

Thus if $M(L)$ has dimension $n+r-k$, rank $\|L_i(u_j)\| = k$ and

$$\sigma U\begin{pmatrix} 1, \cdots, k, k+1, \cdots, n+r \\ L_1, \cdots, L_k, x_1, \cdots, x_{n+r-k} \end{pmatrix} \geqq 0, \tag{2.9}$$

for all $0 \leqq x_1 < \cdots < x_{n+r-k} \leqq 1$, then the set of functions $\{v_l\}_{l=1}^{n+r-k}$ form a weak Chebyshev system on [0, 1].

Furthermore, let us note that if there exists a set of points for which strict inequality holds in (2.9), then it follows that the rank $\|(L_i(u_j)\| = k$ and the dimension of $M(L)$ is $n+r-k$.

Let f be a function defined on [0, 1] such that

$$\sigma \begin{vmatrix} L_1(u_1) & \cdots & L_1(u_{n+r}) & L_1(f) \\ \vdots & & \vdots & \vdots \\ L_k(u_1) & \cdots & L_k(u_{n+r}) & L_k(f) \\ u_1(x_1) & \cdots & u_{n+r}(x_1) & f(x_1) \\ \vdots & & \vdots & \vdots \\ u_1(x_{n+r-k+1}) & \cdots & u_{n+r}(x_{n+r-k+1}) & f(x_{n+r-k+1}) \end{vmatrix} \geqq 0, \tag{2.10}$$

for $0 \leqq x_1 < \cdots < x_{n+r-k+1} \leqq 1$. Then according to Sylvester's determinant identity, we may express the determinant in (2.10) as

$$d^{n+r-k} \begin{vmatrix} v_{i_1'}(x_1) & \cdots & v_{i_1'}(x_{n+r-k-1}) \\ v_{i_2'}(x_1) & \cdots & v_{i_2'}(x_{n+r-k+1}) \\ \vdots & & \vdots \\ v_{i_{n+r-k}'}(x_1) & \cdots & v_{i_{n+r-k}'}(x_{n+r-k+1}) \\ \bar{f}(x_1) & \cdots & \bar{f}(x_{n+r-k+1}) \end{vmatrix}$$

where

$$\bar{f}(t) = \begin{vmatrix} L_1(u_{i_1}) & \cdots & L_1(u_{i_k}) & L_1(f) \\ \vdots & & \vdots & \vdots \\ L_k(u_{i_1}) & \cdots & L_k(u_{i_k}) & L_k(f) \\ u_{i_1}(t) & \cdots & u_{i_k}(t) & f(t) \end{vmatrix}. \tag{2.11}$$

Thus $\bar{f}$ is in the convexity cone of $M(L)$ and $L_i(\bar{f}) = 0$, $i = 1, \cdots, k$. Among all functions which satisfy these relations, (2.11) gives us a correspondence between

the elements in the cone of $M(L)$ and functions for which (2.10) is valid.

Finally, observe that we may expand the determinant in (2.11) by the last column and express $\bar{f}(t)$ in the form

$$\bar{f}(t) = df(t) + \sum_{j=1}^{k} a_j(t) L_j(f),$$

where $a_1(t), \cdots, a_k(t)$ are elements of M.

We now consider some application of our previous results.

3. Quadrature formulae for spline functions with boundary conditions. Let Δ_r denote the partition $0 = \xi_0 < \xi_1 < \cdots < \xi_r < \xi_{r+1} = 1$ of the unit interval $[0, 1]$. The class of spline functions on $[0, 1]$ of degree $n-1$ with simple knots at Δ_r is defined by

$$\mathcal{S} = \mathcal{S}_{n-1}(\Delta_r) = \{S: S \in C^{n-2}[0, 1],\ S|_{(\xi_\nu, \xi_{\nu+1})} \in \prod_{n-1},\ \nu = 0, 1, \cdots, r\},$$

where Π_{n-1} denotes all polynomials of degree $\leqq n-1$. Every element $S \in \mathcal{S}$ has a representation of the form

$$S(t) = \sum_{i=0}^{n-1} a_i t^i + \sum_{i=1}^{r} c_i (t - \xi_i)_+^{n-1},$$

where $t_+ = \max\{0, t\}$ (we shall always assume $n \geqq 2$).

We are interested in the subclass of $\mathcal{S}$ which satisfies boundary conditions of the following form.

Let $n + r = k + m$, and define

$$C_i(f) = \sum_{j=0}^{n-1} A_{ij} f^{(j)}(0) + \sum_{j=0}^{n-1} B_{ij} f^{(j)}(1), \qquad i = 1, \cdots, k. \tag{3.1}$$

Denote by $\mathcal{S}(\mathcal{C}_k)$ the subset of $\mathcal{S}$ satisfying $C_i(S) = 0$, $i = 1, \cdots, k$, and let $C = \|C_{ij}\|_{i=1}^{k}{}_{j=0}^{2n-1}$, where

$$C_{ij} = \begin{cases} A_{ij}(-1)^{j+n+r+1}, & i = 1, \cdots, k, \quad j = 0, 1, \cdots, n-1, \\ B_{i,2n-1-j}, & i = 1, \cdots, k, \quad j = n, \cdots, 2n-1. \end{cases} \tag{3.2}$$

The following conditions on the matrix C are assumed to prevail throughout this paper.

(i) $0 \leqq k \leqq \min\{2n, n+r\}$.

(ii) *There exist* $\{i_1, \cdots, i_s, j_1, \cdots, j_{k-s}\} \subseteq \{0, 1, \cdots, 2n-1\}$, $0 \leqq i_1 < \cdots < i_s \leqq n-1 < j_1 < \cdots < j_{k-s} \leqq 2n-1$, *satisfying* $M_{\nu-1} + m \geqq \nu$, $\nu = m+1, \cdots, n$, *where* M_ν *counts the number of terms in* $\{i_1, \cdots, i_s, 2n-1-j_1, \cdots, 2n-1-j_{k-s}\}$ *less than or equal to* ν, *and*

$$C\begin{pmatrix} 1, & \cdots, & k \\ i_1, \cdots, i_s, & j_1, \cdots, & j_{k-s} \end{pmatrix} \neq 0. \tag{3.3}$$

(iii) *For all* $\{i_1, \cdots, i_s, j_1, \cdots, j_{k-s}\}$ *satisfying* (ii),

$$C\begin{pmatrix} 1, & \cdots, & k \\ i_1, \cdots, i_s, & j_1, \cdots, & j_{k-s} \end{pmatrix},$$

is of one fixed sign.

Remark 3.1. Note that we make no assumptions on the $k \times k$ minors of C for which $M_{\nu-1}+m \geqq \nu$, $\nu = m+1, \cdots, n$ does not hold.

The main theorem of this section is the following result.

THEOREM 3.1. *Given a positive weight function $w(t)>0$ and nonnegative integers n, r, k, l with $n \geqq 2$, and $n+r=k+2l$, then there exists a quadrature formula of the form*

$$\int_0^1 f(t)w(t)\,dt = \sum_{j=1}^{k} c_j C_j(f) + \sum_{j=1}^{l} \lambda_j f(t_j), \tag{3.4}$$

which is exact for all $s \in \mathcal{S}$, where $\lambda_j > 0$, $j = 1, \cdots, l$, and $0 < t_1 < \cdots < t_l < 1$.

We remark that the formula appearing above is of "double precision" since the dimension of $\mathcal{S}$ is $n+r$ while the number of "free" parameters appearing on the right hand side of (3.4) is $k+2l$.

In general, the above quadrature formula is not unique as the following example demonstrates.

Example 3.1. Let $n=2$ and $r=3$ with the knots chosen at $\xi_1=\frac{1}{3}$, $\xi_2=\frac{1}{2}$, $\xi_3=\frac{2}{3}$, and $l=2$ and $k=1$ where the boundary condition is $S(0)+S(1)=0$. This boundary condition satisfies (3.3) and the following two quadrature formulae hold for $f \in \{1, t, (t-\frac{1}{3})_+^1, (t-\frac{1}{2})_+^1, (t-\frac{2}{3})_+^1\}$:

$$\int_0^1 f(t)\,dt = \frac{1}{6}(f(0)+f(1)) + \frac{1}{4}f\left(\frac{1}{3}\right) + \frac{5}{12}f\left(\frac{3}{5}\right),$$

$$\int_0^1 f(t)\,dt = \frac{1}{6}(f(0)+f(1)) + \frac{5}{12}f\left(\frac{2}{5}\right) + \frac{1}{4}f\left(\frac{2}{3}\right).$$

Whether uniqueness persists for all $n \geqq 3$ remains unresolved. However, we will later give some partial results on the uniqueness of (3.4).

The main idea in the proof of Theorem 3.1 is simply to show that the subspace $\mathcal{S}(\mathcal{C}_k)$ has a lower principal representation. The remainder of the section is devoted to the details of the proof of this fact.

Let us write

$$u_i(t) = t^{i-1}, \qquad i = 1, \cdots, n,$$

and

$$u_{n+i}(t) = (t-\xi_i)_+^{n-1}, \qquad i = 1, \cdots, r.$$

Then in the notation of § 2, Melkman [9] (see also [5]) proved.

THEOREM 3.2. *If $n+r=k+m$ and* (3.3) *is valid, then*

$$U\begin{pmatrix} 1, & \cdots, & n+r \\ C_1, \cdots, C_k, & x_1, \cdots, & x_m \end{pmatrix}\sigma \geqq 0, \tag{3.5}$$

where $\sigma = +1$ or -1 fixed, for all choices of $0 < x_1 \leqq \cdots \leqq x_m < 1$ (where at most n of the x_i's coincide), and (3.5) *is strictly positive iff there exists an $\{s, k-s\}$ for which* (ii) *of* (3.3) *is satisfied and*

$$\xi_{\mu+s-n} < x_\mu < \xi_{\mu+s}, \qquad \mu = 1, \cdots, m \tag{3.6}$$

(whenever the inequalities are meaningful).

Thus according to the discussion at the end of § 2, we conclude that $\mathscr{S}(\mathscr{C}_k)$ is a weak Chebyshev subspace of dimension m on $[0, 1]$.

We list below some examples of boundary conditions which satisfy (3.3).

Example 3.2.

$$S^{(i_\mu)}(0) = 0, \qquad \mu = 1, \cdots, p,$$

$$S^{(j_\mu)}(1) = 0, \qquad \mu = 1, \cdots, q,$$

where $p+q=k$, $0 \leqq i_1 < \cdots < i_p \leqq n-1$, $0 \leqq j_1 < \cdots < j_q \leqq n-1$, and $M'_{\nu-1} + m \geqq \nu$, $\nu = m+1, \cdots, n$, where M'_ν counts the number of terms in $\{i_1, \cdots, i_p, j_1, \cdots, j_q\}$ less than or equal to ν.

Example 3.3. $S^{(i)}(0) = S^{(i)}(1)$, $i = 0, 1, \cdots, k-1$, if $k+n+r$ is odd.

Example 3.4. $S^{(i)}(0) = -S^{(i)}(1)$, $i = 0, 1, \cdots, k-1$, if $k+n+r$ is even.

Example 3.5. Separated boundary conditions. Let

$$A_i(f) = \sum_{j=0}^{n-1} A_{ij} f^{(j)}(0) = 0, \qquad i = 1, \cdots, p,$$

$$B_i(f) = \sum_{j=0}^{n-1} B_{ij} f^{(j)}(1) = 0, \qquad i = 1, \cdots, q,$$

where $p+q=k$. It may be easily seen (see [5]) that these boundary conditions satisfy (3.3) provided that

(i) $0 \leqq p, q \leqq n$;

(ii) *there exist* $\{i_1, \cdots, i_p\}$, $\{j_1, \cdots, j_q\} \subseteq \{0, 1, \cdots, n-1\}$ *satisfying* $M'_{\nu-1} + m \geqq \nu$, $\nu = m+1, \cdots, n$, *where* M'_ν *counts the number of terms in* $\{i_1, \cdots, i_p, j_1, \cdots, j_q\}$ *less than or equal to* ν *and*

$$\tilde{A}\begin{pmatrix} 1, \cdots, p \\ i_1, \cdots, i_p \end{pmatrix} B\begin{pmatrix} 1, \cdots, q \\ j_1, \cdots, j_q \end{pmatrix} \neq 0, \tag{3.7}$$

where $\tilde{A} = \|A_{ij}(-1)^j\|_{i=1}^{p} {}_{j=0}^{n-1}$, $B = \|B_{ij}\|_{i=1}^{q} {}_{j=0}^{n-1}$;

(iii) *for all* $\{i_1, \cdots, i_p\}$, $\{j_1, \cdots, j_q\}$ *satisfying* (ii),

$$\tilde{A}\begin{pmatrix} 1, \cdots, p \\ i_1, \cdots, i_p \end{pmatrix} B\begin{pmatrix} 1, \cdots, q \\ j_1, \cdots, j_q \end{pmatrix},$$

is of one fixed sign.

Note that Example 3.5 includes Example 3.2.

Returning to the general case (3.1), let T denote the set of integers s for which (ii) of (3.3) is satisfied. Then we have the following interesting corollary of Theorem 3.2.

COROLLARY 3.1. *If there exists an* $s \in T$ *for which* $\min\{s, k-s\} \geqq r$, *then* $\mathscr{S}(\mathscr{C}_k)$ *has a basis of m functions which form a Chebyshev system on* $(0, 1)$.

It may also be shown that if $s \in T$ is such that $\min\{s, k-s\} \leqq r$, then $M_{\nu-1} + m \geqq \nu$, $\nu = m+1, \cdots, n$ for all $\{i_1, \cdots, i_s, j_1, \cdots, j_{k-s}\}$ satisfying $0 \leqq i_1 < \cdots < i_s \leqq n-1 < j_1 < \cdots < j_{k-s} \leqq 2n-1$. Note that when the boundary conditions are separated (Example 3.5), then $T = \{p\}$.

In the discussion which follows, we set $m = 2l$. Since we wish to prove the existence of a lower principal representation for any positive measure $d\alpha(t)$ relative to $\mathscr{S}(\mathscr{C}_k)$, we shall assume $l > 0$ (i.e., $k < n+r$). If $l = 0$, Theorem 3.1 is

easily proven. To apply Theorem 2.1, we must show that there exists a function in $\mathscr{S}(\mathscr{C}_k)$ which is positive on $(0, 1)$. However, this is not always possible. To circumvent this difficulty, we introduce the following notion of a zero of degree α for the subspace $\mathscr{S}(\mathscr{C}_k)$.

DEFINITION 3.1. If $S \in \mathscr{S}(\mathscr{C}_k)$ implies $S(0) = S'(0) = \cdots = S^{(\alpha-1)}(0) = 0$, while there exists an $S \in \mathscr{S}(\mathscr{C}_k)$ for which $S^{(\alpha)}(0) \neq 0$, then we say that $\mathscr{S}(\mathscr{C}_k)$ has a zero of degree α at 0. If there exists no such α, i.e., $S^{(i)}(0) = 0$, $i = 0, 1, \cdots, n-1$, then we say that $\mathscr{S}(\mathscr{C}_k)$ has a zero of degree n at 0. Similarly we define the degree of the zero of $\mathscr{S}(\mathscr{C}_k)$ at 1.

The following result in the case of separated boundary conditions is to be found in [12]. The proof of the general case below is essentially the same as the proof in [12]. We include it here for completeness.

PROPOSITION 3.1. *For $k < n + r$, $\mathscr{S}(\mathscr{C}_k)$ has a zero of degree α at 0 iff for all $\{i_1, \cdots, i_s, j_1, \cdots, j_{k-s}\}$ satisfying* (ii) *of* (3.3), $i_1 = 0$, $i_2 = 1, \cdots, i_\alpha = \alpha - 1$. *A similar result holds at* 1.

Proof. Assume $\mathscr{S}(\mathscr{C}_k)$ has a zero of degree α at zero, $\alpha > 0$. Assume, as well, that for all $\{i_1, \cdots, i_s, j_1, \cdots, j_{k-s}\}$ satisfying (ii) of (3.3), $i_1 = 0, \cdots, i_\gamma = \gamma - 1$, but that there exists an $\{i_1, \cdots, i_s, j_1, \cdots, j_{k-s}\}$ satisfying (ii) of (3.3) for which $i_{\gamma+1} \neq \gamma$, $0 \leqq \gamma < \alpha$. Consider the matrix $\bar{C} = \|\bar{C}_{ij}\|_{i=1\ j=0}^{k+1\ 2n-1}$, where

$$\bar{C}_{ij} = \begin{cases} C_{ij}, & i = 1, \cdots, k; \quad j = 0, 1, \cdots, 2n-1, \\ \delta_{\gamma j}, & i = k+1; \quad j = 0, 1, \cdots, 2n-1. \end{cases}$$

It is easily shown, since $i_1 = 0, \cdots, i_\gamma = \gamma - 1$ for all $\{i_1, \cdots, i_s, j_1, \cdots, j_{k-s}\}$ satisfying (ii) of (3.3), that $\bar{C}$ satisfies (3.3) unless $k = 2n$. However, if $k = 2n$, then the proposition is immediate. Let $\mathscr{S}(\bar{\mathscr{C}}_{k+1})$ denote the subset of $\mathscr{S}$ satisfying the boundary conditions associated with the matrix $\bar{C}$. Since $\mathscr{S}(\bar{\mathscr{C}}_{k+1})$ satisfies (3.3), $\mathscr{S}(\bar{\mathscr{C}}_{k+1})$ is a weak Chebyshev subspace of dimension $2l-1$ on $(0, 1)$ (recall that $n + r = k + 2l$). However, every $S \in \mathscr{S}(\mathscr{C}_k)$ satisfies $S^{(\gamma)}(0) = 0$ since $\gamma < \alpha$, and thus $\mathscr{S}(\mathscr{C}_k) \subseteq \mathscr{S}(\bar{\mathscr{C}}_{k+1})$. $\mathscr{S}(\mathscr{C}_k)$ is a subspace of dimension $2l$, and a contradiction follows.

Now let us assume that $\mathscr{S}(\mathscr{C}_k)$ has a zero of degree α at 0, and for all $\{i_1, \cdots, i_s, j_1, \cdots, j_{k-s}\}$ satisfying (ii) of (3.3), we have $i_1 = 0, \cdots, i_\alpha = \alpha - 1$, and $i_{\alpha+1} = \alpha$. Construct $\bar{C}$ as above with $\gamma = \alpha$. Then (ii) of (3.3) is not satisfied, and from the analysis of Theorem 3.2 (see [5]) the determinant associated with the conditions $S \in \mathscr{S}(\bar{\mathscr{C}}_{k+1})$, $S(x_i) = 0$, $i = 1, \cdots, 2l-1$, is singular for every choice of $\{x_i\}_{i=1}^{2l-1}$ in $(0, 1)$.

Now $\mathscr{S}(\bar{\mathscr{C}}_{k+1}) \subseteq \mathscr{S}(\mathscr{C}_k)$ and there exists an $S \in \mathscr{S}(\mathscr{C}_k)$ for which $S^{(\alpha)}(0) \neq 0$. Thus $\mathscr{S}(\bar{\mathscr{C}}_{k+1})$ has dimension at most $2l-1$. Since we may choose a basis $\{S_j(x)\}_{j=1}^{2l}$ for $\mathscr{S}(\mathscr{C}_k)$ such that $S_j^{(\alpha)}(0) = 0$, $j = 1, \cdots, 2l-1$, $\mathscr{S}(\bar{\mathscr{C}}_{k+1})$ has dimension $2l-1$. The $\{S_j(x)\}_{j=1}^{2l-1}$ are linearly independent functions and thus there exist points $\{y_i\}_{i=1}^{2l-1}$, $0 < y_1 < \cdots < y_{2l-1} < 1$, such that any $S \in \mathscr{S}(\bar{\mathscr{C}}_{k+1})$, i.e., $S(x) = \sum_{j=1}^{2l-1} a_j S_j(x)$, satisfying $S(y_i) = 0$, $i = 1, \cdots, 2l-1$, implies $S(x) \equiv 0$. This contradicts the fact that the determinant associated with the conditions $S \in \mathscr{S}(\bar{\mathscr{C}}_{k+1})$, $S(x_i) = 0$, $i = 1, \cdots, 2l-1$, is singular for every choice of $\{x_i\}_{i=1}^{2l-1}$. The proposition is proven by applying the same analysis at 1.

From Proposition 3.1, we have

COROLLARY 3.2. *For all* $S \in \mathscr{S}(\mathscr{C}_k)$, $S(t) \equiv 0$ *in* $[0, \xi_1]$ *if and only if* $s = n$ *for all* $s \in T$, *and* $S(t) \equiv 0$ *on* $[\xi_r, 1]$ *for all* $S \in \mathscr{S}(\mathscr{C}_k)$ *if and only if* $k - s = n$ *for all* $s \in T$.

Now, let $d\alpha(t)$ be any positive measure relative to $\mathscr{S}(\mathscr{C}_k)$, and construct, as in the proof of Theorem 2.1, the points $\{t_i\}_{i=1}^{l}$, $0 < t_1 \leqq \cdots \leqq t_l < 1$ for the subspace $\mathscr{S}(\mathscr{C}_k)$. Corollary 3.2 shows that we cannot, in general, expect $\mathscr{S}(\mathscr{C}_k)$ to contain a positive function. However, in the proof of Theorem 2.1, we see that we only require the existence of a function which is positive on the set $\{t_i\}_{i=1}^{l}$ to conclude that $d\alpha$ has a lower principal representation. In our next proposition we will explore the relationship between the $\{t_i\}_{i=1}^{l}$ and the $\{\xi_i\}_{i=1}^{r}$.

Let us note that from the proof of Theorem 2.1, there exists a nontrivial $\hat{S}(t) \in \mathscr{S}(\mathscr{C}_k)$ such that

$$\hat{S}(t_i) = 0, \qquad i = 1, \cdots, l,$$

$$\hat{S}(t) \geqq 0, \qquad t \geqq t_1,$$

and

$$\int_0^1 \hat{S}(t)\, d\alpha(t) = 0.$$

Therefore, since $d\alpha(t)$ is a positive measure with respect to $\mathscr{S}(\mathscr{C}_k)$, $\hat{S}(t) < 0$ for some $t < t_1$. On the basis of this observation, we have

PROPOSITION 3.2. *If* $\mathscr{S}(\mathscr{C}_k)$ *has a zero of degree* n *at* 0, *then* $t_1 > \xi_2$, *while if* $\mathscr{S}(\mathscr{C}_k)$ *has a zero of degree* n *at* 1, *then* $t_l < \xi_{r-1}$.

Proof. If $\mathscr{S}(\mathscr{C}_k)$ has a zero of degree n at zero, then $S \in \mathscr{S}(\mathscr{C}_k)$ implies $S(t) \equiv 0$, $t \in [0, \xi_1]$. Since $S^{(i)}(\xi_1) = 0$, $i = 0, 1, \cdots, n-2$, and $S|_{(\xi_1, \xi_2)} \in \Pi_{n-1}$, if $S(t_1) = 0$ for $t_1 \leqq \xi_2$, then $S(t) \equiv 0$ for $t \leqq t_1$. However, $\hat{S}(t) < 0$ for some $t < t_1$, and thus we conclude that $t_1 > \xi_2$. By an analogous argument we obtain the corresponding result at one.

PROPOSITION 3.3. *If* $k < n + r$, *then there exists an* $S \in \mathscr{S}(\mathscr{C}_k)$ *which is strictly positive on an open interval containing* $[t_1, t_l]$.

Proof. If $k = 2n$, then $r = n + 2l$, and we may easily construct, by the use of B-splines (see [2] or § 5), a spline $S \in \mathscr{S}(\mathscr{C}_k)$ such that $S(t) > 0$, $t \in (\xi_1, \xi_r)$. The result then emanates from Proposition 3.2.

In what follows, we shall assume $n \geqq 3$. For the case $n = 2$, the required spline may be explicitly constructed.

Define, for $\varepsilon, \delta \in (0, 1)$,

$$S_1(t) = \int_{\varepsilon \leqq y_1 < \cdots \leqq y_{l-1} \leqq 1} \cdots \int U\begin{pmatrix} 1, & \cdots, & n+r \\ C_1, \cdots, C_k, \varepsilon, y_1, y_1, \cdots, y_{l-1}, y_{l-1}, t \end{pmatrix} \sigma\, dy_1 \cdots dy_{l-1},$$

$$S_2(t) = \int_{0 \leqq y_1 \leqq \cdots \leqq y_{l-1} \leqq 1-\delta} \cdots \int U\begin{pmatrix} 1, & \cdots, & n+r \\ C_1, \cdots, C_k, y_1, y_1, \cdots, y_{l-1}, y_{l-1}, t, 1-\delta \end{pmatrix} \sigma\, dy_1 \cdots dy_{l-1},$$

and

$$S(t) = S_1(t) + S_2(t).$$

From (3.5), $S_1(t) \geqq 0$ for $t \in (\varepsilon, 1)$, while $S_2(t) \geqq 0$ for $t \in (0, 1-\delta)$.

Case I. There exists an $s \in T$ such that $s, k-s < n$. Let $\varepsilon, \delta > 0$ be chosen arbitrarily small. By (3.6), if $s < n-1$, $S_1(t) > 0$ for $t \in (\varepsilon, 1)$, while if $s = n-1$. $S_1(t) > 0$ for $t \in (\xi_1, 1)$. Similarly, if $k-s < n-1$, $S_2(t) > 0$ for $t \in (0, 1-\delta)$, and if $k-s = n-1$, $S_2(t) > 0$ for $t \in (0, \xi_r)$. Thus it follows that $S(t) > 0$ for $t \in (\varepsilon, 1-\delta)$ for all ε, δ positive and small. Since $t_1 > 0$, $t_l < 1$, the result follows.

Case II. $\mathscr{S}(\mathscr{C}_k)$ has a zero of degree n at 0 or 1. Assume $\mathscr{S}(\mathscr{C}_k)$ has a zero of degree n at 1. Thus for $s \in T$, $k-s = n$. Since we have already considered the case $k = 2n$, we assume $k < 2n$, implying $s \leqq n-1$. Choose $\varepsilon, \delta > 0$, ε small and $\xi_{r-1} < 1-\delta < \xi_r$. If $s < n-1$, then $S_1(t) > 0$ for $t \in (\varepsilon, \xi_r)$, while if $s = n-1$, then $S_1(t) > 0$, for $t \in (\xi_1, \xi_r)$. However, $S_2(t) > 0$ for $t \in (0, \xi_{r-1})$, and the result then follows from Proposition 3.2.

Case III. $\mathscr{S}(\mathscr{C}_k)$ has no zero of degree n at 0 or 1, but for all $s \in T$, either $s = n$ or $k-s = n$.

From Corollary 3.2, it follows that since $\mathscr{S}(\mathscr{C}_k)$ has no zero of degree n at 0 or 1, there exist $s_1, s_2 \in T$ such that $s_1 = n$, $k-s_1 < n$, and $s_2 < n$, $k-s_2 = n$. Obviously, $k-s_1 = s_2$.

If $k-s_1 = s_2 < n-1$, let $\varepsilon, \delta > 0$ be chosen small. Since $s_2 < n-1$, $S_1(t) > 0$ for $t \in (\varepsilon, \xi_r)$, and since $k-s_1 < n-1$, $S_2(t) > 0$ for $t \in (\xi_1, 1-\delta)$. Thus $S(t) > 0$ for $t \in (\varepsilon, 1-\delta)$.

Assume $k-s_1 = s_2 = n-1$. By the above construction, $S(t) > 0$ for $t \in (\xi_1, \xi_r)$. We shall show that $t_1 > \xi_1$ and $t_l < \xi_r$, proving the proposition.

LEMMA 3.1. *Assume, as above, that* $k = 2n-1$ *and* $n, n-1 \in T$. *Then* $t_1 > \xi_1$ *and* $t_l < \xi_r$.

Proof. Let $\mathscr{S}'$ denote the subset of $\mathscr{S}$ satisfying $S^{(i)}(0) = S^{(i)}(1) = 0$, $i = 0, 1, \cdots, n-1$. Since $k = 2n-1$, $\mathscr{S}'$ is a subset of $\mathscr{S}(\mathscr{C}_k)$ of dimension $2l-1$ and every $S \in \mathscr{S}'$ vanishes identically on $[0, \xi_1] \cup [\xi_r, 1]$. For $\xi_1 < \varepsilon < \xi_2$, $S_1(t) < 0$ for $t < \varepsilon$ and $S_1(t) > 0$ for $t > \varepsilon$ where $S_1(t)$ is defined above. Thus $S_1 \in \mathscr{S}(\mathscr{C}_k)$, and $S_1(t) \neq 0$ for all $t \in (0, \xi_1] \cup [\xi_r, 1)$. Therefore the subset of $\mathscr{S}(\mathscr{C}_k)$ which vanishes at some point $x \in (0, \xi_1] \cup [\xi_r, 1)$ has dimension $2l-1$. However, since this subset still contains $\mathscr{S}'$, it must equal $\mathscr{S}'$. Let $\hat{S}(t)$ be as constructed in Theorem 2.1. If $t_1 \leqq \xi_1$, then by the above analysis, $\hat{S} \in \mathscr{S}'$ and thus $\hat{S}(t) \equiv 0$ for $t \leqq t_1$. This contradicts the properties of $\hat{S}$ and therefore $t_1 > \xi_1$. Similarly, $t_l < \xi_r$. The lemma is proven.

We are now ready to prove

THEOREM 3.3. *Assume* $n+r = k+2l$ *and* $d\alpha(t)$ *is a positive measure with respect to the weak Chebyshev subspace* $\mathscr{S}(\mathscr{C}_k)$ *of dimension* $2l$. *Then* $d\alpha(t)$ *has a lower principal representation.*

Proof. The proof of Theorem 3.3 follows Theorem 2.1 and Proposition 3.3.

Remark 3.2. If there exists an $s \in T$ such that $\min\{s, k-s\} \geqq r$, then from Corollary 3.1, $\mathscr{S}(\mathscr{C}_k)$ is a Chebyshev subspace on $(0, 1)$. The existence and uniqueness of the lower principal representation for $\mathscr{S}(\mathscr{C}_k)$ is immediate in this case.

We may now prove Theorem 3.1.

Proof of Theorem 3.1. From Theorem 3.3, there exists a quadrature formula of the form

$$\int_0^1 f(t)\,dt = \sum_{i=1}^{l} \lambda_i f(t_i), \tag{3.8}$$

which is exact for all $S \in \mathcal{S}(\mathcal{C}_k)$, where $\lambda_i > 0$, $i = 1, \cdots, l$ and $0 < t_1 < \cdots < t_l < 1$.

From (2.10) and the subsequent analysis, any $S \in \mathcal{S}$ may be expressed in the form

$$S(t) = d^{-1}\left[\bar{S}(t) - \sum_{j=1}^{k} c_j(t) C_j(S)\right],$$

where $d \neq 0$ and $\bar{S} \in \mathcal{S}(\mathcal{C}_k)$. Substituting this relation with $f = \bar{S}$ into (3.8), we obtain (3.4). The theorem is proven.

If the boundary conditions under consideration are separated (see Example 3.5), then the quadrature formula (3.8) and (3.4) are unique. The proof of this fact is based upon Corollary 2.2 and the following proposition.

PROPOSITION 3.4. *Assume* $n + r = p + q + 2l$, *and that the boundary conditions* (3.1) *are separated and satisfy* (3.7). *If* $f \in C^n[0, 1]$ *and* $C_i(f) = 0$, $i = 1, \cdots, p + q = k$, *then* $f \in [K]$, *where* $[K]$ *is the smallest linear subspace containing* K, *the convexity cone generated by* $\mathcal{S}(\mathcal{C}_k)$.

Proof. Any $f \in C^n[0, 1]$ may be written as $f = f_1 - f_2$, where $f_j^{(n)}(t)(-1)^i \geqq 0$ for $t \in (\xi_{i-1}, \xi_i)$, $i = 1, \cdots, r+1$; $j = 1, 2$, and $f_j \in C^{n-1}[0, 1]$, $f_j \in C^n(\xi_{i-1}, \xi_i)$, $i = 1, \cdots, r+1$. Let $g_1(t) = f_1(t) + S(t)$, where $S \in \mathcal{S}$, such that $C_i(g_1) = 0$, $i = 1, \cdots, k$. Since $C_i(f) = 0$, $i = 1, \cdots, p+q$, and $f = (f_1 + S) - (f_2 + S)$, we conclude that $C_i(f_2 + S) = 0$, $i = 1, \cdots, k$. Let $g_2(t) = f_2(t) + S(t)$.

We shall prove that for any function $g(t)$ which has the form

$$g(t) = S(t) + \frac{1}{(n-1)!}\int_0^1 (t-x)_+^{n+1} g^{(n)}(x)\,dx, \tag{3.9}$$

where $S \in \mathcal{S}$, and which satisfies $g^{(n)}(t)(-1)^i \geqq 0$, $t \in (\xi_{i-1}, \xi_i)$, $i = 1, \cdots, r+1$, and $C_i(g) = 0$, $i = 1, \cdots, k$, then either g or $-g$ lie in K. By Taylor's theorem, both $g_1(t)$ and $g_2(t)$ are of the requisite form. From the properties of $g(t)$, (3.9) may be rewritten in the form

$$g(t) = S(t) + \sum_{i=1}^{r+1} \frac{(-1)^i}{(n-1)!}\int_{\xi_{i-1}}^{\xi_i} (t-x)_+^{n+1} |g^{(n)}(x)|\,dx.$$

Let $u_{n+r+1}(t) = (t - \xi)_+^{n-1}$. An application of Theorem 3.2 and (3.7) yields

$$U\begin{pmatrix} 1, & \cdots, & n+r+1 \\ C_1, \cdots, C_k, & x_1, \cdots, & x_{2l+1} \end{pmatrix}(-1)^i \sigma \geqq 0, \tag{3.10}$$

for $\xi \in (\xi_{i-1}, \xi_i)$, $i = 1, \cdots, r+1$, where $\sigma = +1$ or -1, fixed. Now from (3.9),

$$\begin{vmatrix} C_1(u_1) & \cdots & C_1(u_{n+r}) & C_1(g) \\ \vdots & & \vdots & \vdots \\ C_k(u_1) & \cdots & C_k(u_{n+r}) & C_k(g) \\ u_1(x_1) & \cdots & u_{n+r}(x_1) & g(x_1) \\ \vdots & & \vdots & \vdots \\ u_1(x_{2l+1}) & \cdots & u_{n+r}(x_{2l+1}) & g(x_{2l+1}) \end{vmatrix}$$

$$= \sum_{i=1}^{r+1} \frac{(-1)^i}{(n-1)!} \int_{\xi_{i-1}}^{\xi_i} U\begin{pmatrix} 1, \cdots, n+r+1 \\ C_1, \cdots, C_k, x_1, \cdots, x_{2l+1} \end{pmatrix} |g^{(n)}(\xi)|\, d\xi.$$

From (3.10) it follows that the above determinant is nonnegative (or nonpositive) for all $0 < x_1 \leqq \cdots \leqq x_{2l+1} < 1$. Thus by (2.11), $\bar{g}$ (or $-\bar{g}$) is in K. Since, by assumption $C_i(g) = 0$, $i = 1, \cdots, k$, it follows that $g = \bar{g}$, and the proposition is proven.

Thus we have also proven

THEOREM 3.4. *For separated boundary conditions which satisfy* (3.7), *the quadrature formula* (3.4) *is unique.*

Remark 3.3. In the case of separated boundary conditions, it follows from Corollary 2.3 and Theorem 3.2 that

$$\xi_{2i+p-n} < t_i < \xi_{2i-1+p}, \qquad i = 1, \cdots, l,$$

where the $\{t_i\}_{i=1}^l$ are the nodes ofthe unique quadrature formula (3.4).

Remark 3.4. The analysis of this section also holds for spline functions with knots of multiplicity at most $n-2$ and for Chebyshevian spline functions (see [3] and [6]).

4. Monosplines satisfying boundary conditions. In this section, we shall study the Peano kernel of the quadrature formula of Theorem 3.1 and state our version of the fundamental theorem of algebra for monosplines.

A monospline of degree n with l knots $\{x_i\}_{i=1}^l$, $0 < x_1 < \cdots < x_l < 1$, is a function of the form

$$M(x) = \frac{x^n}{n!} + \sum_{i=0}^{n-1} a_i x^i + \sum_{i=1}^{l} b_i (x - x_i)_+^{n-1}. \tag{4.1}$$

Let $C_i(f)$, $i = 1, \cdots, k$, be boundary conditions of the form (3.1) such that the $k \times 2n$ matrix C satisfies (3.3). Let

$$Q(f) = \sum_{i=1}^{k} c_i C_i(f) + \sum_{i=1}^{l} \lambda_i f(t_i), \tag{4.2}$$

be the quadrature formula constructed in Theorem 3.1, i.e., $\int_0^1 f(x)\, dx = Q(f)$ for all $f \in \mathcal{S}$. Recall that $n + r = k + 2l$. Every $f \in C^n[0, 1]$ has the representation

$$f(t) = \sum_{i=0}^{n-1} f^{(i)}(0) \frac{t^i}{i!} + \frac{1}{(n-1)!} \int_0^1 (t-x)_+^{n-1} f^{(n)}(x)\, dx.$$

Define $R(f)=\int_0^1 f(x)\,dx-Q(f)$. Then for $f\in C^n[0,1]$,

$$R(f)=\int_0^1 R_t((t-x)_+^{n-1})f^{(n)}(x)\,dx.$$

This, of course, is the Peano representation of the remainder $R(f)$. We define

$$\frac{(-1)^n}{(n-1)!}R_t((t-x)_+^{n-1})=M^*(x),$$

and note that $M^*(x)$ is a monospline of degree n with the l knots $\{t_i\}_{i=1}^l$. Thus for all $f\in C^n[0,1]$,

$$\int_0^1 f(x)\,dx=\sum_{i=1}^k c_iC_i(f)+\sum_{i=1}^l \lambda_i f(t_i)+(-1)^n\int_0^1 M^*(x)f^{(n)}(x)\,dx. \tag{4.3}$$

Since $R_t((t-\xi_i)_+^{n-1})=0$, we conclude that $M^*(\xi_i)=0$, $i=1,\cdots,r$.

Let $M(x)$ be any monospline of the form (4.1), and $f\in C^n[0,1]$. Then integration by parts yields

$$\begin{aligned}\int_0^1 f(x)\,dx=&\sum_{i=0}^{n-1}(-1)^{i+1}f^{(i)}(0)M^{(n-1-i)}(0)\\ &+\sum_{i=0}^{n-1}(-1)^i f^{(i)}(1)M^{(n-1-i)}(1)\\ &-\sum_{i=1}^l (n-1)!b_i f(x_i)+(-1)^n\int_0^1 M(x)f^{(n)}(x)\,dx.\end{aligned} \tag{4.4}$$

Thus from (4.3) and (4.4) we obtain

LEMMA 4.1. *The monospline $M^*(x)$ defined above satisfies*

$$\begin{aligned}\sum_{i=1}^k c_iC_i(f)=&\sum_{i=0}^{n-1}(-1)^{i+1}f^{(i)}(0)(M^*)^{(n-1-i)}(0)\\ &+\sum_{i=0}^{n-1}(-1)^i f^{(i)}(1)(M^*)^{(n-1-i)}(1),\end{aligned} \tag{4.5}$$

for all $f\in C^n[0,1]$.

Since the $k\times 2n$ matrix C has rank k, we may construct a $2n\times 2n$ nonsingular matrix whose first k rows agree with C. We shall also denote this enlarged matrix by C. Define $D=(C^T)^{-1}$, and let

$$\bar{\mathbf{g}}=((-1)^{n+r+1}g(0),(-1)^{n+r+2}g'(0),\cdots,(-1)^r g^{(n-1)}(0),g^{(n-1)}(1),\cdots,g(1))$$

and

$$\hat{\mathbf{g}}=((-1)^{n+r}g^{(n-1)}(0),\cdots,(-1)^{n+r}g(0),(-1)^{n-1}g(1),(-1)^{n-2}g'(1),\cdots,g^{(n-1)}(1)).$$

Thus $(\bar{\mathbf{f}},\hat{\mathbf{M}})$ (the inner product of the vectors $\bar{\mathbf{f}}$ and $\hat{\mathbf{M}}$) represents the right-hand side of (4.5), and $(C\bar{\mathbf{f}})_i=C_i(f)$, $i=1,\cdots,k$.

LEMMA 4.2. *For $M^*(x)$ as above,*

$$(D\hat{\mathbf{M}}^*)_i=0,\qquad i=k+1,\cdots,2n.$$

Proof. The proof follows from Lemma 4.1 and the equation

$$(\bar{\mathbf{f}}, \hat{\mathbf{M}}^*) = (C\bar{\mathbf{f}}, D\hat{\mathbf{M}}^*) = \sum_{i=1}^{2n} (C\bar{\mathbf{f}})_i (D\hat{\mathbf{M}})_i.$$

THEOREM 4.1. *If $n+r=k+2l$ and $D=(C^T)^{-1}$, where the matrix composed of the first k rows of C satisfies* (3.3), *then given $\{\xi_i\}_{i=1}^r$, $0<\xi_1<\cdots<\xi_r<1$, there exists a monospline $M(x)$ of degree n with l knots such that*

$$\begin{aligned} M(\xi_i) &= 0, & i &= 1, \cdots, r, \\ (D\hat{\mathbf{M}})_i &= 0, & i &= k+1, \cdots, 2n. \end{aligned} \tag{4.6}$$

Furthermore, the knots of the monospline $M(x)$ are the nodes of the quadrature formula (3.4), *and $M(x)$ is unique if and only if the corresponding quadrature formula is unique.*

Proof. From the above analysis, every quadrature formula of the form (3.4) gives rise to a monospline $M(x)$ satisfying (4.6).

If $M(x)$ satisfies (4.6), then (4.5) holds for $f \in \mathcal{S}$. Let

$$Q^0(f) = \sum_{i=1}^{k} c_i C_i(f) - \sum_{i=1}^{l} (n-1)! b_i f(x_i)$$

and

$$R^0(f) = \int_0^1 f(x)\,dx - Q^0(f).$$

From (4.4), $R^0(f)=0$ for $f \in \{1, t, \cdots, t^{n-1}\}$, and since $R_t^0((t-x)_+^{n-1}) = M(x)$, the theorem is proven.

The following two theorems represent a partial converse to Theorem 4.1. To prove these theorems, we demand an additional assumption on the $k \times 2n$ matrix C (see Remark 3.1).

(4.7) Assume the $k \times 2n$ matrix C satisfies (3.3) and *all* nonzero $k \times k$ minors of C are of one sign.

THEOREM 4.2. *If D is as above, where C satisfies* (4.7), *and if $M(x)$ is a monospline of degree n with l knots for which $(D\hat{\mathbf{M}})_i = 0$, $i = k+1, \cdots, 2n$, then $M(x)$ has at most $r+1$ distinct zeros in* (0, 1).

Proof. Assume $M(x)$ has $r+2$ distinct zeros $\{\xi_i\}_{i=1}^{r+2}$ in (0, 1). Then there exists a quadrature formula

$$\int_0^1 f(t)\,dt = \sum_{i=1}^{k} c_i C_i(f) + \sum_{i=1}^{l} \lambda_i f(t_i),$$

which is exact for $f \in \mathcal{S}^* = \{1, t, \cdots, t^{n-1}, (t-\xi_1)_+^{n-1}, \cdots, (t-\xi_{r+2})_+^{n-1}\}$. The $k \times 2n$ matrix C satisfies (3.3) with respect to n, k, $2l$ and r. Since r and $r+2$ are of the same parity and C satisfies (4.7), it follows that $\mathcal{S}^*(\mathcal{C}_k)$ is a weak Chebyshev subspace of dimension $2l+2$, and

$$\int_0^1 f(t)\,dt = \sum_{i=1}^{l} \lambda_i f(t_i),$$

for all $f \in \mathcal{S}^*(\mathcal{C}_k)$. This is impossible since we may construct, by smoothing (see

Theorem 2.1), a nonnegative nontrivial $S \in \mathscr{S}^*(\mathscr{C}_k)$ which vanishes at the nodes $\{t_i\}_{i=1}^{l}$. The theorem is proven.

THEOREM 4.3. *Under the assumptions of Theorem* 4.2, *if the boundary conditions represented by C are separated, then* $M(x)$ *has at most r distinct zeros in* $(0, 1)$.

Proof. The proof is the same as that of Theorem 4.2, where we use (3.7) and note that for separated boundary conditions, the parity of r plays no role. Thus the addition of one more knot to $\mathscr{S}(\mathscr{C}_k)$ gives rise to a weak Chebyshev subspace of dimension $2l+1$.

Remark 4.1. The bound given in Theorem 4.2 is sharp as the following example indicates. Consider the case $n=2$ and $r=1$ with the knot $\xi=\frac{1}{3}$. Let $l=1$ and $k=1$, with the boundary condition $S(0)+S(1)=0$ which satisfies (4.7). The quadrature formula

$$\int_0^1 f(t)\,dt=\frac{1}{6}(f(0)+f(1))+\frac{2}{3}f\left(\frac{1}{2}\right),$$

holds for $f \in \{1, t, (t-\frac{1}{3})_+^1\}$ and $f(t)=(t-\frac{2}{3})_+^1$. The associated monospline

$$M(x)=\frac{x^2}{2}-\frac{x}{6}-\frac{2}{3}\left(x-\frac{1}{2}\right)_+^1,$$

satisfies $M(0)=M(1)=M'(0)+M'(1)=0$, and $M(\frac{1}{3})=M(\frac{2}{3})=0$.

Remark 4.2. As previously commented upon in Remark 3.4, Theorem 3.1 extends to spline functions with knots of multiplicity $\leqq n-2$. Thus Theorems 4.1–4.3 extend to the case of multiple zeros of order at most $n-2$ for $M(x)$.

Let $\mathscr{D}_{2n-k}$ denote the set of boundary conditions $(D\hat{\mathbf{M}})_i=0$, $i=k+1, \cdots, 2n$, where $D=(C^T)^{-1}$, and the first k rows of C satisfy (3.3). Our present goal is to present a more workable definition of $\mathscr{D}_{2n-k}$. To this end, define

$$G_i(f)=\sum_{j=0}^{n-1} E_{ij}f^{(j)}(0)+\sum_{j=0}^{n-1} F_{ij}f^{(j)}(1), \qquad i=1, \cdots, 2n-k, \tag{4.8}$$

and let $G=\|G_{ij}\|_{i=1\ j=0}^{2n-k\ 2n-1}$, where

$$G_{ij}=\begin{cases} E_{ij}(-1)^{j+n+1}, & i=1, \cdots, 2n-k;\ j=0, 1, \cdots, n-1, \\ F_{i,2n-j-1}, & i=1, \cdots, 2n-k;\ j=n, \cdots, 2n-1. \end{cases} \tag{4.9}$$

Let $\mathscr{G}_{2n-k}$ denote the set of boundary conditions $G_i(M)=0$, $i=1, \cdots, 2n-k$, where G satisfies (3.3) with $m=r$; i.e.,

(i) $\max\{0, n-r\} \leqq 2n-k \leqq 2n$,

(ii) *there exist* $\{i_1, \cdots, i_s, j_1, \cdots, j_{2n-k-s}\} \subseteq \{0, 1, \cdots, 2n-1\}$, *satisfying* $M_{\nu-1}+r \geqq \nu$, $\nu=r+1, \cdots, n$, *and*

$$G\begin{pmatrix} 1, & \cdots, & 2n-k \\ i_1, \cdots, i_s, & j_1, \cdots, & j_{2n-k-s} \end{pmatrix} \neq 0, \tag{4.10}$$

(iii) *for all* $\{i_1, \cdots, i_s, j_1, \cdots, j_{2n-k-s}\}$ *satisfying* (ii),

$$G\begin{pmatrix} 1, & \cdots, & 2n-k \\ i_1, \cdots, i_s, & j_1, \cdots, & j_{2n-k-s} \end{pmatrix},$$

is of one sign.

Then we have

PROPOSITION 4.1. $\mathscr{D}_{2n-k} = \mathscr{G}_{2n-k}$.

Proof.

$$(D\hat{\mathbf{M}})_i = \sum_{j=0}^{n-1} D_{ij} M^{(n-j-1)}(0)(-1)^{n+r} + \sum_{j=0}^{n-1} D_{i,j+n}(-1)^{n-j-1} M^{(j)}(1)$$

$$= \sum_{j=0}^{n-1} D_{i,n-j-1}(-1)^{r+j+1}[M^{(j)}(0)(-1)^{n+j+1}]$$

$$+ \sum_{j=0}^{n-1} D_{i,2n-j-1}(-1)^{j}[M^{(n-j-1)}(1)], \qquad i = k+1, \cdots, 2n.$$

Let

$$H_{ij} = \begin{cases} D_{i+k,n-j-1}(-1)^{r+j+1}, & i = 1, \cdots, 2n-k; j = 0, 1, \cdots, n-1, \\ D_{i+k,3n-j-1}(-1)^{n+j}, & i = 1, \cdots, 2n-k; j = n, \cdots, 2n-1. \end{cases}$$

Then Proposition 4.1 is valid if we can prove that $H = \|H_{ij}\|_{i=1}^{2n-k}{}_{j=0}^{2n-1}$ satisfies (4.10) if and only if C satisfies (3.3).

Let $\{i'_m\}_{m=1}^{n-s}$ denote the complementary set of indices to $\{n-i_m-1\}_{m=1}^{s}$ in $\{0, 1, \cdots, n-1\}$, and $\{j'_m\}_{m=1}^{k+s-n}$ denote the complementary set of indices to $\{3n-j_m-1\}_{m=1}^{2n-k-s}$ in $\{n, \cdots, 2n-1\}$.

The following two lemmas prove Proposition 4.1.

LEMMA 4.3. *With the above definitions,*

$$H\begin{pmatrix} 1, & \cdots, & 2n-k \\ i_1, \cdots, i_s, & j_1, \cdots, & j_{2n-k-s} \end{pmatrix} = C\begin{pmatrix} 1, & \cdots, & k \\ i'_1, \cdots, i'_{n-s}, & j'_1, \cdots, & j'_{k+s-n} \end{pmatrix}.$$

Proof.

$$H\begin{pmatrix} 1, & \cdots, & 2n-k \\ i_1, \cdots, i_s, & j_1, \cdots, & j_{2n-k-s} \end{pmatrix}$$

$$= D\begin{pmatrix} k+1, & \cdots, & 2n \\ n-i_1-1, \cdots, n-i_s-1, & 3n-j_1-1, \cdots, & 3n-j_{2n-k-s}-1 \end{pmatrix}(-1)^{\varepsilon_1}$$

$$= D\begin{pmatrix} k+1, & \cdots, & 2n \\ n-i_s-1, \cdots, n-i_1-1, & 3n-j_{2n-k-s}-1, \cdots, & 3n-j_1-1 \end{pmatrix}(-1)^{\varepsilon_1+\varepsilon_2}$$

$$= C\begin{pmatrix} 1, & \cdots, & k \\ i'_1, \cdots, i'_{n-s}, & j'_1, \cdots, & j'_{k+s-n} \end{pmatrix}(-1)^{\varepsilon_1+\varepsilon_2+\varepsilon_3},$$

where

$$\varepsilon_1 = (r+1)s + n(2n-k-s) + \sum_{m=1}^{s} i_m + \sum_{m=1}^{2n-k-s} j_m,$$

$$\varepsilon_2 = \frac{s(s-1)}{2} + \frac{(2n-k-s)(2n-k-s-1)}{2},$$

and

$$\varepsilon_3 = \frac{2n(2n+1)}{2} - \frac{k(k+1)}{2} + s(n-1) + (3n-1)(2n-k-s) - \sum_{m=1}^{s} i_m - \sum_{m=1}^{2n-k-s} j_m.$$

Utilizing the fact that $n+r-k=2l$, it follows that $(-1)^{\varepsilon_1+\varepsilon_2+\varepsilon_3}\equiv 1$. The lemma is proven.

LEMMA 4.4. *$\{i_1,\cdots,i_s,j_1,\cdots,j_{2n-k-s}\}$ satisfies $M_{\nu-1}+r\geqq\nu$, $\nu=r+1,\cdots,n$, if and only if $\{i'_1,\cdots,i'_{n-s},j'_1,\cdots,j'_{k+s-n}\}$ satisfies $M_{\nu-1}+2l\geqq\nu$, $\nu=2l+1,\cdots,n$.*

Proof. Due to the symmetry of the analysis we prove only one direction. Assume $\{i_1,\cdots,i_s,j_1,\cdots,j_{2n-k-s}\}$ is such that $M_\mu+r\leqq\mu$ for some $\mu=r,\cdots,n-1$. Note that since $M_{n-1}=2n-k$, and $2n-k+r=n+2l\geqq n$, $\mu<n-1$. Let i_γ, $2n-j_\beta-1\leqq\mu<i_{\gamma+1}$, $2n-j_{\beta-1}-1$. Thus $M_\mu+r=\gamma+(2n-k-s-\beta+1)+r\leqq\mu$, i.e., $\gamma-s-\beta+1+n+2l-\mu\leqq 0$. Now $n-i_\gamma-1$, $j_\beta-n\geqq n-\mu-1>n-i_{\gamma+1}-1$, $j_{\beta-1}-n$, and therefore

$$i'_{n-\mu-s+\gamma},\ 2n-1-j'_{\beta+k+s-2n+\mu}>n-\mu-2\geqq i'_{n-\mu-s+\gamma-1},\ 2n-1-j'_{\beta+k+s-2n+\mu+1},$$

and for $\{i'_1,\cdots,i'_{n-s},j'_1,\cdots,j'_{k+s-n}\}$,

$$\begin{aligned}M_{n-\mu-2}+2l&=2l+(n-\mu-s+\gamma-1)+(n-\beta-\mu)\\&=(\gamma-s-\beta+1+n+2l-\mu)+(n-\mu-2)\leqq n-\mu-2,\end{aligned}$$

since $\gamma-s-\beta+1+n+2l-\mu\leqq 0$. The lemma is proven.

Utilizing Proposition 4.1, Theorems 4.1–4.3 may be restated in terms of the boundary conditions (4.8), where G satisfies (4.10).

Remark 4.3. Note that from the proof of Proposition 4.1, it is easily seen that the boundary forms (4.8) are separated if and only if the corresponding boundary forms (3.1) are also separated. Thus Theorem 4.1 in conjunction with Proposition 4.1 extends the results in [3].

5. An example and a further application of moment theory. In this section we discuss an interesting example of Theorem 4.1, as well as present another application of Theorem 2.1.

We begin by recalling that the nth Bernoulli polynomial, $B_n(x)$, is determined by the relations

$$\begin{aligned}&B_0(x)=1,\qquad B'_n(x)=nB_{n-1}(x), && n=1,2,\cdots,\\ &B_n(x)=(-1)^nB_n(1-x)\\ &B_n(0)=0, && n=3,5,\cdots.\end{aligned}\tag{5.1}$$

The periodic extension of period one of the Bernoulli polynomial which we denote by $\bar{B}_n$ is, according to (5.1), a monospline of degree n with knots at the integers. $\bar{B}_n(x)$ is the Peano kernel for the Euler–Maclaurin quadrature formula

$$\begin{aligned}\int_0^N f(x)\,dx=&\frac{1}{2}f(0)+f(1)+\cdots+f(N-1)+\frac{1}{2}f(N)\\&+\sum_{0<2\nu\leqq n}\frac{B_{2\nu}(0)}{(2\nu)!}\left(f^{(2\nu-1)}(0)-f^{(2\nu-1)}(N)\right)\\&+\frac{(-1)^n}{n!}\int_0^N\bar{B}_n(x)f^{(n)}(x)\,dx.\end{aligned}\tag{5.2}$$

When n is even, $n=2m$, we may rewrite (5.2) in the form

$$\int_0^N f(x)\,dx=\frac{1}{2}f(0)+f(1)+\cdots+f(N-1)+\frac{1}{2}f(N)$$
$$+\sum_{\nu=1}^{m-1}\frac{B_{2\nu}(0)}{(2\nu)!}[f^{(2\nu-1)}(0)-f^{(2\nu-1)}(1)] \tag{5.3}$$
$$+\frac{1}{(2m)!}\int_0^N M(x)f^{(2m)}(x)\,dx,$$

where $M(x)=\bar{B}_{2m}(x)-B_{2m}(0)$. M has a double zero at every integer. Furthermore, it satisfies the boundary conditions

$$M^{(2m-1-i)}(0)=M^{(2m-1-i)}(N)=0, \qquad i\notin\{0,1,3,\cdots,2m-3\},$$

which are adjoint to the boundary terms appearing in (5.3). Thus we see that the Euler–Maclaurin quadrature formula (5.3) is exact for all spline functions of degree $2m-1$ with double knots at $1,2,\cdots,N-1$. Thus it is of double precision. In the notation of Thorem 4.1, $n=2m$, $l=N-1$, $k=2m$ and $r=2(N-1)$.

Similarly, the odd degree Bernoulli monospline $M(x)=\bar{B}_{2m-1}(x)$ is the Peano kernel of the (odd degree) Euler–Maclaurin quadrature formula

$$\int_0^N f(x)\,dx=\frac{1}{2}f(0)+f(1)+\cdots+f(N-1)+\frac{1}{2}f(N)$$
$$+\sum_{\nu=1}^{m-1}\frac{B_{2\nu}(0)}{(2\nu)!}[f^{(2\nu-1)}(0)-f^{(2\nu-1)}(N)] \tag{5.4}$$
$$-\frac{1}{(2m-1)!}\int_0^N M(x)f^{(2m-1)}(x)\,dx.$$

In this case, M has a simple zero at each integer and half integer. Also, M satisfies

$$M^{(2m-2-i)}(0)=M^{(2m-2-i)}(N)=0, \qquad i\notin\{0,1,3,\cdots,2m-3\}.$$

Thus (5.4) is of double precision and corresponds to Theorem 4.1 with $n=2m-1$, $l=N-1$, $k=2m$ and $r=2N-1$.

The following theorem was suggested to us by A. A. Melkman who indicated a method of proof similar to that used in [6].

THEOREM 5.1. *Let data $y_1,\cdots,y_{n+2r+1}$ and points $x_1<x_2<\cdots<x_{n+2r+1}$ be given. Suppose that the divided differences $[y_i,\cdots,y_{i+n}]$ of the data over the points $x_i,\cdots,x_{i+n}$, $i=1,\cdots,2r+1$, strictly alternates in sign and $n\geqq 2$. Then there exists a monospline $M(x)$ of degree n with r knots and a nonzero constant λ such that*

$$M(x_i)=\lambda y_i, \qquad i=1,\cdots,n+2r+1.$$

Furthermore, if $2r\leqq n$, then $M(x)$ is unique.

Proof. Assume without loss of generality that $x_1=0$ and $x_{n+2r+1}=1$. Consider the space $\mathcal{S}_0$ of spline functions of the form

$$S(t)=\sum_{i=1}^{2r+1}c_iB(x_i,\cdots,x_{i+n};t),$$

where $\sum_{i=1}^{2r+1} c_i[y_i, \cdots, y_{i+n}] = 0$, and $B(x_i, \cdots, x_{i+n}; t)$ is (the B-spline) defined to be the nth divided difference of $(x-t)_+^{n-1}$ at $x = x_i, \cdots, x_{i+n}$. It is well known that any subset of B-splines form a weak Chebyshev system (cf. [2]). Since the divided difference of the data strictly alternates, we conclude that $\mathcal{S}_0$ is a weak Chebyshev subspace of dimension $2r$. To prove this fact, let us set $B(x_i, \cdots, x_{i+n}; t) = u_i(t)$, $i = 1, \cdots, 2r+1$, and $[y_i, \cdots, y_{i+n}] = z_i$, $i = 1, \cdots, 2r+1$, and consider the functions $v_i(t) = u_i(t) - (z_i/z_{2r+1})u_{2r+1}(t)$, $i = 1, \cdots, 2r$. Note that

$$z_{2r+1} V\begin{pmatrix} 1, \cdots, 2r \\ t_1, \cdots, t_{2r} \end{pmatrix} = \begin{vmatrix} u_1(t_1) & \cdots & u_1(t_{2r}) & z_1 \\ \vdots & & \vdots & \vdots \\ u_{2r+1}(t_1) & \cdots & u_{2r+1}(t_{2r}) & z_{2r+1} \end{vmatrix}$$

$$= \sum_{i=1}^{2r+1} (-1)^{i+2r+1} z_i U\begin{pmatrix} 1, \cdots, i-1, i+1, \cdots, 2r+1 \\ t_1 \quad \cdots, \quad t_{2r} \end{pmatrix}.$$

Now

$$U\begin{pmatrix} 1, \cdots, i-1, i+1, \cdots, 2r+1 \\ t_1, \quad \cdots, \quad t_{2r} \end{pmatrix} \geqq 0,$$

and $z_i(-1)^i\sigma > 0$, $\sigma^2 = 1$, fixed. Thus $\{v_i(t)\}_{i=1}^{2r}$ is a weak Chebyshev system of dimension $2r$ on (0, 1) which spans the set $\mathcal{S}_0$.

Since $z_i(-1)^i\sigma > 0$, $i = 1, \cdots, 2r+1$, we may always find positive numbers $c_1^0, \cdots, c_{2r+1}^0$ such that $\sum_{i=1}^{2r+1} c_i^0 z_i = \sum_{i=1}^{2r+1} c_i^0 [y_i, \cdots, y_{i+n}] = 0$. Thus the function $S(t) = \sum_{i=1}^{2r+1} c_i^0 B(x_i, \cdots, x_{i+n}; t)$ is strictly positive on (0, 1), and from Theorem 2.1, there exist points $0 < \xi_1 < \cdots < \xi_r < 1$, and $\mu_i > 0$, $i = 1, \cdots, r$, such that

$$\int_0^1 f(t)\, dt = \sum_{i=1}^{r} \mu_i f(\xi_i), \tag{5.5}$$

for all $f \in \mathcal{S}_0$. It easily follows that there is a constant λ for which

$$\int_0^1 B(x_j, \cdots, x_{j+n}; t)\, dt = \sum_{i=1}^{r} \mu_i B(x_j, \cdots, x_{j+n}; \xi_i) + \lambda[y_j, \cdots, y_{j+n}],$$

for $j = 1, \cdots, 2r+1$.

Let

$$M(x) = \sum_{j=0}^{n-1} a_j x^j + \int_0^1 (x-t)_+^{n-1}\, dt - \sum_{i=1}^{r} \mu_i (x-\xi_i)_+^{n-1},$$

where $a_0, a_1, \cdots, a_{n-1}$ are chosen so that $M(x_j) = \lambda y_j$, $j = 1, \cdots, n$. Now

$$M(x_j, \cdots, x_{j+n}) = \int_0^1 B(x_j, \cdots, x_{j+n}; t)\, dt - \sum_{i=1}^{r} \mu_i B(x_j, \cdots, x_{j+n}; \xi_i)$$

$$= \lambda[y_j, \cdots, y_{j+n}], \qquad j = 1, \cdots, 2r+1.$$

Since $M(x_j)=\lambda y_j$, $j=1,\cdots,n$, it then follows that $M(x_j)=\lambda y_j$, $j=1,\cdots,n+2r+1$. Note that if $\lambda=0$, then $M(x)$ has $n+2r+1$ zeros, an impossibility (Theorem 4.3). Thus $\lambda\neq 0$ and $M(x)$ is the desired monospline.

In [11], it is proven that the functions $1, t, \cdots, t^{n-1}$ are each contained in the smallest linear subspace containing the convexity cone generated by $\mathcal{S}_0$. Since uniqueness of the monospline $M(x)$ is equivalent to the uniqueness of the quadrature formula (5.5), we conclude from Corollary 2.2 that M is unique when $2r\leqq n$. This completes the proof of the theorem.

Remark 5.1. In the statement of Theorem 5.1, we assumed $z_i(-1)^i\sigma>0$, $i=1,\cdots,2r+1$. This was done to insure that $\mathcal{S}_0$ is a weak Chebyshev subspace of dimension $2r$ which contains a positive function. In order that $\mathcal{S}_0$ be a weak Chebyshev subspace of dimension $2r$, it is sufficient that $z_i(-1)^i\sigma\geqq 0$, $i=1,\cdots,2r+1$, and at least one of the z_i is nonzero. Assuming that this is the case and if the sets $\{i\colon z_i>0\}$ and $\{i\colon z_i<0\}$ are both nonempty, then we may construct, as in the proof of Theorem 5.1, an element of $\mathcal{S}_0$ which is strictly positive on (0, 1). If one of the above two sets is empty, but the other does not contain either 1 or $2r+1$, and if $n\geqq 3$, then we may still construct a positive function in $\mathcal{S}_0$. These conditions suffice for Theorem 5.1 to hold.

In particular, if we choose $y_i=\delta_{i,n+2r}$, $i=1,\cdots,n+2r+1$, we obtain

COROLLARY 5.1. *Given any points $s_1<\cdots<s_{n+2r}$, there exists a monospline $M(x)$ of degree n with r knots such that*

$$M(s_i)=0, \qquad i=1,\cdots,n+2r.$$

This is the fundamental theorem of algebra for monosplines as it appears in [6] and [13]. This result is also a special case of Theorem 4.1 with $k=2n$. The uniqueness as well as the converse, i.e., $M(x)$ has no more than $n+2r$ zeros, are also results of Theorems 4.1 and 4.3.

REFERENCES

[1] H. BURCHARD, *Interpolation and approximation by generalized convex functions*, Ph.D. dissertation, Purdue Univ., West Lafayette, Ind., 1968.

[2] S. KARLIN, *Total Positivity*, Vol. I, Stanford University Press, Stanford, Calif., 1968.

[3] S. KARLIN AND C. A. MICCHELLI, *The fundamental theorem of algebra for monosplines satisfying boundary conditions*, Israel J. Math., 11 (1972), pp. 405–451.

[4] S. KARLIN AND A. PINKUS, *Gaussian quadrature formulae with multiple nodes*, Studies in Spline functions and Approximation Theory, S. Karlin, C. A. Micchelli, A. Pinkus and I. J. Schoenberg, eds., Academic Press, New York, 1976.

[5] ———, *Interpolation by Splines with Mixed Boundary Conditions*, Academic Press, New York, 1976.

[6] S. KARLIN AND L. SCHUMAKER, *The fundamental theorem of algebra for Tchebycheffian monosplines*, J. Analyse Math., 20 (1967), pp. 233–270.

[7] S. KARLIN AND W. J. STUDDEN, *Tchebycheff Systems: with Applications in Analysis and Statistics*, Interscience, New York, 1966.

[8] M. G. KREIN, *The ideas of P. L. Chebyshev and A. A. Markov in the theory of limiting values of integrals and their further developments*, Amer. Math. Soc. Transl., 12 (1951), pp. 1–122.

[9] A. A. MELKMAN, *Interpolation by splines satisfying mixed boundary conditions*, Israel J. Math., 19 (1974), pp. 369–381.

[10] C. A. MICCHELLI, *The fundamental theorem of algebra for monosplines with multiplicities*, Linear Operators and Approximation, Proc. Conf. in Oberwolfach, P. L. Butzer, J. P. Kahane and B. Sz. Nagy, eds., Birkhaüser Verlag, Basel, Switzerland, 1971, pp. 14–22.

[11] ———, *Best L^1-approximation by weak Chebyshev systems and the uniqueness of interpolating perfect splines*, IBM Res. Rep. #5388, Thomas J. Watson Res. Center, Yorktown Heights, N.Y., 1975, to appear J. Approx. Theory.

[12] A. PINKUS, *Representation theorems for Tchebycheffian polynomials with boundary conditions and their applications*, Israel J. Math., 17 (1974), pp. 11–34.

[13] I. J. SCHOENBERG, *Spline functions, convex curves and mechanical quadrature*, Bull. Amer. Math. Soc., 64 (1958), pp. 352–357.

[14] C. A. MICCHELLI AND W. L. MIRANKER, *High order search methods for finding roots*, J. Assoc. Comput. Mach., 22 (1975), pp. 51–60.

SIAM J. MATH. ANAL.
Vol. 8, No. 2, April 1977

ELLIPTIC INTEGRALS OF THE FIRST KIND*

B. C. CARLSON†

Abstract. The reciprocal square root of any real polynomial with known zeros and degree not exceeding four is integrated in terms of a standard integral by a new quadratic transformation which preserves symmetry in the zeros. If at least one zero is real, this method, unlike earlier methods, leads to a single standard integral instead of a difference of two standard integrals even when neither limit of integration is a zero. If no zero is real, a particular point on the real line has special significance. Formulas listed in integral tables are unified and generalized.

1. The case of real zeros. The general elliptic integral of the first kind is

$$\int_y^x \frac{dt}{[(a+\alpha t)(b+\beta t)(c+\gamma t)(d+\delta t)]^{1/2}}. \tag{1}$$

The quartic polynomial is assumed to be real-valued, although linear factors might be conjugate complex. The zeros of the polynomial are $-a/\alpha, \cdots, -d/\delta$, and if $d \neq 0$ and $\delta \to 0$ the last zero tends to infinity. The quartic then becomes a cubic polynomial, which we regard as a quartic with one zero at infinity, and similarly for polynomials of lower degree. To make the integral well-defined, we assume that the integrand is strictly positive on the open interval of integration, which cannot contain a zero. If both limits of integration are zeros, the integral is called complete; otherwise it is incomplete.

In this section we assume the finite zeros are real, deferring discussion of complex zeros to § 3 and § 4. Therefore we may suppose that $a+\alpha t, \cdots, d+\delta t$ are strictly positive on the open interval of integration. The cubic and quartic cases occur frequently in many parts of applied mathematics but, aside from a few technical exceptions, they are listed in present integral tables only if one limit of integration is a zero. If neither limit is a zero, the integral must be split in two, each part having a zero as one limit. In most tables sixteen cases are distinguished according to whether the polynomial is cubic or quartic and whether the upper or lower limit of integration is each in turn of the four zeros arranged in order on the real line (extended real line in the cubic case). A typical one of the sixteen formulas is [2, 252.00]

$$\int_d^x [(a-t)(b-t)(c-t)(t-d)]^{-1/2}\,dt = 2[(a-c)(b-d)]^{-1/2}F(\varphi, k), \tag{2}$$

$$\sin\varphi = \left[\frac{(a-c)(x-d)}{(c-d)(a-x)}\right]^{1/2}, \quad k^2 = \frac{(a-b)(c-d)}{(a-c)(b-d)},$$

$$a > b > c \geqq x > d,$$

where $F(\varphi, k)$ is Legendre's standard integral of the first kind,

$$F(\varphi, k) = \int_0^{\varphi} (1-k^2\sin^2\theta)^{-1/2}\,d\theta. \tag{3}$$

* Received by the editors September 4, 1975.

† Ames Laboratory-ERDA and Departments of Mathematics and Physics, Iowa State University, Ames Iowa 50011.

The formal symmetry of (1) in the zeros has of course been impaired by choosing one zero as the lower limit of integration in (2), but even the symmetry in the remaining zeros has been concealed on the right side of (2) by adopting Legendre's notation. The integral is complete if $x = c$, which implies $\varphi = \pi/2$.

Nellis and Carlson [6, Table I] likewise require one limit of integration to be a zero but preserve formal symmetry in all other finite zeros, thereby reducing the sixteen cases to four:

$$\int_d^x [(a+\alpha t)(b+\beta t)(c+\gamma t)(t-d)]^{-1/2}\,dt$$
$$= 2[(a+\alpha d)(b+\beta d)(c+\gamma d)]^{-1/2}(x-d)^{1/2}R_F\left(\frac{a+\alpha x}{a+\alpha d}, \frac{b+\beta x}{b+\beta d}, \frac{c+\gamma x}{c+\gamma d}\right), \tag{4}$$

$$\int_y^d [(a+\alpha t)(b+\beta t)(c+\gamma t)(d-t)]^{-1/2}\,dt$$
$$= 2[(a+\alpha d)(b+\beta d)(c+\gamma d)]^{-1/2}(d-y)^{1/2}R_F\left(\frac{a+\alpha y}{a+\alpha d}, \frac{b+\beta y}{b+\beta d}, \frac{c+\gamma y}{c+\gamma d}\right), \tag{5}$$

$$\int_{-\infty}^x [(a+\alpha t)(b+\beta t)(c+\gamma t)]^{-1/2}\,dt$$
$$= 2R_F[(a+\alpha x)\beta\gamma, (b+\beta x)\gamma\alpha, (c+\gamma x)\alpha\beta], \tag{6}$$

$$\int_y^\infty [(a+\alpha t)(b+\beta t)(c+\gamma t)]^{-1/2}\,dt$$
$$= 2R_F[(a+\alpha y)\beta\gamma, (b+\beta y)\gamma\alpha, (c+\gamma y)\alpha\beta], \tag{7}$$

where R_F is the standard symmetric integral of the first kind [3],

$$R_F(u, v, w) = \frac{1}{2}\int_0^\infty [(t+u)(t+v)(t+w)]^{-1/2}\,dt. \tag{8}$$

The factor $\frac{1}{2}$ makes $R_F(1, 1, 1) = 1$. The standard integral is symmetric and homogeneous of degree $-\frac{1}{2}$ in u, v, w. It is related to Legendre's integral by

$$R_F(u, v, w) = (w-u)^{-1/2}F\left[\arccos\left(\frac{u}{w}\right)^{1/2}, \left(\frac{w-v}{w-u}\right)^{1/2}\right],$$
$$F(\varphi, k) = (\sin\varphi)R_F(\cos^2\varphi, 1-k^2\sin^2\varphi, 1). \tag{9}$$

The linear transformations of Legendre's integral are equivalent to the permutation symmetry of R_F, which makes it unnecessary to order the zeros in equations (4) to (7). In (4) and (5) the cubic case is the case $\gamma = 0$. All four integrals are complete if $a+\alpha x = a+\alpha y = 0$, which implies that one argument of R_F is 0. To deduce (2) from (4), put $\alpha = \beta = \gamma = -1$ and use (9) with $u = (c-x)/(c-d)$, $v = (b-x)/(b-d)$, $w = (a-x)/(a-d)$ so that $u < v < w$ under the conditions stated in (2). However, there is no advantage in returning to (2) since algorithms for computing R_F are given in [4] and a FORTRAN program is available on request. The fifteen companions of (2) can be obtained similarly.

To evaluate (1) if neither x nor y is a zero, we may split the integral into two parts. For example, if no zero lies in the interval $d<t<x$, we can use (4) to evaluate

$$\int_y^x [(a+\alpha t)(b+\beta t)(c+\gamma t)(t-d)]^{-1/2}\,dt = \int_d^x - \int_d^y \tag{10}$$

as a difference of two standard integrals, each symmetric in all zeros but d. The point of this paper is to observe that explicit symmetry in all four zeros is restored if the two standard integrals are combined into one by the addition theorem. We do this in a special case of (1) to obtain a fundamental quadratic transformation (13) which is then applied to the general case of (1). The result is

$$\int_y^x [(a+\alpha t)(b+\beta t)(c+\gamma t)(d+\delta t)]^{-1/2}\,dt = 2R_F(U^2, V^2, W^2), \qquad x>y,$$

$$\begin{aligned}
(x-y)U &= [(a+\alpha x)(b+\beta x)(c+\gamma y)(d+\delta y)]^{1/2} \\
&\quad + [(a+\alpha y)(b+\beta y)(c+\gamma x)(d+\delta x)]^{1/2}, \\
(x-y)V &= [(a+\alpha x)(b+\beta y)(c+\gamma x)(d+\delta y)]^{1/2} \\
&\quad + [(a+\alpha y)(b+\beta x)(c+\gamma y)(d+\delta x)]^{1/2}, \\
(x-y)W &= [(a+\alpha x)(b+\beta y)(c+\gamma y)(d+\delta x)]^{1/2} \\
&\quad + [(a+\alpha y)(b+\beta x)(c+\gamma x)(d+\delta y)]^{1/2}, \\
V^2-W^2 &= (a\beta-b\alpha)(c\delta-d\gamma), \quad W^2-U^2 = (a\gamma-c\alpha)(d\beta-b\delta), \\
U^2-V^2 &= (a\delta-d\alpha)(b\gamma-c\beta).
\end{aligned} \tag{11}$$

Since U, V, W have finite limits as $x\to\infty$ or $y\to-\infty$, (11) has an unambiguous meaning if x or y is infinite. They cannot both be infinite because $a+\alpha t, \cdots, d+\delta t$ are assumed to be strictly positive on the open interval of integration.

The variables U, V, W correspond to the three ways of pairing the four zeros, and permutations of the zeros induce permutations of U, V, W, leaving R_F unchanged. Most methods of reducing (1) (see [5, § 13.5] and [1, § 17.8]) introduce asymmetry through transformed limits of integration if not through the transformed integrand. The right side of each of the equations defining U, V, W is the sum of two nonnegative terms which differ by interchange of x and y. One of the terms in each equation vanishes if either x or y is a zero, and one of the variables U, V, W vanishes if both x and y are zeros, the integral being then complete. To recover (5) put $x=d$ and $\delta=-1$, whence $d+\delta x=0$ and $d+\delta y=d-y$, and use the homogeneity of R_F. A similar procedure leads to (4) except that the sign of d must first be changed throughout (11) to adapt the notation to (4). Equations (6) and (7) are recovered by putting $\delta=0$ and $d=1$ and taking the limits of U, V, W as $x\to\infty$ or $y\to-\infty$. (In both cases the assumption of positivity requires α, β, γ to have the same sign as the quantity which tends to infinity.) Therefore (11) reproduces all sixteen formulas in which one limit of integration is a zero.

The case in which the quartic in t (possibly with conjugate imaginary zeros) is a quadratic in t^2 with real zeros occurs often in practice and is widely used as a canonical form. Unless the limits of integration have opposite signs, the integral

can first be transformed to the cubic case by putting $t = s^{1/2}$ and then evaluated by using (11). The result is

$$\int_y^x [(at^2+c)(\alpha t^2+\gamma)]^{-1/2}\,dt = R_F(U^2, V^2, W^2), \qquad 0 \leqq y < x \leqq \infty,$$

$$\begin{aligned}
(x^2-y^2)U &= x(ay^2+c)^{1/2}(\alpha y^2+\gamma)^{1/2} + y(ax^2+c)^{1/2}(\alpha x^2+\gamma)^{1/2},\\
(x^2-y^2)V &= x(ax^2+c)^{1/2}(\alpha y^2+\gamma)^{1/2} + y(ay^2+c)^{1/2}(\alpha x^2+\gamma)^{1/2},\\
(x^2-y^2)W &= x(ay^2+c)^{1/2}(\alpha x^2+\gamma)^{1/2} + y(ax^2+c)^{1/2}(\alpha y^2+\gamma)^{1/2},\\
V^2-W^2 &= a\gamma - c\alpha, \quad W^2-U^2 = c\alpha, \quad U^2-V^2 = -a\gamma.
\end{aligned} \tag{12}$$

We assume that at^2+c and $\alpha t^2+\gamma$ are strictly positive on the open interval of integration. The variables U, V, W have finite limits as $x \to \infty$. Twelve special cases in which one limit of integration corresponds to a zero of the cubic polynomial in $s = t^2$ are listed in integral tables (see for example [1, p. 596]).

Formulas analogous to (11) and (12) can be obtained for integrals of the second and third kinds but will not be discussed here.

2. The fundamental transformation. Equation (11) will be deduced after proving the special case,

$$\begin{aligned}
&\int_0^\infty [(t+A)(t+B)(t+C)(t+D)]^{-1/2}\,dt = \int_0^\infty [(t+X^2)(t+Y^2)(t+Z^2)]^{-1/2}\,dt,\\
&X = (AB)^{1/2}+(CD)^{1/2}, \quad Y = (AC)^{1/2}+(BD)^{1/2},\\
&\qquad\qquad Z = (AD)^{1/2}+(BC)^{1/2}, \quad A, B, C, D > 0.
\end{aligned} \tag{13}$$

This fundamental quadratic transformation from the quartic to the cubic case is new. The upper limit of integration is a zero of the cubic but not of the quartic.

By symmetry we may suppose that A is not greater than B, C, or D. Since each integral is jointly continuous if A, B, C, D are strictly positive, it suffices to prove (13) when A is strictly less than B, C and D. By (4) and the homogeneity of R_F, the integral on the left side is

$$\begin{aligned}
&\int_{-A}^\infty [(t+A)(t+B)(t+C)(t+D)]^{-1/2}\,dt - \int_{-A}^0 [(t+A)(t+B)(t+C)(t+D)]^{-1/2}\,dt\\
&= 2[(B-A)(C-A)(D-A)]^{-1/2}\\
&\quad\cdot\left[R_F\left(\frac{1}{B-A}, \frac{1}{C-A}, \frac{1}{D-A}\right) - A^{1/2}R_F\left(\frac{B}{B-A}, \frac{C}{C-A}, \frac{D}{D-A}\right)\right]\\
&= 2R_F[(C-A)(D-A), (D-A)(B-A), (B-A)(C-A)]\\
&\quad - 2R_F\left[\frac{B}{A}(C-A)(D-A), \frac{C}{A}(D-A)(B-A), \frac{D}{A}(B-A)(C-A)\right].
\end{aligned} \tag{14}$$

The two standard integrals can be combined by the addition theorem [7, § 8],

$$\begin{aligned}
&R_F(x, y, z) = R_F(x+\lambda, y+\lambda, z+\lambda) + R_F(x+\mu, y+\mu, z+\mu),\\
&(\lambda\mu - xy - yz - zx)^2 = 4xyz(\lambda+\mu+x+y+z),\\
&\mu = \lambda^{-1}(xy+yz+zx) + 2\lambda^{-2}xyz + 2\lambda^{-2}[xyz(x+\lambda)(y+\lambda)(z+\lambda)]^{1/2}.
\end{aligned} \tag{15}$$

Putting $x=(C-A)(D-A)$, $y=(D-A)(B-A)$, $z=(B-A)(C-A)$ and $\lambda=(B-A)(C-A)(D-A)/A$, we find

$$\begin{aligned}&\mu=A(B+C+D-A)+2(ABCD)^{1/2},\\ &x+\mu=X^2,\quad y+\mu=Y^2,\quad z+\mu=Z^2,\end{aligned}\tag{16}$$

where X, Y, Z were defined above. Hence the left side of (13) equals $2R_F(X^2, Y^2, Z^2)$, which by (8) equals the right side. It is not known whether (13) can be proved more directly by a change of integration variable. See the note added in proof. Evaluating the left side by [6, (T.2)], we can put (13) in the alternative form,

$$\begin{aligned}&R_{-1}(\tfrac12,\tfrac12,\tfrac12,\tfrac12;A,B,C,D)=2R_F(X^2,Y^2,Z^2), && A,B,C,D>0,\\ &X=(AB)^{1/2}+(CD)^{1/2},\quad Y=(AC)^{1/2}+(BD)^{1/2}, && Z=(AD)^{1/2}+(BC)^{1/2},\\ &Y^2-Z^2=(A-B)(C-D),\quad Z^2-X^2=(A-C)(D-B), && X^2-Y^2=(A-D)(B-C).\end{aligned}\tag{17}$$

In place of the notation $R(1;\frac12,\frac12,\frac12,\frac12)$ in [6], we use $R_{-1}(\frac12,\frac12,\frac12,\frac12)$ here, the subscript being the degree of homogeneity.

By [6, (T.1)] (with one more linear factor inserted),

$$\begin{aligned}&\int_y^x[(a+\alpha t)(b+\beta t)(c+\gamma t)(d+\delta t)]^{-1/2}\,dt\\ &\quad=(x-y)[(a+\alpha y)(b+\beta y)(c+\gamma y)(d+\delta y)]^{-1/2}\\ &\qquad\cdot R_{-1}\left(\frac12,\frac12,\frac12,\frac12;\frac{a+\alpha x}{a+\alpha y},\frac{b+\beta x}{b+\beta y},\frac{c+\gamma x}{c+\gamma y},\frac{d+\delta x}{d+\delta y}\right).\end{aligned}\tag{18}$$

Transformation by (17) and use of homogeneity prove (11).

Since an R-function is holomorphic in each of its variables on the plane cut along the nonpositive real axis, the permanence of functional relations implies that (17) holds if A, B, C, D lie in the cut plane and are such that X, Y, Z lie in the open right half-plane. In calculating X, Y, Z from A, B, C, D we take the square root of a product to be the product of the principal square roots of the factors. Therefore a sufficient but not necessary condition is that A, B, C, D lie in the open right half-plane, since $(AB)^{1/2}, \cdots, (CD)^{1/2}$ also lie in this half-plane.

In § 3 some of the variables will indeed be complex, and we shall arrive at real variables by using Landen's transformation [4, (2.3)],

$$\begin{aligned}&R_F(X^2,Y^2,Z^2)=2R_F(L^2,M^2,N^2),\\ &M=Y+Z,\quad N+L=2(X+Y)^{1/2}(X+Z)^{1/2},\\ &N-L=2(X-Y)^{1/2}(X-Z)^{1/2},\\ &N^2-M^2=[(X^2-Y^2)^{1/2}+(X^2-Z^2)^{1/2}]^2,\\ &L^2-M^2=[(X^2-Y^2)^{1/2}-(X^2-Z^2)^{1/2}]^2,\quad LN=2MX.\end{aligned}\tag{19}$$

The transformation is valid if X, Y, Z, L, M, N are in the open right half-plane. If $X>0$, $Z=\bar{Y}$, and $\operatorname{Re} Y>0$, where a bar signifies complex conjugation, then $0<L\leqq M\leqq N$. If $X\to 0$, then $L\to 0$ by the last equation of (19). Since each

function R_F is continuous in this limit, equality still holds between the two complete integrals. Both functions become infinite if Re $Y \to 0$.

Applying (19) to (17) we find

$$
\begin{aligned}
&R_{-1}(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}; A, B, C, D) = 4R_F(L^2, M^2, N^2),\\
&M = (A^{1/2}+B^{1/2})(C^{1/2}+D^{1/2}),\\
&N+L = 2[(A^{1/2}+C^{1/2})(A^{1/2}+D^{1/2})(B^{1/2}+C^{1/2})(B^{1/2}+D^{1/2})]^{1/2},\\
&N-L = 2[(A^{1/2}-C^{1/2})(A^{1/2}-D^{1/2})(B^{1/2}-C^{1/2})(B^{1/2}-D^{1/2})]^{1/2},\\
&N^2-M^2 = [(A-C)^{1/2}(B-D)^{1/2}+(A-D)^{1/2}(B-C)^{1/2}]^2,\\
&L^2-M^2 = [(A-C)^{1/2}(B-D)^{1/2}-(A-D)^{1/2}(B-C)^{1/2}]^2.
\end{aligned}
\tag{20}
$$

The transformation is valid if A, B, C, D are in the plane cut along the nonpositive real axis and L, M, N are in the open right half-plane. If $A>0$, $B>0$, $D=\bar{C}$, and C is in the cut plane, (17) implies $X>0$, $Z=\bar{Y}$, and Re $Y>0$, whence $0<L\leqq M\leqq N$. If $A\to 0$ or $B\to 0$, the equations still hold by continuity and L remains strictly positive. If both A and B tend to zero, so do L and M, and the functions become infinite.

An alternative way of reaching real variables is to apply the inverse of Landen's transformation to the right side of (17). The result is

$$
\begin{aligned}
&R_{-1}(\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}; A, B, C, D) = 2R_F(L_1^2, M_1^2, N_1^2),\\
&M_1 = \frac{(AB)^{1/2}(C+D)+(CD)^{1/2}(A+B)}{(AB)^{1/2}+(CD)^{1/2}},\\
&N_1^2-M_1^2 = N^2-M^2, \quad L_1^2-M_1^2 = L^2-M^2,
\end{aligned}
\tag{21}
$$

where N^2-M^2 and L^2-M^2 are specified in (20). The transformation is valid if A, B, C, D are in the cut plane and L_1, M_1, N_1 are in the open right-half plane. If $A>0$, $B>0$, $D=\bar{C}$, and C is in the cut plane, then $M_1 = YZ/X>0$ and $L_1\leqq M_1\leqq N_1$. However, since $XL_1 = X^2-|Y^2-X^2|$, L_1 is positive if and only if $|A-C|^{1/2}|B-C|^{1/2}<(AB)^{1/2}+|C|$, an inequality which fails, for example, if A is small and B large compared to $|C|$. Even when the inequality holds, it seems preferable to use (20), which is related to (21) by the duplication theorem [7, (8.7), (8.13)],

$$
\begin{aligned}
&R_F(L_1^2, M_1^2, N_1^2) = 2R_F(L^2, M^2, N^2),\\
&L^2 = L_1^2+\lambda, \quad M^2 = M_1^2+\lambda, \quad N^2 = N_1^2+\lambda, \quad \lambda = M_1N_1+N_1L_1+L_1M_1.
\end{aligned}
\tag{22}
$$

This means that the fractional differences of L, M, N are less than those of L_1, M_1, N_1, and hence (20) is a little better for numerical calculation.

3. Two conjugate complex zeros. In (18) we suppose that x, y, $a+\alpha t$, $b+\beta t$ are real, while $c+\gamma t$ and $d+\delta t$ are complex conjugates. Putting $A = (a+\alpha x)/(a+\alpha y), \cdots, D = (d+\delta x)/(d+\delta y)$, we assume for the moment that neither x nor y is a zero of the quartic, so that $A>0, B>0, D=\bar{C}$, and $|\text{ph } C|<\pi$.

Then (20) can be applied to the R-function on the right side of (18) to obtain

$$
\begin{aligned}
&\int_y^x [(a+\alpha t)(b+\beta t)(c+\gamma t)(d+\delta t)]^{-1/2}\,dt = 4R_F(L^2, M^2, N^2),\\
&(x-y)M = [(a+\alpha x)^{1/2}(b+\beta y)^{1/2} + (a+\alpha y)^{1/2}(b+\beta x)^{1/2}]\\
&\qquad\qquad \cdot [(c+\gamma x)^{1/2}(d+\delta y)^{1/2} + (c+\gamma y)^{1/2}(d+\delta x)^{1/2}],\\
&N^2 - M^2 = [(a\gamma - c\alpha)^{1/2}(b\delta - d\beta)^{1/2} + (a\delta - d\alpha)^{1/2}(b\gamma - c\beta)^{1/2}]^2,\\
&L^2 - M^2 = [(a\gamma - c\alpha)^{1/2}(b\delta - d\beta)^{1/2} - (a\delta - d\alpha)^{1/2}(b\gamma - c\beta)^{1/2}]^2.
\end{aligned}
\tag{23}
$$

The discussion following (20) shows that $0 < L \leqq M \leqq N$. By continuity we may now allow x and y to be zeros of the quartic, which means that L may possibly be 0.

In order to rewrite (23) in terms of real quantities, we put

$$
\begin{aligned}
&(c+\gamma t)(d+\delta t) = pt^2 + qt + r,\\
&\xi^2 = (c+\gamma x)(d+\delta x) = px^2 + qx + r,\\
&\eta^2 = (c+\gamma y)(d+\delta y) = py^2 + qy + r,\\
&\rho^2 = (a\gamma - c\alpha)(a\delta - d\alpha) = pa^2 - qa\alpha + r\alpha^2,\\
&\sigma^2 = (b\gamma - c\beta)(b\delta - d\beta) = pb^2 - qb\beta + r\beta^2,\\
&\tau^2 = p(a-b)^2 - q(a-b)(\alpha-\beta) + r(\alpha-\beta)^2.
\end{aligned}
\tag{24}
$$

Then

$$
\begin{aligned}
[(c+\gamma x)^{1/2}(d+\delta y)^{1/2} + (c+\gamma y)^{1/2}(d+\delta x)^{1/2}]^2 &= 2pxy + q(x+y) + 2r + 2\xi\eta\\
&= (\xi+\eta)^2 - p(x-y)^2,\\
N^2 - M^2 = 2pab - q(a\beta + b\alpha) + 2r\alpha\beta + 2\rho\sigma &= (\rho+\sigma)^2 - \tau^2.
\end{aligned}
$$

The final result is

$$
\begin{aligned}
&\int_y^x [(a+\alpha t)(b+\beta t)(pt^2+qt+r)]^{-1/2}\,dt = 4R_F(L^2, M^2, N^2), \qquad q^2 \leqq 4pr,\\
&(x-y)M = [(a+\alpha x)^{1/2}(b+\beta y)^{1/2} + (a+\alpha y)^{1/2}(b+\beta x)^{1/2}]\\
&\qquad\qquad \cdot [(\xi+\eta)^2 - p(x-y)^2]^{1/2},
\end{aligned}
$$

$$
\begin{aligned}
&N^2 - M^2 = (\rho+\sigma)^2 - \tau^2, \quad M^2 - L^2 = \tau^2 - (\rho-\sigma)^2, \quad N^2 - L^2 = 4\rho\sigma,\\
&\xi^2 = px^2 + qx + r, \qquad \eta^2 = py^2 + qy + r,\\
&\rho^2 = pa^2 - qa\alpha + r\alpha^2, \quad \sigma^2 = pb^2 - qb\beta + r\beta^2,\\
&\tau^2 = p(a-b)^2 - q(a-b)(\alpha-\beta) + r(\alpha-\beta)^2.
\end{aligned}
\tag{25}
$$

All quantities are real, all square roots are nonnegative, and $0 \leqq L \leqq M \leqq N$. The integrand is assumed to be strictly positive on the open interval of integration. As

x or y becomes infinite, M has a finite limit:

$$(26)\qquad \begin{aligned} &M\to[\alpha^{1/2}(b+\beta y)^{1/2}+(a+\alpha y)^{1/2}\beta^{1/2}](2p^{1/2}\eta+2py+q)^{1/2}, && x\to\infty,\\ &M\to[(a+\alpha x)^{1/2}(-\beta)^{1/2}+(-\alpha)^{1/2}(b+\beta x)^{1/2}](2p^{1/2}\xi-2px-q)^{1/2}, && y\to-\infty. \end{aligned}$$

As an example consider [2, 242.00],

$$(27)\qquad \int_y^\infty (t^3-1)^{-1/2}\,dt = 3^{-1/4}F(\varphi,k),\quad \cos\varphi=\frac{y-1-\sqrt{3}}{y-1+\sqrt{3}},\quad k^2=\frac{2-\sqrt{3}}{4},\quad y\geqq 1.$$

If $1\leqq y<1+\sqrt{3}$, then $\pi/2<\varphi\leqq\pi$ and $F(\varphi,k)=2K(k)-F(\pi-\varphi,k)$. If the interval of integration is $2\leqq t\leqq 3$, we need the values of one complete and two incomplete integrals:

$$\int_2^3 (t^3-1)^{-1/2}\,dt=\int_2^\infty-\int_3^\infty=3^{-1/4}[2K(k)-F(\varphi_1,k)]-3^{-1/4}F(\varphi_2,k)$$
$$=0.2697,$$
$$\cos\varphi_1=2-\sqrt{3},\quad \cos\varphi_2=(2-\sqrt{3})^2,\quad k=\sin 15^\circ.$$

Alternatively, since $t^3-1=(t-1)(t^2+t+1)$, we put $a=-1$, $\beta=0$, $\alpha=b=p=q=r=1$ in (25) to find

$$\int_y^x (t^3-1)^{-1/2}\,dt=4R_F(M^2-3-2\sqrt{3},\,M^2,\,M^2-3+2\sqrt{3}),$$
$$(28)\qquad (x-y)^2M^2=[(x-1)^{1/2}+(y-1)^{1/2}]^2[(\xi+\eta)^2-(x-y)^2],$$
$$\xi=(x^2+x+1)^{1/2},\quad \eta=(y^2+y+1)^{1/2},\quad 1\leqq y<x\leqq\infty.$$

Choosing $y=2$ and $x=3$ we calculate $4R_F(215.5, 221.9, 222.4)=0.2697$. Since the ratios of the arguments are close to unity, the computation is quick even by expansion in power series. For comparison with (27) we find in the limit as $x\to\infty$ that $M^2=1+2y+2\eta$ and

$$\int_y^\infty (t^3-1)^{-1/2}\,dt=2^{3/2}R_F(z-1-\sqrt{3},\,z+\tfrac{1}{2},\,z-1+\sqrt{3}),$$
$$z=y+(y^2+y+1)^{1/2},\quad y\geqq 1$$

If $y\geqq 1+\sqrt{3}$ application of (21) instead of (20) to the right side of (18) gives

$$\int_y^\infty (t^3-1)^{-1/2}\,dt=2(y-1)^{1/2}R_F[(y-1-\sqrt{3})^2,\,y^2+y+1,\,(y-1+\sqrt{3})^2],$$
$$y\geqq 1+\sqrt{3},$$

which by (9) agrees with (27). The inequality discussed after (21) fails in the interval $1 \leqq y < 1+\sqrt{3}$, where $\varphi > \pi/2$ in (27). By contrast the variables of R_F in (29) are strictly positive for $y > 1$.

4. Two pairs of conjugate complex zeros. This is the only case in which the integral (1) sometimes cannot be reduced to a single standard integral, but we consider first the circumstances in which it can. We may suppose that $a+\alpha t$ and $b+\beta t$ are complex conjugates, and likewise $c+\gamma t$ and $d+\delta t$. In (11) we take each square root of a product to be the product of the square roots of the factors. The square root of each linear factor such as $a+\alpha t$ is continuous in t, and conjugate factors have conjugate square roots. Then U, V, W are real and $U>0$, but (11) is not valid if either V or W is negative (see § 2). Their sum is given by

$$\begin{aligned}(x-y)(V+W) &= [(a+\alpha x)^{1/2}(b+\beta y)^{1/2}+(a+\alpha y)^{1/2}(b+\beta x)^{1/2}] \\ &\quad \cdot [(c+\gamma x)^{1/2}(d+\delta y)^{1/2}+(c+\gamma y)^{1/2}(d+\delta x)^{1/2}].\end{aligned} \tag{30}$$

The quantity

$$\begin{aligned}(a+\alpha x)^{1/2}(b+\beta y)^{1/2} &= \left(\frac{a+\alpha x}{a+\alpha y}\right)^{1/2}(a+\alpha y)^{1/2}(b+\beta y)^{1/2} \\ &= \left(\frac{a+\alpha x}{a+\alpha y}\right)^{1/2}|a+\alpha y|\end{aligned} \tag{31}$$

lies in the open right half-plane, and so does $(c+\gamma x)^{1/2}(d+\delta y)^{1/2}$. Therefore both quantities in square brackets are positive, and $V+W>0$. It follows that (11) is valid if and only if $VW \geqq 0$, which is equivalent to

$$\begin{aligned}&[(a+\alpha x)(b+\beta y)+(a+\alpha y)(b+\beta x)][(c+\gamma x)(d+\delta x)(c+\gamma y)(d+\delta y)]^{1/2} \\ &\quad +[(c+\gamma x)(d+\delta y)+(c+\gamma y)(d+\delta x)] \\ &\quad \cdot [(a+ax)(b+\beta x)(a+ay)(b+\beta y)]^{1/2} \geqq 0.\end{aligned} \tag{32}$$

We now eliminate all complex quantities by a change of notation, replacing $(a+\alpha t)(b+\beta t)$ by at^2+bt+c and $(c+\gamma t)(d+\delta t)$ by $\alpha t^2+\beta t+\gamma$ and subsequently defining $d=(4ac-b^2)^{1/2}$ and $\delta=(4\alpha\gamma-\beta^2)^{1/2}$. Thus a, c, α, γ, d, δ are henceforth strictly positive. The validity condition (32) becomes

$$\begin{aligned}&[axy+\tfrac{1}{2}b(x+y)+c](\alpha x^2+\beta x+\gamma)^{1/2}(\alpha y^2+\beta y+\gamma)^{1/2} \\ &\quad +[\alpha xy+\tfrac{1}{2}\beta(x+y)+\gamma](ax^2+bx+c)^{1/2}(ay^2+by+c)^{1/2} \geqq 0,\end{aligned} \tag{33}$$

while (11) becomes

$$\begin{aligned}&\int_y^x (at^2+bt+c)^{-1/2}(\alpha t^2+\beta t+\gamma)^{-1/2}\,dt = 2R_F(U^2, V^2, W^2), \\ &(x-y)U = (ax^2+bx+c)^{1/2}(\alpha y^2+\beta y+\gamma)^{1/2} \\ &\qquad +(ay^2+by+c)^{1/2}(\alpha x^2+\beta x+\gamma)^{1/2}, \\ &V^2-W^2 = d\delta, \quad W^2-U^2 = -a\gamma-c\alpha+\tfrac{1}{2}b\beta-\tfrac{1}{2}d\delta, \\ &U^2-V^2 = a\gamma+c\alpha-\tfrac{1}{2}b\beta-\tfrac{1}{2}d\delta, \\ &\qquad d=(4ac-b^2)^{1/2}, \quad \delta=(4\alpha\gamma-\beta^2)^{1/2}, \quad a, c, d, \alpha, \gamma, \delta > 0.\end{aligned} \tag{34}$$

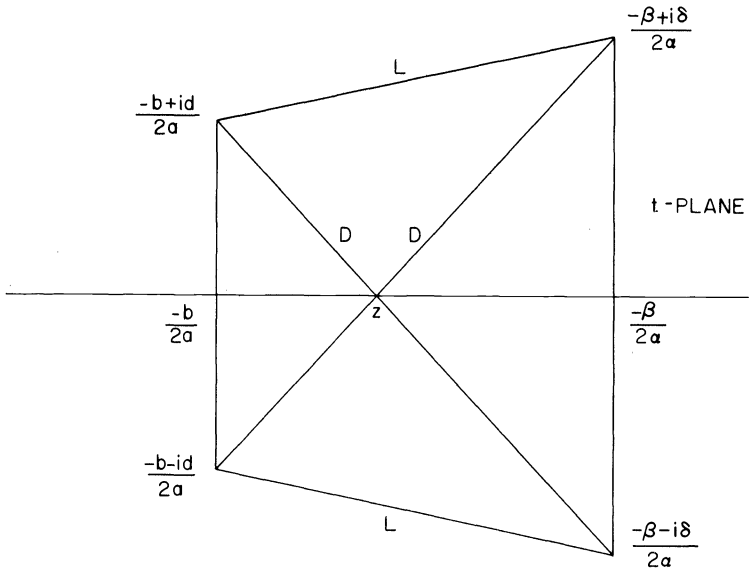


FIG. 1

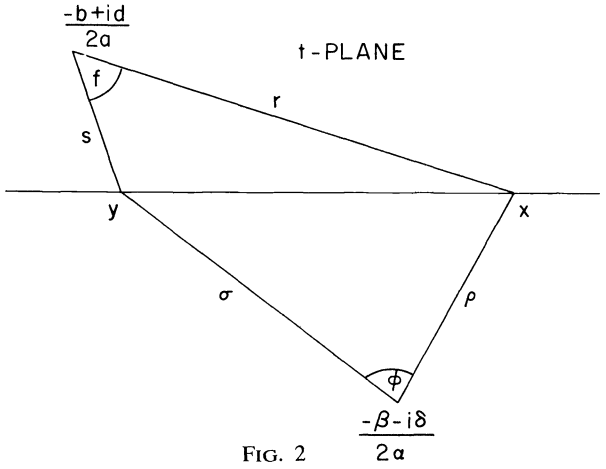


FIG. 2

Although (34) is satisfactory for numerical purposes when (33) is satisfied, it is enlightening to use the notation shown in Figs. 1 and 2. The four complex zeros are the vertices of a quadrilateral with sides of length L, d/a, δ/α and diagonals of length D. The distances of the zeros from x are r and ρ, the distances from y are s and σ, and the angles subtended by the interval of integration are f and φ. By simple analytic geometry,

$$
\begin{aligned}
a\alpha L^2 &= a\gamma + c\alpha - \tfrac{1}{2}b\beta - \tfrac{1}{2}d\delta, & a\alpha D^2 &= a\gamma + c\alpha - \tfrac{1}{2}b\beta + \tfrac{1}{2}d\delta,\\
ar^2 &= ax^2 + bx + c, & \alpha\rho^2 &= \alpha x^2 + \beta x + \gamma,\\
as^2 &= ay^2 + by + c, & \alpha\sigma^2 &= \alpha y^2 + \beta y + \gamma,\\
ars\cos f &= axy + \tfrac{1}{2}b(x+y) + c, & \alpha\rho\sigma\cos\varphi &= \alpha xy + \tfrac{1}{2}\beta(x+y) + \gamma.
\end{aligned}
\tag{35}
$$

The validity condition (33) reduces to $\cos f+\cos\varphi \geqq 0$ or equivalently

$$f+\varphi \leqq \pi, \tag{36}$$

while (34) becomes

$$\int_y^x (at^2+bt+c)^{-1/2}(\alpha t^2+\beta t+\gamma)^{-1/2}\,dt = 2(a\alpha)^{-1/2}R_F(T^2, T^2-L^2, T^2-D^2),$$
$$(x-y)T = r\sigma + s\rho. \tag{37}$$

Consider the intersection z of the diagonals of the quadrilateral and the real axis (Fig. 1). If we choose $y = z$, the line segments of length s and σ in Fig. 2 lie on a diagonal (with $s+\sigma = D$), and f and φ are two angles of a triangle. Then, as $x \to \infty$, $f+\varphi$ increases monotonically to the limiting value π. We conclude that (36) is satisfied if $z \leqq y < x \leqq \infty$, and likewise if $-\infty \leqq y < x \leqq z$. In summary, the validity condition (33) or (36) is satisfied if the open interval of integration does not contain z. If it does contain z, the condition will be satisfied only if the interval of integration is finite and sufficiently short. This conclusion can of course be verified algebraically by using (33) and the formula

$$z = \frac{-(b\delta+\beta d)}{2(a\delta+\alpha d)}, \tag{38}$$

which results from considering similar triangles in Fig. 1.

If (33) is not satisfied, we can split the integral into two parts,

$$\int_y^x = \int_y^z + \int_z^x, \qquad y < z < x, \tag{39}$$

and evaluate each part by (34) or (37). In the second part the value z of the lower limit implies $s+\sigma = D$ and $s/\sigma = \alpha d/a\delta$, whence

$$s = \frac{\alpha dD}{a\delta+\alpha d}, \quad \sigma = \frac{a\delta D}{a\delta+\alpha d}, \quad (x-z)T = \frac{a\delta r+\alpha d\rho}{a\delta+\alpha d}D. \tag{40}$$

In the first part the value z of the upper limit implies

$$r = \frac{\alpha dD}{a\delta+\alpha d}, \quad \rho = \frac{a\delta D}{a\delta+\alpha d}, \quad (z-y)T = \frac{a\delta s+\alpha d\sigma}{a\delta+\alpha d}D. \tag{41}$$

Since $r/x \to 1$ and $\rho/x \to 1$ as $x \to \infty$, it is clear from (40) that $T \to D$ as $x \to \infty$ in the second part, and likewise $T \to D$ as $y \to -\infty$ in the first part. Therefore,

$$\begin{aligned}
&\int_z^\infty (at^2+bt+c)^{-1/2}(\alpha t^2+\beta t+\gamma)^{-1/2}\,dt \\
&\quad = \int_{-\infty}^z (at^2+bt+c)^{-1/2}(\alpha t^2+\beta t+\gamma)^{-1/2}\,dt \\
&\quad = 2(a\alpha)^{-1/2}R_F(0, D^2-L^2, D^2) \\
&\quad = 2R_F(0, d\delta, a\gamma+c\alpha-\tfrac{1}{2}b\beta+\tfrac{1}{2}d\delta).
\end{aligned} \tag{42}$$

The point z, which does not figure in previous treatments of this problem, divides the real line so that the integrals over the two half-lines are equal. The value of the integral over the whole real line can easily be checked by closing the contour with a large semicircle in the upper half-plane, deforming the contour so that it follows the edges of a cut joining the two zeros in the upper half-plane, and evaluating the integral along an edge of the cut by [6, (T.1), (3.6)].

Note added in proof. A direct proof of (13) by change of integration variable is given in [8].

REFERENCES

[1] M. ABRAMOWITZ AND I. STEGUN, EDS., *Handbook of Mathematical Functions*, U.S. Government Printing Office, Washington, D.C., 1964.

[2] P. F. BYRD AND M. D. FRIEDMAN, *Handbook of Elliptic Integrals for Engineers and Scientists*, 2nd ed., Springer-Verlag, Berlin, 1971.

[3] B. C. CARLSON, *Normal elliptic integrals of the first and second kinds*, Duke Math. J., 31 (1964), pp. 405–419.

[4] ———, *On computing elliptic integrals and functions*, J. Math. and Phys., 44 (1965), pp. 36–51.

[5] A. ERDÉLYI, W. MAGNUS, F. OBERHETTINGER AND F. G. TRICOMI, *Higher Transcendental Functions*, McGraw-Hill, New York, 1953.

[6] W. J. NELLIS AND B. C. CARLSON, *Reduction and evaluation of elliptic integrals*, Math. Comp., 20 (1966), pp. 223–231.

[7] D. G. ZILL AND B. C. CARLSON, *Symmetric elliptic integrals of the third kind*, Ibid., 24 (1970), pp. 199–214.

[8] B. C. CARLSON, *Short proofs of three theorems on elliptic integrals*, submitted.

SIAM J. MATH. ANAL.
Vol. 8, No. 2, April 1977

THE NUMERICAL-VALUED FOURIER TRANSFORM IN THE TWO-SIDED OPERATIONAL CALCULUS*

RAIMOND A. STRUBLE†

Abstract. The classical and distributional Fourier transform is extended to the ultimate setting in which it can be considered to be numerical-valued. It is extended as a ring isomorphism *onto* the ring of all measurable and finite almost everywhere functions under ordinary (pointwise) addition and multiplication of functions. The essential technique needed for this extension is the familiar algebraic procedure (first applied by Mikusiński to the operational calculus in the late 40's) of imbedding a given ring in a larger ring of fractions, the denominators being nondivisors of zero. Where Mikusiński's application resulted in a field of one-sided operators, the present application results in a ring of two-sided operators. Beyond this, only classical Fourier analysis is needed, though the extension of the latter to distributions is very useful in identifying many of the operators.

A descriptive subtitle for this paper would be *Basic definitions and theorems*, with applications to be considered later. For reasons of motivation and practical emphasis, the operational calculus is developed in a slightly more restrictive setting where the Fourier transforms are continuous almost everywhere.

1. Introduction. As is well-known, Mikusiński obtained [2] certain generalized functions by considering an algebraic field of fractions for the convolution ring of continuous functions on the half-line $[0, \infty)$. The elements of the field are called operators and provide for a one-sided operational calculus which possesses all of the advantages of rigor supplied by the Laplace transform method and none of the limitations imposed by the underlying analysis of the latter.

Recently [7], [8], the writer has constructed a two-sided operational calculus using the same algebraic technique, but in a setting for which the field of operators (called exponential operators) becomes isomorphic to a field of functions. These functions turn out to be meromorphic in (various) neighborhoods of the real axis $\mathscr{R}$ of the complex plain C, and one can use the classical analytic function theory and arithmetic in their study. The isomorphism introduced is an extension of the classical Fourier transform and agrees with the distributional Fourier transform of Schwartz [4] on a certain subclass of operators which are tempered distributions. The setting allows for all of the distributional Laplace transform theory associated with analytic functions, such as in [3], [9], [10], and considerably more.

In this paper we exploit the Fourier transform technique of Schwartz and the algebraic technique of Mikusiński even more fully and construct a large ring of two-sided operators which includes the field of exponential operators, the ring of integrable distributions, and considerably more. We are able to further extend the Fourier transform so that it becomes a ring isomorphism onto the ring $\mathscr{F}$ of all ordinary functions which are continuous almost everywhere. In fact, we can just as easily extend the Fourier transform as a ring isomorphism onto the ring of all functions which are measurable and finite-valued almost everywhere, which we illustrate in the final section of this paper. The latter extension is done here mainly for the sake of generalization. However, this last ring (notably) does contain *all*

* Received by the editors September 4, 1975, and in revised form January 2, 1976.
† Department of Mathematics, North Carolina State University, Raleigh, North Carolina 27607.

regular ultradistributions and appears to represent the ultimate setting in which the Fourier transformation can be used to exploit the arithmetic of ordinary functions. This aspect of the work may be the most interesting development at this time. The increased utility of the ring F of operators constructed here (or of its extension constructed in § 5), say in comparison with other two-sided operator calculi, has yet to be assessed, and here we only undertake to establish the framework for such an assessment and to provide for (hopefully) new applications. The development is presented in the one-dimensional case, but applies equally well in the higher-dimensional cases.

In § 2, we review some standard notation and terminology concerning distributions. Here, and throughout the paper, the readers are assumed to be familiar with elementary aspects of distributions and their Fourier transforms (such as given in [1], [9]). In § 3, we consider some preliminary results which come principally from recent work in references [3] and [6]. This section is used mainly for motivation, but includes important definitions of rings of distributions and related functions. The two-sided operational calculus is developed mostly in § 4. Important definitions of the rings of operators and related functions and of the extended Fourier transform between them are given. A few, perhaps noteworthy, theorems and corollaries are proved. Here the emphasis is mostly on the ring of functions $\mathscr{F}$ and what it means to the operational calculus. Some topological and convergence concepts are considered in § 4. Finally in § 5, the present development is extended to the ring of all measurable and finite almost everywhere functions. This final section also shows that distributions can be avoided entirely and that the final ring of operators, as well as that of their Fourier transforms, is isomorphic with a ring of quotients of *ordinary* functions under addition and convolution. Thus, only the Mikusiński (algebraic) technique is really needed in all of this. However, it is very convenient to be able to interpret many of the operators (fractions) as distributions.

2. Notation and terminology. Much of the notation to be used in this paper is standard; some of it is not. Throughout the paper we let R denote the real line, C denote the complex plane and $\mathscr{R}$ denote the real C-axis; the latter to be distinguished from R. We let $\mathscr{D} = \mathscr{D}(R)$ denote the space of infinitely differentiable test functions $\phi(t) = \phi$ of the real variable $t \in R$, with compact supports, together with the standard topology given by Schwartz [4]. We let $Z(C) = Z$ denote the space of entire functions $\psi(z) = \psi$ of the complex variable $z = \omega + i\rho \in C$, which are the Fourier transforms of the elements of $\mathscr{D}$, together with the standard topology for which the Fourier transform from $\mathscr{D}$ onto Z, (see [7], [8]),

$$\phi(t) \mapsto \tilde{\phi}(z) = \int_{-\infty}^{\infty} e^{-izt}\phi(t)\,dt, \tag{1}$$

and the inverse Fourier transform from Z onto $\mathscr{D}$,

$$\tilde{\phi}(z) \mapsto \phi(t) = \frac{1}{2\pi}\int_{\mathrm{Im}\, z = \text{const.}} e^{izt}\tilde{\phi}(z)\,dz = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t}\tilde{\phi}(\omega)\,d\omega, \tag{2}$$

become topological vector space isomorphisms. We recall that Z is the collection

of entire functions $\tilde{\phi}$ which satisfy families of inequalities of the form

$$|z^k\tilde{\phi}(z)| \leqq c_{\phi,k}\, e^{a_\phi|\mathrm{Im}\, z|}, \qquad k = 1, 2, \cdots, \tag{3}$$

for positive constants $c_{\phi,k}$ and a_ϕ (depending upon $\tilde{\phi}$, k and $\tilde{\phi}$ respectively).

The topological duals of these spaces are denoted, as usual, by $\mathscr{D}'(R) = \mathscr{D}'$ and $Z'(C) = Z'$, and an element $f(t) = f \in \mathscr{D}'$ is called a distribution, and an element $g(\omega) = g \in Z'$ is called an ultradistribution. The value at $\phi \in \mathscr{D}$ of the distribution $f \in \mathscr{D}'$ is denoted by $\langle f, \phi\rangle = \langle f(t), \phi(t)\rangle$, and the value at $\psi \in Z$ of the ultradistribution $g \in Z'$ is denoted by $\langle g, \psi\rangle = \langle g(\omega), \psi(\omega)\rangle$. Observe that each entire function $\psi \in Z$ is determined by its restriction to the real C-axis $\mathscr{R} = -\infty < \omega < \infty$, and so we may (and shall) consider an ultradistribution g as a generalized function on $\mathscr{R}$.

We shall make use of the well-known subspaces of tempered distributions $\mathscr{S}'(R)$ and $\mathscr{S}'(\mathscr{R})$, (on the two real lines R and $\mathscr{R}$), as well as the not so well-known subspace of integrable distributions $\mathscr{B}_0'(R) = \mathscr{B}_0'$ [3]. We recall that a distribution f is tempered iff (if and only if) there exist positive integers M and N such that

$$|\langle f, \phi\rangle| \leqq M \sup_t |(1+t^{2N})\phi^{(N)}(t)|, \qquad t \in R, \tag{4}$$

holds for all $\phi \in \mathscr{D}$, and is integrable iff there exists a positive integer K such that

$$|\langle f, \phi\rangle| \leqq K \max_j \sup_t |\phi^{(j)}(t)|, \qquad t \in R, \quad 0 \leqq j \leqq K, \tag{5}$$

holds for all $\phi \in \mathscr{D}$ (see [3]). Here $\phi^{(j)}$ denotes the ordinary jth order derivative of ϕ. Clearly, $\mathscr{B}_0' \subset \mathscr{S}'$, and we shall see (Prop. 7) that if $f \in \mathscr{S}'$, then $f(t)/(1+t^{2N}) \in \mathscr{B}_0'$ for some integer $N \geqq 0$.

If $f(t) = f$ is a distribution, then $\tilde{f}(\omega) = \tilde{f}$ will denote its Fourier transform (as an ultradistribution) which satisfies the Parseval relation

$$2\pi\langle f, \check{\phi}\rangle = \langle \tilde{f}, \tilde{\phi}\rangle \tag{6}$$

for all $\tilde{\phi} \in Z$, where $\check{\phi}(t) = \phi(-t)$. With this definition, the Fourier transform $f \mapsto \tilde{f}$ becomes a vector space topological isomorphism from $\mathscr{D}'$ onto Z' (with respect to their weak topologies, say), and satisfies

$$\widetilde{f * \phi} = \tilde{f}\tilde{\phi} \qquad \text{(with } \tilde{\phi} \text{ as a multiplier on } Z') \tag{7}$$

for all $f \in \mathscr{D}'$ and $\phi f \in \mathscr{D}$, where $*$ denotes, as usual, convolution. We recall that $f \in \mathscr{S}'(R)$ iff $\tilde{f} \in \mathscr{S}'(\mathscr{R})$.

Finally, if n is a positive integer, then U_n will denote the dilatation transformation on $\mathscr{D}'$ defined by $U_n f(t) = nf(nt)$, where $\langle nf(nt), \phi(t)\rangle = \langle f(t), \phi(t/n)\rangle$, and a function will be said to be *slowly increasing* if it is bounded by a polynomial. All algebraic rings in this paper are commutative, and are complex vector spaces as well.

3. Preliminary results and definitions. We shall need several results from references [3] and [6].

PROPOSITION 1. *If $f \in \mathcal{D}'$, and if its Fourier transform $\tilde{f}$ is a regular ultradistribution which is slowly increasing and is continuous at a point $\omega_0 \in \mathcal{R}$, then*

$$\lim_{n\to\infty} U_n e^{-i\omega_0 t} f(t) = \tilde{f}(\omega_0)\,\delta(t), \tag{8}$$

where the convergence is in $\mathcal{D}'$ and $\delta(t)$ is the delta function, of course.

This proposition was stated in [6] for everywhere continuous $\tilde{f}$, but the proof applies equally well to the above case.

PROPOSITION 2. *If $f \in \mathcal{B}_0'$, then $\tilde{f}$ is continuous on $\mathcal{R}$ and is slowly increasing.*

This proposition was only stated in [6], but follows readily from the work in [3]. It appears to be unknown whether or not the converse holds. Propositions 1 and 2 have the following obvious corollary.

COROLLARY 1. *If $f \in \mathcal{B}_0'$, then $\lim_{n\to\infty} U_n e^{-i\omega t} f(t) = \tilde{f}(\omega)\,\delta(t)$, for every $\omega \in \mathcal{R}$.*

The following was proved in [3].

PROPOSITION 3. *If $f_1, f_2 \in \mathcal{B}_0'$, then the convolution $f_1 * f_2$ exists and belongs to $\mathcal{B}_0'$.*

And finally, the following was proved in [6].

PROPOSITION 4. *If $f_1, f_2 \in \mathcal{B}_0'$, then $\widetilde{f_1 * f_2} = f_1 f_2$, with juxtaposition denoting the ordinary multiplication of the two functions on the right.*

From these results we are led rather naturally to consider the collection $\mathcal{T}$ of all functions g which are defined, are continuous and are slowly increasing a.e. (almost everywhere) on $\mathcal{R}$. The latter, of course, means that such a g is defined, is continuous and is bounded by a polynomial on $\mathcal{R}$ except for possibly a subset of measure zero. As usual, two such functions which agree a.e. will be identified. Each of these functions may be (and shall be) considered as a regular ultradistribution (in fact, a regular tempered ultradistribution in $\mathcal{S}'(\mathcal{R})$). Moreover, by Proposition 1, each of these functions is the Fourier transform of a distribution which satisfies the limit condition in (8) for almost every ω_0 of $\mathcal{R}$. Also by Proposition 2, if $f \in \mathcal{B}_0'$, then $\tilde{f} \in \mathcal{T}$. Let us, therefore, introduce the following definition for functions on $\mathcal{R}$.

DEFINITION 1. A function g, defined a.e. on $\mathcal{R} = -\infty < \omega < \infty$, is called a *tempered* function if it is continuous and slowly increasing a.e. on $\mathcal{R}$. Two such functions will be identified if they agree a.e. on $\mathcal{R}$. The collection of all tempered functions will be denoted by $\mathcal{T}$, and will be considered a ring under ordinary addition and multiplication of functions. The subcollection of $\mathcal{T}$ consisting of those functions which are nonzero a.e. on $\mathcal{R}$ will be denoted by $\mathcal{T}_0$.

It is easy to see that the subcollection $\mathcal{T}_0$ is closed under multiplication and, moreover, constitutes a subcollection of the nondivisors of zero in the ring $\mathcal{T}$. This means that $g_1 g_2 = 0$ for $g_1 \in \mathcal{T}_0$ and $g_2 \in \mathcal{T}$, implies that $g_2 = 0$. The divisors of zero in $\mathcal{T}$ are precisely those functions which vanish on some open interval. Since there are (even) continuous functions which vanish on sets of positive measure but do not vanish on any open intervals (consider a function which vanishes on a Cantor set with positive measure and is nonzero on its compliment, for example), $\mathcal{T}_0$ does not contain all of the nondivisors of zero in $\mathcal{T}$. It will become apparent later why we are not interested in the collection of all nondivisors of zero in $\mathcal{T}$.

The following is the companion definition for distributions on R.

DEFINITION 2. A distribution f on $R = -\infty < t < \infty$ is called a *$\mathcal{T}$-tempered* distribution if its Fourier transform $\tilde{f}$ is a tempered function. The collection of all $\mathcal{T}$-tempered distributions will be denoted by T, and will be considered a ring under ordinary addition of distributions and *extended* convolution of distributions defined by

$$f_3 = f_1 \circledast f_2, \quad \text{whenever } \tilde{f}_3 = \tilde{f}_1 \tilde{f}_2. \tag{9}$$

The subcollection of T consisting of those distributions f for which $\tilde{f} \in \mathcal{T}_0$ will be denoted by T_0.

By Proposition 2, we have $\mathcal{B}_0' \subset T$, and by Proposition 4, it follows that the operation $\circledast$ is an extension to all of T of ordinary convolution $*$ of distributions in $\mathcal{B}_0'$. It is unknown whether or not $\circledast$ is ordinary convolution in T. In any case, T_0 consists of a subcollection of nondivisors of zero in T with respect to the operation $\circledast$. It contains all the distributions with compact supports and, more generally, all the distributions whose Fourier transforms are analytic on $\mathcal{R}$ and are slowly increasing. These include (suitable shifts of) all the Laplace transformable distributions of Schwartz [5].

By these definitions, the Fourier transform becomes a ring isomorphism from T onto $\mathcal{T}$, and henceforth, we shall usually denote an element (function) of $\mathcal{T}$ by $\tilde{f}$, meaning that it is the Fourier transform of an element (distribution) f of T. The Fourier transform can be defined directly on T using (8).

PROPOSITION 5. *If $f \in T$, then for any $\phi \in \mathcal{D}$ with $\phi(0) \neq 0$,*

$$\tilde{f}(\omega) = \frac{1}{\phi(0)} \lim_{n \to \infty} \langle U_n e^{-i\omega t} f(t), \phi(t) \rangle \quad \text{for a.e. } \omega \in \mathcal{R}.$$

(*Here a.e. reads, almost every.*)

Proof. Since $f \in T$, $\tilde{f}$ is continuous a.e. on $\mathcal{R}$ and so (8) holds for a.e. $\omega_0 \in \mathcal{R}$. Applying these distributions to the test function ϕ yields Proposition 5.

The following is a well-known result [4].

PROPOSITION 6. *If $\tilde{g} \in \mathcal{S}'(\mathcal{R})$, then there exists an integer $N \geqq 0$ and an $\tilde{f} \in \mathcal{T}$ such that $\tilde{g} = \tilde{f}^{(2N)}$, where the latter is the (ultra) distributional $2N$-th derivative of $\tilde{f}$.*

Thus every tempered distribution on $\mathcal{R}$ is a finite order derivative of a tempered function. Through the Fourier transform (for which we have $\widehat{t^{2N} f(t)} = -\tilde{f}^{(2N)}$), it follows that every tempered distribution h on R can be expressed in the form $h(t) = t^{2N} f(t)$ for some integer $N \geqq 0$ and some $\mathcal{T}$-tempered distribution f. This means that $h(t)/(1+t^{2N}) = f(t) - f(t)/(1+t^{2N}) \in T$. We can improve upon this last result as follows.

PROPOSITION 7. *If $h \in \mathcal{S}'(R)$, then $h(t)/(1+t^{2N}) \in \mathcal{B}_0'$ for some integer $N \geqq 0$.*

Proof. Because $h \in \mathcal{S}'(R)$, it satisfies (4) for suitable integers M and N. Hence for any $\phi \in \mathcal{D}$,

$$\begin{aligned} |\langle h(t)/(1+t^{2N}), \phi(t) \rangle| &= |\langle h(t), \phi(t)/(1+t^{2N}) \rangle| \\ &\leqq M \sup_t |(1+t^{2N})[\phi(t)/(1+t^{2N})]^{(N)}| \\ &\leqq M \sum_{j=0}^{N} M_j \sup_t |\phi^{(N-j)}(t)| \leqq L \max_j \sup_t |\phi^{(j)}(t)|, \end{aligned}$$

where $M_j = \sup_t \binom{N}{j}(1+t^{2N})[1/(1+t^{2N})]^{(j)}$ and $L = (N+1)M \max_j M_j$. Thus if K is any integer greater than L and N, we have

$$|\langle h(t)/(1+t^{2N}), \phi(t)\rangle| \leqq K \max_j \sup_t |\phi^{(j)}(t)|, \qquad t \in R, \quad 0 \leqq j \leqq K,$$

which means (see (5)) that $h(t)/(1+t^{2N}) \in \mathcal{B}_0'$.

This proposition applies, in particular, to $\mathcal{T}$-tempered distributions. Hence we have the following corollary.

COROLLARY 2. *If $f \in T$ (alternatively, if $\tilde{f} \in \mathcal{T}$), then there exists an integrable distribution g and an integer $N \geqq 0$ such that $f(t) = (1+t^{2N})g(t)$ (alternatively, such that $\tilde{f}(\omega) = \tilde{g}(\omega) - \tilde{g}^{(2N)}(\omega)$).*

Thus the elements of T are obtainable from the elements of $\mathcal{B}_0'$ through multiplication by powers of t, and the elements of $\mathcal{T}$ are obtainable from the continuous members of $\mathcal{T}$ by (generalized) differentiations. It seems likely that one differentiation should suffice since one (ordinary) integration of an element of $\mathcal{T}$ results in a continuous function which is differentiable a.e. Such would be the case if the converse of Proposition 2 were to be true.

Examples. The distribution p.v. $1/t$ is defined by

$$\left\langle \text{p.v.}\frac{1}{t}, \phi(t)\right\rangle = \lim_{\varepsilon \to 0^+} \int_{-\infty}^{-\varepsilon} + \int_{\varepsilon}^{\infty} \frac{\phi(t)}{t}\, dt.$$

It belongs to T (in fact, T_0), but it does not belong to $\mathcal{B}_0'$. Its Fourier transform is the step function $-i\pi$ sign ω. The generalized derivative of the latter is $-2i\pi\delta(\omega)$, which belongs to $\mathcal{S}'(\mathcal{R})$, but does not belong to $\mathcal{T}$. The (ordinary) integral of it, $-i\pi|\omega|$ is continuous and is the Fourier transform of the finite part Fp $-i/t^2$ (see [9]) which belongs to $\mathcal{B}_0'$. The Fourier transforms of the delta function $\delta(t)$, the differentiation operator s (equivalently, $\delta^{(1)}(t)$) and the translation operator $e^{\lambda s}$ (equivalently, $\delta(t-\lambda)$) are, respectively, the functions 1, $i\omega$ and $e^{i\lambda\omega}$.

4. Construction of the two-sided operator calculus. We shall now consider the collection F of (formal) fractions g/f with $g \in T$ and $f \in T_0$, where as usual, $g_1/f_1 = g_2/f_2$ iff $g_1 * f_2 = g_2 * f_1$. Because T_0 is closed under $*$ and consists of nondivisors of zero in T, these fractions can be added and multiplied, just as ordinary numerical fractions can be, but where multiplication becomes extended convolution. Under these operations, F becomes a ring; its elements are called operators, following Mikusiński's example. The ring T itself may (and shall) be considered as a subring of F by identifying each $g \in T$ with the fraction $(g * f)/f$ for any $f \in T_0$. Thus all $\mathcal{T}$-tempered distributions are operators.

Similarly, we shall also consider the collection $\mathcal{F}$ of functions of the form $\tilde{g}/\tilde{f}$ with $\tilde{g} \in \mathcal{T}$ and $\tilde{f} \in \mathcal{T}_0$. Since the denominators $\tilde{f}$ can vanish only on sets of measure zero, such functions are defined and are continuous a.e. on $\mathcal{R}$. $\mathcal{F}$ becomes a ring under ordinary addition and multiplication of functions and contains $\mathcal{T}$ as a subring, since the function which is identically 1 belongs to $\mathcal{T}_0$.

It is clear that the two rings of fractions F and $\mathcal{F}$ are isomorphic via the mapping which sends $g/f \in F$ to $\tilde{g}/\tilde{f} \in \mathcal{F}$. This mapping is an extension to F of the

distributional Fourier transform on T. We shall formalize these important considerations with our principal definition and our first theorem.

DEFINITION 3. The ring of all fractions $x = g/f$ with $g \in T$ and $f \in T_0$ is denoted by F and its elements are called *operators*. The fractions are identified, added and multiplied, as ordinary numerical fractions are, with the operations corresponding to addition and extended convolution (Def. 2) in T. The ring of all functions $\tilde{x} = \tilde{g}/\tilde{f}$ with $\tilde{g} \in \mathcal{T}$ and $\tilde{f} \in \mathcal{T}_0$, under ordinary addition and multiplication of functions is denoted by $\mathcal{F}$. These functions are defined and are continuous a.e. on $\mathcal{R}$, and are identified if they agree a.e. on $\mathcal{R}$. The mapping $x = g/f \mapsto \tilde{g}/\tilde{f} = \tilde{x}$ from F onto $\mathcal{F}$ is called the *Fourier transform*.

THEOREM 1. *The Fourier transform, as defined in Definition* 3, *is an extension to F of the distributional Fourier transform on the subspace T of $\mathcal{T}$-tempered distributions. It is a ring isomorphism from the ring F onto the ring $\mathcal{F}$.*

We can now explain why we do not work with the ring of all nondivisors of zero in $\mathcal{T}$, as might be expected. For to do so would require that we treat the fractions $\tilde{g}/\tilde{f}$ *formally* (as we must do in any case with the distributions), rather than numerically as functions. For example, if $\tilde{f}$ were to vanish on a set of positive measure, where $\tilde{g}$ does also, then as a function, the fraction $\tilde{g}/\tilde{f}$ would be indeterminant on a set of positive measure and the formal ring operations would not correspond to the ordinary arithmetical ones. We have selected, as denominators, precisely those elements in $\mathcal{T}$ for which the arithmetical operations are preserved.

It is, of course, easier to examine the arithmetical properties of the ring $\mathcal{F}$ of functions than it is the convolution properties of the ring F of operators. This is why we introduce the Fourier transform in the first place. However, the operators are the objects of primary concern in the operational calculus and we shall consider them more fully in (hopefully) subsequent papers. The following theorem dispenses with the need for our fractional notation in $\mathcal{F}$.

THEOREM 2. *A function $\tilde{x}$ belongs to the ring $\mathcal{F}$ iff $\tilde{x}$ is continuous a.e. on $\mathcal{R}$.*

Proof. We need only establish the "if" part of this theorem. So suppose $\tilde{x}(\omega)$ is continuous a.e. on $\mathcal{R}$. Then the function $\tilde{f}(\omega) = 1/(1+|\tilde{x}(\omega)|)$ is also, and moreover, $\tilde{f}(\omega)$ is bounded and nonzero a.e. on $\mathcal{R}$. Furthermore, the product $\tilde{x}(\omega)\tilde{f}(\omega)$ is continuous and bounded a.e. on $\mathcal{R}$. Thus $\tilde{x}(\omega) = \tilde{x}(\omega)\tilde{f}(\omega)/\tilde{f}(\omega) \in \mathcal{F}$, since $\tilde{x}(\omega)\tilde{f}(\omega) \in \mathcal{T}$ and $\tilde{f}(\omega) \in \mathcal{T}_0$.

Two immediate corollaries of this theorem are themselves of some interest.

COROLLARY 3. *The ring $\mathcal{F}$ contains all regular ultradistributions defined by locally (Riemann) integrable functions.*

Of course, many of these ultradistributions are not tempered, so we have certainly enlarged the collection of ultradistributions at our disposal. We shall, of course, identify those distributions with the corresponding operators whose Fourier transforms are regular ultradistributions in $\mathcal{F}$. These distributions are given by $\langle f, \check{\phi} \rangle = (1/(2\pi)) \int_{-\infty}^{\infty} \tilde{f}(\omega)\tilde{\phi}(\omega)\, d\omega$, whenever $\tilde{f}$ is a regular ultradistribution.

It is always of interest to identify those elements of a ring which are invertible. Such elements are usually called units, and because of Theorem 2, are easily identified in $\mathcal{F}$.

COROLLARY 4. *A function $\tilde{x}$ in the ring $\mathcal{F}$ is a unit iff $\tilde{x}$ is nonzero a.e. on $\mathcal{R}$.*

Though this last corollary appears to be rather mundane here, it is of considerable importance for the ring F of operators. Therefore we introduce an appropriate definition.

DEFINITION 4. The subset of all functions $\tilde{x}$ in $\mathcal{F}$ which are nonzero a.e. on $\mathcal{R}$ is denoted by $\mathcal{U}$. The elements of $\mathcal{U}$ are called *units*. The subset of all operators x in F for which $\tilde{x} \in \mathcal{U}$ is denoted by U. The elements of U are called *unitary* operators.

Examples. Suppose that $H(z)$ is a function which for some $b>0$ is meromorphic in the neighborhood $N_b = \{z : |\operatorname{Im} z| < b\}$ of the real C-axis $\mathcal{R}$. Then for any real ρ with $|\rho| < b$, the function $\omega \mapsto H(\omega + i\rho)$ (all $\omega \in \mathcal{R}$) is an element of $\mathcal{F}$. Moreover, every such element is a unit. In particular, if $H(z) = Q(z)/P(z)$ is a rational function with P and Q polynomials, then for any real ρ the function $\omega \mapsto Q(\omega + i\rho)/P(\omega + i\rho)$ is an element of $\mathcal{F}$; in fact, it is an element of $\mathcal{T}$, whenever ρ is chosen so that the poles of $H(z)$ are avoided, since it is then slowly increasing. If this is the case for $\rho = 0$, then $H(\omega)$ is the Fourier transform of a distributional solution $f(t)$ of the differential equation $P(-id/dt)f(t) = Q(-id/dt)\,\delta(t)$, since the Fourier transform of this equation is $P(\omega)\tilde{f}(\omega) = Q(\omega)$, and $P(\omega)$ is a unit. For other values of ρ, $H(\omega + i\rho)$ is the Fourier transform of the shift $e^{\rho t}f(t)$ of the distribution $f(t)$. Thus, $H(z)$ is the Laplace transform of $f(t)$ (rotated 90°, of course), where $f(t)$ is a solution of this differential equation in T. A two-sided operational calculus was developed recently [7] in which the Fourier transforms of the operators (called exponential operators) are meromorphic functions in various neighborhoods N_b. Hence these operators all belong to the ring F and form a subfield in the subset U of unitary operators.

The first example above can be generalized immediately to the following.

COROLLARY 5. *Let $P(s)$ be any polynomial in the differentiation operator s with complex coefficients. Then $P(s)$ is a unitary operator and for any operator $y \in F$, the fraction $x = y/P(s)$ is the unique solution of the equation $P(s) * x = y$ in F. If $y \in T$, and if $P(z)$ has no pure imaginary zeros, then $x \in T$.*

Proof. The proof is trivial, since $\widetilde{P(s)} = P(i\omega)$ is a unit in $\mathcal{F}$ and $\tilde{y}(\omega)/P(i\omega)$ is slowly increasing, if $y \in T$ and $P(i\omega) \neq 0$ for all ω.

The second part of this corollary, of course, yields a distributional solution x of the differential equation $P(d/dt)x(t) = y(t)$. The arithmetic (operator method) is overwhelmingly simpler than the corresponding analysis, but we obtain less information. However, we can use this simple arithmetic just as readily to treat differential equations of *infinite* order.

THEOREM 3. *Let $\mathcal{P}$ be a function which is analytic on the imaginary C-axis. Then the operator $\mathcal{P}(s)$, defined as the Fourier inverse of the function $\mathcal{P}(i\omega)$, is unitary, and for any operator $y \in F$, the fraction $x = y/\mathcal{P}(s)$ is the unique solution of the equation $\mathcal{P}(s) * x = y$ in F.*

The operator $\mathcal{P}(s)$ in this theorem is, of course, an infinite order differential operator unless $\mathcal{P}$ is a polynomial. For example, if $\mathcal{P}$ is an entire function with power series $\sum_{j=0}^{\infty} c_j z^j$, then $\mathcal{P}(s) = \sum_{j=0}^{\infty} c_j s^j$. If y is a distribution in T and if $1/\mathcal{P}(i\omega)$ is an ultradistribution in $\mathcal{T}$ (as in the above example), then x is a distribution and satisfies $\sum_{j=0}^{\infty} c_j x^{(j)} = y$. Consider the interesting example, $\mathcal{P}(z) = e^{-z^2}$. We note in passing that $s * x = g^{(1)}/f$, whenever $x = g/f$ with $g \in T$, $f \in T_0$, and it is appropriate to define the fraction $g^{(1)}/f$ as the *derivative* $x^{(1)}$ of the

operator x. Then $\widetilde{x^{(1)}} = i\omega\tilde{x}(\omega)$ holds for a.e. $\omega \in \mathcal{R}$ and we can write $P(s) \ast x = P(d/dt)x = y$ in Corollary 5.

It would be of interest to consider some topological and convergence notions in $\mathcal{F}$. One possible topology for $\mathcal{F}$ is a metric one defined by the family of chordal pseudo-metrics

$$\rho_k(\tilde{x}, \tilde{y}) = \text{ess. sup}_{-k \leqq \omega \leqq k} \frac{|\tilde{x}(\omega) - \tilde{y}(\omega)|}{\sqrt{1+|\tilde{x}(\omega)|^2}\sqrt{1+|\tilde{y}(\omega)|^2}}, \qquad k = 1, 2, \cdots. \tag{10}$$

We recall that the ess. sup $\tilde{f}$ is the infinimum of the positive numbers M for which the set $\{\omega : |\tilde{f}(\omega)| \geqq M\}$ has measure zero. If there are no such numbers, then ess. sup $\tilde{f} = \infty$. For functions which are essentially bounded on compact sets, this is the topology of uniform convergence a.e. on compact sets.

The following appears to be a rather rouine result. Its corollary shows differently.

THEOREM 4. *Let $\tilde{x} \in \mathcal{F}$. Then with respect to the metric topology* (10),

$$\lim_{n\to\infty} \tilde{x}\left(\frac{\omega}{n} + \omega_0\right) = \tilde{x}(\omega_0) \quad \textit{for a.e. } \omega_0 \in \mathcal{R}. \tag{11}$$

In particular, this holds at every point ω_0 of continuity of $\tilde{x}$.

Proof. The proof is immediate from (10), since for every k, it is clear that

$$\lim_{n\to\infty} \rho_k\left(\tilde{x}\left(\frac{\omega}{n} + \omega_0\right), \tilde{x}(\omega_0)\right) = 0$$

if ω_0 is a point of continuity of $\tilde{x}$.

The following is the analogue for operators of Proposition 1.

COROLLARY 6. *Let x be an operator in F. Then with respect to the metric topology induced on F,*

$$\lim_{n\to\infty} U_n e^{-i\omega_0 t} x(t) = \tilde{x}(\omega_0)\delta(t) \quad \textit{for a.e. } \omega_0 \in \mathcal{R}. \tag{12}$$

Here we need to explain the notation and the terminology used. For each $x \in F$, we define

$$e^{-i\omega_0 t} x(t) = \frac{e^{-i\omega_0 t} g(t)}{e^{-i\omega_0 t} f(t)} \tag{13}$$

and

$$U_n x(t) = \frac{U_n g(t)}{U_n f(t)}, \tag{14}$$

where $x = g/f$, with $g \in T$ and $f \in T_0$. It is then easy to verify that these definitions extend the multipliers $e^{-i\omega_0 t}$ and the dilatations U_n to automorphisms on the ring F. Since the Fourier transform is one-to-one from F onto $\mathcal{F}$, we can transform the metric topology of $\mathcal{F}$ to a metric topology for F simply by defining the pseudo-metrics $\rho_k(x, y) = \rho_k(\tilde{x}, \tilde{y})$ in F. This topology is then said to be induced on F. The proof of Corollary 6 is simply the observation that (13) and (14) imply that $\widetilde{U_n e^{-i\omega_0 t} x(t)}(\omega) = \tilde{x}((\omega/n) + \omega_0)$ holds for every $x \in F$.

We also have an analogue of the numerical result in Proposition 5.

COROLLARY 7. *If $x \in F$, then for any $\phi \in \mathcal{D}$ with $\phi(0) \neq 0$,*

$$\tilde{x}(\omega_0) = \frac{1}{\phi(0)} \langle \lim_{n\to\infty} U_n e^{-\omega_0 t} x(t), \phi(t) \rangle \quad \text{for a.e. } \omega_0 \in \mathcal{R}.$$

Proof. This follows immediately from Corollary 6, where the limit is a distribution.

The next theorem gives a useful criteria for convergence in $\mathcal{F}$ with respect to the metric topology. The main difficulty associated with this topology is that the space $\mathcal{F}$ contains (essentially) unbounded functions. It would seem, therefore, that if, for example, the limit $\tilde{x}$ of a sequence is (essentially) bounded, then the nature of the convergence should be more readily understood than if $\tilde{x}$ is unbounded. This is part of the message conveyed by the following theorem. It says that a sequence $\mathcal{F}$ converges if it can be represented in fractional form so that its numerator and denominator sequences converge nicely to a fraction which is (essentially) nowhere indeterminant (see (15)). In such a case, the limit could be unbounded, but if it is bounded, then the denominator is (essentially) bounded away from zero.

THEOREM 5. *Suppose that a sequence $\{\tilde{x}_n\}$ of functions in $\mathcal{F}$ is such that $\tilde{x}_n = \tilde{g}_n/\tilde{f}_n$ with $\tilde{g}_n \in \mathcal{F}$, $\tilde{f}_n \in \mathcal{U}$ for all n, and that $\lim_{n\to\infty} \tilde{g}_n = \tilde{g}$, and $\lim_{n\to\infty} \tilde{f}_n = \tilde{f} \in \mathcal{U}$, with respect to the topology of uniform convergence (a.e.) on compact sets of $\mathcal{R}$. Then $\lim_{n\to\infty} \tilde{x}_n = \tilde{x} = \tilde{g}/\tilde{f}$, with respect to the metric topology* (10), *provided*

$$\operatorname*{ess.\,inf}_{-k\leq\omega\leq k} \sqrt{|\tilde{f}(\omega)|^2 + |\tilde{g}(\omega)|^2} > 0 \quad \text{for every } k. \tag{15}$$

Proof. Condition (15) means that for every k, there exists a positive M such that the set $\{\omega : |\omega| \leq k, \sqrt{|\tilde{f}(\omega)|^2 + |\tilde{g}(\omega)|^2} \leq M\}$ is of measure zero. Now for each k we have

$$\begin{aligned}\rho_k(\tilde{x}_n, \tilde{x}) &= \operatorname*{ess.\,sup}_{-k\leq\omega\leq k} \frac{|\tilde{f}(\omega)\tilde{g}_n(\omega) - \tilde{f}_n(\omega)\tilde{g}(\omega)|}{\sqrt{|\tilde{f}_n(\omega)|^2 + |\tilde{g}_n(\omega)|^2}\sqrt{|\tilde{f}(\omega)|^2 + |\tilde{g}(\omega)|^2}} \\ &\leq \operatorname*{ess.\,sup}_{-k\leq\omega\leq k} \frac{|\tilde{g}_n(\omega) - \tilde{g}(\omega)| + |\tilde{f}_n(\omega) - \tilde{f}(\omega)|}{\sqrt{|\tilde{f}(\omega)|^2 + |\tilde{g}(\omega)|^2}}.\end{aligned}$$

Hence it follows from (15) that $\lim_{n\to\infty} \rho_k(\tilde{x}_n, \tilde{x}) = 0$, since the sequences $\{\tilde{g}_n\}$ and $\{\tilde{f}_n\}$ converge to $\tilde{g}$ and $\tilde{f}$, uniformly on $[-k, k]$. Since this holds for every k, the conclusion of the theorem holds.

Examples. Let $\{\phi_n\}$ be a delta function sequence in $\mathcal{D}$. This means that the supports of all the members of the sequence are contained in some fixed compact subset of R and that $\lim_{n\to\infty} \phi_n * \phi = \phi$ with respect to the topology of uniform convergence on compact subsets of R, for every $\phi \in \mathcal{D}$. Then $\lim_{n\to\infty} \phi_n(t) = \delta(t)$ in $\mathcal{D}'$, hence the name for such a sequence. On the other hand, since $\widetilde{\phi_n * \phi} = \tilde{\phi}_n\tilde{\phi}$, the first limit statement means that $\lim_{n\to\infty} \tilde{\phi}_n\tilde{\phi} = \tilde{\phi}$ with respect to the topology of uniform convergence on compact subsets of $\mathcal{R}$, for every $\tilde{\phi} \in Z$. But this can be so only if $\lim_{n\to\infty} \tilde{\phi}_n$ also exists with respect to the latter topology. Thus it exists with respect to the metric topology of $\mathcal{F}$ and is, in fact, 1. Suppose that the members of

the sequence $\{\psi_n\}$ are continuous functions on R, which are all bounded by $e^{-|t|}$, and converge uniformly on compact subsets of R to ψ. Then these belong to $\mathcal{B}'_0$, and by essentially the same arguments, their Fourier transforms $\tilde{\psi}_n$ converge to $\tilde{\psi}$ with respect to the topology of uniform convergence on compact subsets of $\mathcal{R}$. Thus by Theorem 5, $\lim_{n\to\infty}(\psi_n/\phi_n)=\psi$ in T with respect to the metric topology, since (15) is certainly satisfied in this case.

The inverse Fourier transform does not materialize nearly so easily as the Fourier transform did in Corollary 6; it must be obtained as a "fractionalized" limit in $\mathcal{F}$. Let $\tilde{x}\in\mathcal{F}$ and $\tilde{x}=\tilde{g}/\tilde{f}$ with $\tilde{g}\in\mathcal{T}$, $\tilde{f}\in\mathcal{T}_0$. Then if $\tilde{\sigma}$ is a nonzero member of Z, the convolution integral

$$I(\tilde{g},\omega,t,n)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{it\xi}\tilde{g}(\xi)\tilde{\sigma}(\xi)\tilde{\sigma}\left(\omega-\frac{\xi}{n}\right)d\xi$$

exists, and for fixed values of t and n, is an infinitely differentiable function of ω which (it can be verified) belongs to $\mathcal{T}$. For fixed values of ω and n, it is an infinitely differentiable function of t. Since

$$\widetilde{(\sigma * g) * (U_n e^{i\omega t}\sigma(-t))}(\xi)=\tilde{g}(\xi)\tilde{\sigma}(\xi)\tilde{\sigma}\left(\omega-\frac{\xi}{n}\right),$$

by (2) this integral equals $(\sigma * g) * U_n e^{i\omega t}\sigma(-t)$.

It can be verified that for each fixed value of t,

$$\lim_{n\to\infty} I(\tilde{g},\omega,t,n)=(\sigma * g)(t)\tilde{\sigma}(\omega),$$

uniformly in ω on compact subsets of $\mathcal{R}$, and hence by Theorem 5, with respect to the metric topology of $\mathcal{F}$. On the other hand, for each fixed value of ω, this limit also exists with respect to the metric topology of F. This gives us the following,

COROLLARY 8. *Let $\tilde{x}$ be a function in $\mathcal{F}$. Then with respect to the metric topology in $\mathcal{F}$ (or in F),*

$$x=\lim_{n\to\infty} I(\tilde{g},\omega,t,n)/\lim_{n\to\infty} I(\tilde{f},\omega,t,n),$$

where $\tilde{x}=\tilde{g}/\tilde{f}$ with $\tilde{g}\in\mathcal{T}$, $\tilde{f}\in\mathcal{T}_0$, and the limits are evaluated for some ω satisfying $\tilde{\sigma}(\omega)\neq 0$. (Here / denotes the formal fraction in F.)

Proof. The formal fraction $(\sigma * g)(t)\tilde{\sigma}(\omega)/(\sigma * f)(t)\tilde{\sigma}(\omega)=(\sigma * g)(t)/(\sigma * f)(t)=g(t)/f(t)=x$ in F, provided $\tilde{\sigma}(\omega)\neq 0$. Thus the conclusion follows from the above discussion. Note that the result does not depend upon the particular choice of $\tilde{\sigma}$ nor ω, so long as the number $\tilde{\sigma}(\omega)\neq 0$.

In connection with differentiability in $\mathcal{F}$, we have the following consistency theorem.

THEOREM 6. *Suppose that the function $\tilde{x}\in\mathcal{F}$ is continuously differentiable everywhere on $\mathcal{R}$ and that $\tilde{x}^{(1)}=\tilde{y}$. Then*

$$\lim_{\xi\to 0}\frac{\tilde{x}_\xi-\tilde{x}}{\xi}=\tilde{y} \tag{16}$$

with respect to the metric topology in $\mathcal{F}$. Here $\tilde{x}_\xi(\omega)=\tilde{x}(\omega+\xi)$.

Proof. Because each of $\tilde{x}$ and $\tilde{y}$ is continuous, we have

$$\rho_k\left(\frac{\tilde{x}_\xi - \tilde{x}}{\xi}, \tilde{y}\right) \leqq \sup_{-k \leqq \omega \leqq k} \left|\frac{\tilde{x}(\omega+\xi)-\tilde{x}(\omega)}{\xi} - \tilde{y}(\omega)\right|$$

for all $\xi \neq 0$ and k. Since $\tilde{x}$ is continuously differentiable with derivative $\tilde{y}$, it follows that

$$\lim_{\xi \to 0} \rho_k\left(\frac{\tilde{x}_\xi - \tilde{x}}{\xi}, \tilde{y}\right) = 0$$

for every k. For operators, the corresponding result is as follows.

THEOREM 7. *If the operator $x \in F$ is such that its Fourier transform $\tilde{x}$ is continuously differentiable everywhere on $\mathcal{R}$, then the limit*

$$\lim_{\xi \to 0} \frac{e^{-i\xi t}x - x}{\xi} = Dx \tag{17}$$

exists with respect to the metric topology of F.

Proof. Equation (17) follows from (16) since $\widetilde{e^{-i\xi t}x(t)}(\omega) = \tilde{x}(\omega+\xi)$.

The limit in (17) defines the infinitesimal generator D of the one parameter group of exponential shifts (automorphisms of F) $e^{-i\xi t}$. By this theorem, the domain of the generator D includes the space (subring) of all operators x for which $\tilde{x}$ is continuously differentiable on $\mathcal{R}$. If f is a distribution such that $g(t) = -itf(t) \in \mathcal{B}_0'$, then by Proposition 2, $\tilde{g}$ is continuous on $\mathcal{R}$ and $\tilde{f}^{(1)} = \tilde{g}$ (see [6]). Hence, f is in the domain of D and $Df(t) = g(t)$, i.e., $Df(t) = -itf(t)$. The infinitesimal generator D of the group of exponential shifts, therefore, is simply an extension of the familiar algebraic derivative [2]. By Corollary 2, we see that if $f \in T$, then $f = (1+(-1)^N D^{2N})g$ for some $g \in \mathcal{B}_0'$ and some integer $N \geqq 0$.

These last two theorems are of considerable practical importance since every linear operator differential equation on R with polynomial coefficients in t transforms to an ordinary differential equation on $\mathcal{R}$ with polynomial coefficients in ω (and conversely). These theorems then show that every continuously differentiable (classical) solution of the latter on $\mathcal{R}$ transforms back to an operator solution of the former on R. If the solution on $\mathcal{R}$ is an ultradistribution (for example, if it is slowly increasing), then the solution on R is a distribution.

Examples. Let $\tilde{f}, \tilde{g} \in \mathcal{T}$, and suppose that the function $\tilde{f}$ satisfies the $\mathcal{R}$-differential equation $\sum a_{jk}\omega^j \tilde{f}^{(k)}(\omega) = \tilde{g}(\omega)$ for a.e. $\omega \in \mathcal{R}$, where the a_{jk} are complex numbers and the sum is over finitely many nonnegative integers j and k. Then the corresponding distribution $f \in T$ satisfies the R-differential equation $\sum (-i)^{j+k} a_{jk}(t^k f(t))^{(j)} = g(t)$. Now suppose that $\tilde{x}(\omega)$ is a continuous almost periodic function of $\omega \in \mathcal{R}$. Then it is the uniform limit on $\mathcal{R}$ of a sequence of trigonometric polynomials $\tilde{p}_n(\omega) = \sum a_{nj} e^{it_{nj}\omega}$. The corresponding operator $x(t)$ is, therefore, the limit in $\mathcal{F}$, with respect to the metric topology, of the sequence of distributions $p_n(t) = \sum a_{nj}\delta(t - t_{nj})$ in T, since $\widetilde{\delta(t-t_{nj})}(\omega) = e^{it_{nj}\omega}$.

5. Extension of the operator calculus. Let $\mathcal{K}$ denote the collection of all (Lebesque) measurable functions $\tilde{x}$ which are finite a.e. on $\mathcal{R}$. $\mathcal{K}$ becomes a ring under ordinary addition and multiplication of functions, and $\tilde{x}$ is a unit (is invertible) in $\mathcal{K}$ iff $\tilde{x}$ is nonzero a.e. on $\mathcal{R}$.

A simple preliminary lemma is all that is really needed to obtain our extension. It is the counterpart here of Theorem 2, and the proof is essentially the same.

LEMMA. *Let $\tilde{x} \in \mathcal{K}$. Then there exists a unit $\tilde{y} \in \mathcal{K}$ such that both $\tilde{y}$ and the product $\tilde{x}\tilde{y}$ are bounded.*

Proof. Since $\tilde{x}$ is measurable and finite a.e., the function $\tilde{y}(\omega) = 1/(1+|\tilde{x}(\omega)|)$ is measurable, bounded and nonzero a.e., i.e., $\tilde{y}$ is a bounded unit in $\mathcal{K}$. The product $\tilde{x}(\omega)\tilde{y}(\omega)$ is measurable and bounded.

The following theorem tells us that we may identify the elements of the ring $\mathcal{K}$ with fractions composed of ultradistributions, all of which are elements of $\mathcal{K}$.

THEOREM 8. *If $\tilde{x} \in \mathcal{K}$, then $\tilde{x}$ may be expressed in the form $\tilde{x} = \tilde{w}/\tilde{y}$ with $\tilde{w} \in \mathcal{K}$ and $\tilde{y} \in \mathcal{K}$ both regular ultradistributions and $\tilde{y}$ a unit. If $\tilde{v}$ is a regular ultradistribution, then $\tilde{v} \in \mathcal{K}$.*

Proof. The first conclusion follows directly from the lemma with $\tilde{w} = \tilde{x}\tilde{y}$. The second conclusion follows from the definition of a regular ultradistribution, which cannot be infinite on a set of positive measure.

We now let K denote the collection of all (formal) fractions g/f with $\tilde{g} \in \mathcal{K}$ and $\tilde{f} \in \mathcal{K}$ both regular ultradistributions and $\tilde{f}$ a unit in $\mathcal{K}$. Thus, g and f are always distributions. As before, such fractions are identified, added and multiplied, just as ordinary numerical fractions are, with the operations corresponding to addition for distributions and extended convolution for distributions defined by $f_1 * f_2 = f_3$ iff $\tilde{f}_3 = \tilde{f}_1\tilde{f}_2$.

The mapping which sends the fraction $x = g/f \in K$ to the function $\tilde{x} = \tilde{g}/\tilde{f} \in \mathcal{K}$ will be called the *Fourier transform*. Our next theorem is by now allegory and is a direct result of the above definitions.

THEOREM 9. *The rings F and $\mathcal{F}$ are subrings, respectively, of the rings K and $\mathcal{K}$. The Fourier transform defined above from K onto $\mathcal{K}$ is a ring isomorphism which extends the Fourier transform defined (in Definition 3) on F. In particular, the inverse Fourier transform from $\mathcal{K}$ onto K is a one-to-one, ring isomorphic, extension of the distributional inverse Fourier transform on the subspace of all regular ultradistributions.*

It is clear from the proof of the lemma that if $\tilde{x} \in \mathcal{K}$, then it can be further expressed as a fraction $\tilde{x} = \tilde{w}/\tilde{y}$ with $\tilde{w} \in \mathcal{K}$ and $\tilde{y} \in \mathcal{K}$ both regular ultradistributions which are, in fact, absolutely integrable. (Just divide the numerator and the denominator in Theorem 8 by $1+\omega^2$.) In such a case, the inverse Fourier transforms w and y are the ordinary continuous functions given by (2). Thus *distributions can be avoided altogether* and we now state the final, rather startling result of this paper.

THEOREM 10. *The ring K of operators is isomorphic with the (extended) convolution ring of fractions of continuous functions on R, each of which is the classical inverse Fourier transform of a function on $\mathcal{R}$, which is absolutely integrable.*

Examples. Suppose that the function $H(z)$ is analytic for $-b < \operatorname{Im} z < 0$, and that $\lim_{\rho \to 0^-} H(\omega + i\rho)$ exists (and is finite) for a.e. $\omega \in \mathcal{R}$. Then, of course, $\tilde{x}(\omega) = \lim_{\rho \to 0^-} H(\omega + i\rho)$ belongs to $\mathcal{K}$. Further, suppose that for each $\rho (-b < \rho < 0)$, the function $\omega \mapsto H(\omega + i\rho)$ is slowly increasing, and thus belongs to $\mathcal{T}$. Then, as before, there exists a distribution f such that for each ρ $(-b < \rho < 0)$, the distribution $e^{\rho t} f(t)$ belongs to T and satisfies $\widehat{e^{\rho t} f(t)}(\omega) = H(\omega + i\rho)$; f itself need

not belong to T. None-the-less, the distribution f may be identified with the operator x determined by $\tilde{x}(\omega)$. For the distributional Fourier transform $\tilde{f}$ of f is that ultradistribution defined by

$$\tilde{\phi} \mapsto \langle \tilde{f}(\omega), \tilde{\phi}(\omega) \rangle = \langle \tilde{f}(\omega + i\rho), \tilde{\phi}(\omega + i\rho) \rangle = \int_{-\infty}^{\infty} H(\omega + i\rho)\tilde{\phi}(\omega + i\rho)\, d\omega,$$

for any ρ $(-b < \rho < 0)$, and the function H satisfies $\lim_{\rho \to 0^-} H(\omega + i\rho) = \tilde{f}(\omega)$ in $Z'(C)$, since

$$\lim_{\rho \to 0^-} \int_{-\infty}^{\infty} H(\omega + i\rho)\tilde{\phi}(\omega)\, d\omega = \lim_{\rho \to 0^-} \int_{-\infty}^{\infty} H(\omega + i\rho + i\rho_0)\tilde{\phi}(\omega + i\rho_0)\, d\omega$$

$$= \int_{-\infty}^{\infty} H(\omega + i\rho_0)\tilde{\phi}(\omega + i\rho_0)\, d\omega$$

for any ρ_0 $(-b < \rho_0 < 0)$ and any $\tilde{\phi} \in Z$. In this situation, the limit function $\tilde{x}(\omega)$ is defined as the Fourier transform *in* $\mathcal{K}$ of f. It need not be a regular ultradistribution. For the case of a rational function $H(z) = Q(z)/P(z)$, which may have poles on $\mathcal{R}$, $x = f$ is always a distributional solution of the differential equation $P(-id/dt)f(t) = Q(-id/dt)\delta(t)$, but need not belong to T. This extends the result in an earlier example. Of course, if it happened that $\lim_{\rho \to 0^-} H(\omega + i\rho) = \tilde{x}(\omega)$ with respect to the metric topology of $\mathcal{F}$, then we would again identify that distribution f, for which $\widetilde{e^{\rho t} f(t)}(\omega) = H(\omega + i\rho)$, with the operator x determined by $\tilde{x}(\omega)$. However, metric convergence does not transpire in the above case if $H(z)$ actually has poles on $\mathcal{R}$. The Fourier transforms in $\mathcal{K}$ of the operators identified in this way with the Heaviside function $\mathbf{H}(t) = (1 + \operatorname{sgn} t)/2$, $t^n \mathbf{H}(t)$ and $t^n e^{\rho_0 t}\mathbf{H}(t)$ are, respectively, $-i/\omega$, $(-i)^{n+1} n!/\omega^{n+1}$ and $(-i)^{n+1} n!(\omega + i\rho_0)^{n+1}$ for $\rho_0 < 0$.

On the other hand, if $H(z)$ is analytic for $0 < \operatorname{Im} z < b$ and $\lim_{\rho \to 0^+} H(\omega + i\rho)$ exists a.e. on $\mathcal{R}$, then we let $\tilde{x}(\omega) = \lim_{\rho \to 0^+} H(\omega + i\rho)$, and identify the operator x with the distribution f (if any) for which $\widetilde{e^{\rho t} f(t)}(\omega) = H(\omega + i\rho)$ holds for $0 < \rho < b$. The Fourier transforms in $\mathcal{K}$ of $\mathbf{H}(-t)$, $t^n \mathbf{H}(-t)$, $t^n e^{\rho_0 t}\mathbf{H}(-t)$ become, respectively, i/ω, $-(-i)^{n+1} n!/\omega^{n+1}$ and $-(-i)^{n+1} n!/(\omega + i\rho_0)^{n+1}$ for $0 < \rho_0$, by virtue of this latter identification. Thus different distributions can have the same Fourier transforms in $\mathcal{K}$ and are equivalent as operators. In particular, $\mathbf{H}(t)$ and $-\mathbf{H}(-t)$ become equivalent as operators in K, though they are distinct as distributions, since their common Fourier transform in $\mathcal{K}$ is the function $-i/\omega$. (This should be expected since their Laplace transform is the analytic function $1/z$, restricted to the two half-planes $\operatorname{Re} z > 0$ and $\operatorname{Re} z < 0$.) On the other hand, the *distributional* Fourier transform of $\mathbf{H}(t)$ is p.v. $(-i/\omega) + \pi\delta(\omega)$, which is singular and cannot belong to $\mathcal{K}$. We note that the function $-i/\omega$ is also the arithmetical inverse in $\mathcal{K}$ of the function $i\omega$, and so $\mathbf{H}(t)$ (as well as $-\mathbf{H}(-t)$) is identified with the operator inverse in K of the differentiation operator s. Therefore, it is appropriate here to call $\mathbf{H}(t)$ the *integration* operator and to write $\mathbf{H}(t) = 1/s$. Now let $\tilde{x} \in \mathcal{K}$ be absolutely continuous (on every finite $\mathcal{R}$-interval). Then

$$\lim_{\xi \to 0} \frac{\tilde{x}(\xi + \omega) - \tilde{x}(\omega)}{\xi} = \tilde{x}^{(1)}(\omega)$$

for a.e. $\omega \in \mathcal{R}$. Therefore, $\lim_{\xi \to 0} (e^{-i\xi t}x - x)/\xi = Dx$ exists in K with respect to the

topology induced from a.e. convergence on $\mathscr{R}$. This topology, of course, is coarser than the metric topology of F used in Theorem 7 and extends the domain of the algebraic derivative D. More generally, if the function $\tilde{y} \in \mathscr{K}$ satisfies the *nonlinear* differential equation $\sum \tilde{a}_{jk}(\omega)(\tilde{y}^{(k)}(\omega))^j = \tilde{b}(\omega)$ for a.e. $\omega \in \mathscr{R}$, where $\tilde{a}_{jk}, \tilde{b} \in \mathscr{K}$ and the summation is over finitely many positive integers j and k, then the corresponding operator $y \in K$ satisfies the extended convolution equation $\sum a_{jk} \circledast (D^k y) \overbrace{\circledast \cdots \circledast}^{j \text{ factors}} (D^k y) = b$.

REFERENCES

[1] I. M. GELFAND AND G. E. SHILOV, *Generalized Functions. I*, Academic Press, New York, 1964.

[2] J. MIKUSIŃSKI, *Operational Calculus*, Pergamon Press, New York, 1959.

[3] D. B. PRICE, *On the Laplace transform for distributions*, SIAM J. Math. Anal., 6 (1975), pp. 49–80.

[4] L. SCHWARTZ, *Théorie des Distributions*, Hermann, Paris, 1966.

[5] ———, *Transformation de Laplace des distributions*, Math. Seminar, Université Lund, tome supplementaire dédié à M. Riesz, 1953, pp. 196–206.

[6] R. A. STRUBLE, *Representations of Fourier transforms for distributions*, Bull. Inst. Math. Academia Sinica, 2 (1974), pp. 191–206.

[7] ———, *A two-sided operational calculus*, Studia Math., to appear.

[8] ———, *Analytical and algebraic aspects of the operational calculus*, SIAM Rev., to appear.

[9] A. H. ŻEMANIAN, *Distribution Theory and Transform Analysis*, McGraw-Hill, New York, 1965.

[10] ———, *The distributional Laplace and Mellin transformations*, SIAM J. Appl. Math., 14 (1966), pp. 41–59.

SIAM J. MATH. ANAL.
Vol. 8, No. 2, April 1977

LINEAR TRANSFORMATIONS IN THE OPERATIONAL CALCULUS*

RAIMOND A. STRUBLE†

Abstract. A new development of linear transformations in the one-sided operational calculus is presented. The setting for this is a (noncommutative) ring $\mathscr{R}$ of continuous linear transformations on a familiar test function space from distribution theory. Included in $\mathscr{R}$, both algebraically and topologically, are all the right-sided distributions, all the traditional transformations of the operational calculus, such as the exponential shifts, the dilatations and the algebraic derivative, all translations and multiplications by infinitely differentiable functions, and many other new transformations. The development parallels that given by E. Gesztelyi for linear transformations of the Mikusiński operator field, but is cast in a simpler and more flexible setting. The main tool of the investigation is a representation theorem of the type introduced by V. Dolezal and the results are primarily theorems concerning commutativity properties in $\mathscr{R}$. It is shown that a linear transformation (i) commutes with all translations iff it is a distribution (ii) commutes with differentiation iff it is a distribution (iii) commutes with the algebraic derivative iff it is a multiplier and (iv) is a distribution and commutes with every dilatation iff it is a number. Because of the latter, it becomes reasonable to define a Laplace transform in $\mathscr{R}$ which encompasses (for right-sided distributions) that given by D. Price. Some results on inversion in $\mathscr{R}$ are given and a number of unsettled problems, perhaps amenable to solutions in this setting, are mentioned.

1. Introduction. In an earlier paper [4], E. Gesztelyi has studied continuous linear transformations on the field $\mathscr{M}$ of Mikusiński operators [6]. The familiar examples of such transformations are (see § 3) (a) the field elements themselves (including complex numbers) acting under multiplication, (b) the dilatations (c) the exponential shifts, (d) the algebraic derivative, and (e) the various combinations of those in (a), (b), (c) or (d). It remains an open question whether there are any others. Nonetheless, he has given a rather thorough treatment of the (noncommutative) ring $\mathscr{T}$ (under addition and composition) of all such transformations of $\mathscr{M}$ including the proofs of numerous commutativity theorems and a generalization of the Laplace transform for operators. Gesztelyi's definition of continuity, though nontopological, is a reasonable one in the Mikusiński operator case and allows for an interesting and useful representation theorem in $\mathscr{T}$ analogous to the spectral theorem of self-adjoint operators in Hilbert space. His representation theorem is the main tool used in the study of the transformations in $\mathscr{T}$.

Because of the uncertainty of the existence of interesting continuous linear transformations on $\mathscr{M}$ (other than those listed above) and of the technical complications accompanying analytical work with the Mikusiński field, it seems appropriate to look for a more flexible setting in which numerous linear transformations in the operational calculus can be studied with relative analytical ease. It is felt that such a setting is presented in this paper; a setting in which all (right-sided) distributions and all transformations listed above are included together with their usual algebraic and analytical properties. Moreover, many more familiar and not so familiar transformations are included and all are continuous in a strictly topological sense, which is equivalent to that associated with distributions. Furthermore, instead of following the traditional procedure of

* Received by the editors June 16, 1975, and in revised form October 29, 1975.

† Department of Mathematics, North Carolina State University, Raleigh, North Carolina 27607.

first introducing operators or distributions, say at one level, and then certain transformations of them (such as (b), (c), (d) and (e) above) at a second level, in the setting adopted in this paper, we work at one single level with a ring $\mathscr{R}$ (under addition and composition) of continuous linear transformations, called operator transformations (Definition 1). These transformations act on a familiar space of test functions $\mathscr{C}$, with a familiar topology, and include at one and the same time all the (right-sided) distributions (as convolutors) *and* the linear transformations traditionally associated with them and, moreover, other, nontraditional, continuous linear transformations.

The main tool used here in the study of the operator transformations of $\mathscr{R}$ is also a representation theorem (Theorem 1) which is different from Gesztelyi's, but is used in much the same fashion as Gesztelyi used his. Our representation theorem is merely another example of the type of representation theorems obtained recently by Dolezal [1], [2]. The proof for our case is not given since the arguments used are essentially the same as those used by Dolezal which need only be modified so as to take into account the fact that we deal here with *right-sided* test functions (infinitely differentiable functions with supports bounded on the left).

We first list some traditional linear transformations, (a), (b), (c), (d), (f), (g) in § 3 and then characterize distributions (Theorem 2 and Definition 2) with respect to our representation theorem in § 4. We use the latter to construct and study some examples of nonstandard operator transformations. This leads to a "generalized" convolution between certain pairs of elements of $\mathscr{R}$ which relates to the traditional approach (mentioned above) of applying linear transformations from a second level to distributions on a first level. For this purpose, we introduce a special notation (square brackets) designating the transformations of distributions, which is reserved for this special situation throughout the remainder of the paper. In § 5 we prove the main commutativity results (Theorems 3, 4, 8, 9 and Corollaries 1 and 3) for operator transformations which cover the same situations Gesztelyi covered. These include the statements that an operator transformation F commutes with the algebraic derivative D iff it is a multiplier or commutes with the differentiation operator s iff it is a distribution. Multiplicative operator transformations, such as the exponential shifts T^p and the dilatations U_k, are introduced in Definition 3 and are shown to form a multiplicative semigroup in $\mathscr{R}$. Moreover, it is found that among the multipliers, only the exponential shifts are multiplicative. Special commutativity results for distributions include the fact that only the numerical operators commute with D or with all dilatations U_k. Also some analytical concepts in $\mathscr{R}$ are introduced so that, for example, an analogue (Theorem 5) of Gesztelyi's representation (spectral) theorem can be established and that the differentiation of a one-parameter family of operator transformations (i.e., an operator transformation-valued function of a real variable) can be effected. Other analytical topics concern some sequential limits in $\mathscr{R}$ where it is shown (Corollary 4) that if $\lim_{n\to\infty} U_n F U_{1/n} = L$ in $\mathscr{R}$ exists for some distribution F, then L is a numerical operator transformation and (§ 6) that $\lim_{n\to\infty} U_n T^{-p} F T^p U_{1/n} = L(p)$ is an appropriate definition in $\mathscr{R}$ for the Laplace transform of F; a definition which for (right-sided) distributions encompasses the Schwartz–Laplace transform theory as has been developed recently in [7] by D.

Price. Finally in § 7 some results concerning inversion in $\mathscr{R}$ are obtained. These include the facts (Theorem 10) that a distribution which corresponds to a bijection on $\mathscr{C}$ is invertible and (Theorem 11) that the commutator equation $[D, F] = G$ is solvable in $\mathscr{R}$ for F whenever G is a distribution. Mikusiński operators are used to invert all right-sided distributions. Numerous unsettled problems are mentioned at various points throughout the paper.

2. Preliminary considerations. Let $\mathscr{C}$ denote the vector space of all infinitely differentiable complex-valued functions on the real line $-\infty<t<\infty=(-\infty, \infty)$ with supports bounded on the left, i.e., right-sided infinitely smooth functions. If $\phi, \psi \in \mathscr{C}$, then the convolution $\phi*\psi$ is the function

$$\sigma(t) = \int_{-\infty}^{\infty} \phi(u)\psi(t-u)\, du$$

and belongs to $\mathscr{C}$. Under addition and convolution, $\mathscr{C}$ becomes a commutative ring and, according to Titchmarsh's theorem on convolution [12], it is devoid of zero divisors.

We shall adopt the usual convergence notion in $\mathscr{C}$, which is that of compact convergence of all derivatives together with uniformly left bounded supports, so that $\mathscr{C}$ becomes the familiar space of test functions for distributions with right bounded supports [8], [14]. Thus for example, a sequence $\{\phi_n\}$ of elements of $\mathscr{C}$ *converges in* $\mathscr{C}$ (as $n \to \infty$) iff for each natural number j and each compact set $K \subseteq (-\infty, \infty)$, the function sequence $\{\phi_n^{(j)}\}$ of ordinary jth derivatives converges uniformly on K and there exists a real number t_0 such that the support of ϕ_n is contained in the right half-line (t_0, ∞) for every n. Clearly, such a sequence converges to an element of $\mathscr{C}$, and so $\mathscr{C}$ is complete with respect to this convergence notion. We remark that convergence in $\mathscr{C}$ is topological in the sense that there is a (locally convex) topology with respect to which convergence in $\mathscr{C}$ is topological convergence, but it is not necessary for our purposes to specify such a topology.

The above space of distributions with right bounded supports (left-sided distributions) will be denoted by $\mathscr{D}'_L$. If $f \in \mathscr{D}'_L$ and if $\phi \in \mathscr{C}$, then we denote by $\langle f, \phi\rangle = \langle f(t), \phi(t)\rangle$ the value of the distribution f for the test function ϕ. Similarly, the space of distributions with left bounded supports (right-sided distributions) will be denoted by $\mathscr{D}'_R$. If $f \in \mathscr{D}'_R$ and if $\phi \in \mathscr{C}$, then f and ϕ may be "convoluted", and we denote their convolution by $f * \phi = \psi$, where $(f * \phi)(\tau) = \langle f(\tau - t), \phi(t)\rangle$, and observe that $\psi \in \mathscr{C}$. Moreover, the mapping $\phi \mapsto \psi = f * \phi$ from $\mathscr{C}$ into $\mathscr{C}$ is continuous (in the sense of convergence in $\mathscr{C}$) and is linear. An extension of Titchmarsh's theorem says that $f * \phi = 0$ iff $f = 0$ or $\phi = 0$. Here and throughout this paper "0" always denotes the zeros (relative to addition) of the various vector spaces, rings and fields employed.

3. Operator transformations.

DEFINITION 1. A continuous and linear mapping $F: \mathscr{C} \to \mathscr{C}$ (from $\mathscr{C}$ into $\mathscr{C}$) will be called an *operator transformation* and the image ψ of ϕ under F will be denoted by $F(\phi) = \psi$ or, with variables indicated, by $F(\phi(t))(\tau) = \psi(\tau)$ with t and τ real numbers.

The collection of all such mappings will be denoted by $\mathscr{R}$, and under the usual addition and composition of mappings, $\mathscr{R}$ becomes a ring. (It will also be considered a vector space over the complex field.) Composition in $\mathscr{R}$ will be denoted by juxtaposition FG, or by $F \circ G$, and addition in $\mathscr{R}$ will be denoted, as usual, by a plus sign, $F+G$. In general, composition in $\mathscr{R}$ is *not* commutative.

We list below a number of familiar types of operator transformations together with their (more or less) traditional symbols and names.

		Operator transformation F	Distributional representation $f_\tau(t)$
(a)	*numerical*	$a\colon \phi(t)\mapsto a\phi(t)$,	$a\delta(\tau-t)$, (complex a),
	differentiation	$s\colon \phi(t)\mapsto \phi'(t)$,	$\delta'(\tau-t)$,
	integration	$h\colon \phi(t)\mapsto \int_{-\infty}^{t} \phi(u)\,du$,	$H(\tau-t)$,
	translations	$e^{\lambda s}\colon \phi(t)\mapsto \phi(t+\lambda)$,	$\delta(\tau+\lambda-t)$, (real λ),
(b)	*dilatations*	$U_k\colon \phi(t)\mapsto k\phi(kt)$,	$k\delta(k\tau-t)$, $(k>0)$,
(c)	*exponential shifts*	$T^p\colon \phi(t)\mapsto e^{pt}\phi(t)$,	$e^{p\tau}\delta(\tau-t)$, (complex p),
(d)	*algebraic derivative*	$D\colon \phi(t)\mapsto -t\phi(\ )$,	$-\tau\delta(\tau-t)$,

The linearity and continuity in all of these types is readily verified. The latter two cases illustrate a large class of operator transformations called multipliers.

(f)	*multipliers*	$\mu\colon \phi(t)\mapsto \mu(t)\phi(t)$,	$\mu(\tau)\,\delta(\tau-t)$.

Here μ is any infinitely differentiable function including, of course, any member of $\mathscr{C}$ or number (numerical operator transformation). Each member of the test function space $\mathscr{C}$ also induces an operator transformation through convolution. More generally, each right-sided distribution $f \in \mathscr{D}'_R$ induces an operator transformation through convolution.

(g)	*convolutors*	$f\colon \phi\mapsto f*\phi$,	$f(\tau-t)$.

A right-sided distribution is uniquely characterized (in $\mathscr{D}'_R$) by the induced (convolution) mapping. We observe that these latter mappings commute with convolution in $\mathscr{C}$, that is,

$$f*(\phi*\psi)=(f*\phi)*\psi \tag{1}$$

holds for all $f \in \mathscr{D}'_R$ and $\phi, \psi \in \mathscr{C}$. In a slightly different setting [10], [11], these have been referred to as operator homomorphisms because of this special property. Thus the ring $\mathscr{R}$ includes (isomorphic images of) all right-sided distributions, where composition becomes convolution in $\mathscr{D}'_R$. The numerical, differentiation, integration and translation operator transformations above correspond to distributions (under convolution) while the dilatations, exponential shifts, algebraic derivative and (nonnumerical, i.e., nonconstant) multipliers do not. However,

there is a sense in which all operator transformations can be expressed in terms of distributions, and we now consider this important concept.

4. Distributional representations of operator transformations. With almost identical arguments as those used in [1], [2] (especially those used in the proofs of Theorems 1.2 and 1.3 of [1]), one can prove the following.

THEOREM 1. *A mapping $F: \mathscr{C} \to \mathscr{C}$ is linear and continuous iff there exists a one-parameter family $\{f_\tau\}$ of left-sided distributions (in $\mathscr{D}'_L$) such that*

$$F(\phi(t))(\tau) = \langle f_\tau(t), \phi(t)\rangle, \tag{2}$$

holds for every $\phi \in \mathscr{C}$ and real τ. (Proof omitted).

With this theorem, then, we can conveniently express every operator transformation in terms of left-sided distributions. This is illustrated for each of the above types where $\delta(t)$ and $H(t)$ are, respectively, the Heaviside delta function and step function. Conversely, with every one-parameter family $\{f_\tau\}$ of left-sided distributions such that $\psi(\tau) = \langle f_\tau(t), \phi(t)\rangle$ defines a member $\psi \in \mathscr{C}$ for every $\phi \in \mathscr{C}$, we can obtain an operator transformation F using (2). In particular, if $f(t)$ is a right-sided distribution in $\mathscr{D}'_R$, then for each fixed τ, $f(\tau - t) = f_\tau(t)$ is a left-sided distribution and the operator transformation F associated with f through convolution satisfies $F(\phi(t))(\tau) = \langle f(\tau - t), \phi(t)\rangle = \langle f_\tau(t), \phi(t)\rangle$ for every $\phi \in \mathscr{C}$ and real τ. In general, operator transformations F can be considered to be represented by a modified type of convolution through (2). Actually, (2) constitutes ordinary convolution exactly when the mapping F commutes with convolution and $f_\tau(t) = f(\tau - t)$ for some right-sided distribution f. In fact, this special property is expressed more simply as follows.

THEOREM 2. *An operator transformation F commutes with all translation operators $e^{\lambda s}$ iff there exists a right-sided distribution f (in $\mathscr{D}'_R$) such that $f_\tau(t) = f(\tau - t)$ holds for all t and τ, where $f_\tau(t)$ is defined by* (2).

Proof. We need only prove the "only if" part since the translation operators all commute with convolution. Assume F commutes with all translation operators $e^{\lambda s}$. This means that

$$F(\phi(t - \lambda))(\tau) = F(\phi(t))(\tau - \lambda)$$

holds for all $\phi \in \mathscr{C}$ and real t, τ and λ. Using (2) this, in turn, means that

$$\langle f_\tau(t), \phi(t - \lambda)\rangle = \langle f_{\tau-\lambda}(t), \phi(t)\rangle$$

holds for all $\phi \in \mathscr{C}$ and t, τ and λ (for a suitable family $\{f_\tau(t)\}$ of left-sided distributions). In particular, for $\lambda = \tau$, we have

$$\langle f_\tau(t), \phi(t - \tau)\rangle = \langle f_0(t), \phi(t)\rangle = \langle f_0(t - \tau), \phi(t - \tau)\rangle,$$

for all $\phi \in \mathscr{C}$ and τ. Hence $f_\tau(t) = f_0(t - \tau) = f(\tau - t)$, where $f(t) = f_0(-t)$ is a right-sided distribution, and the proof is completed.

We may express this last theorem conveniently by saying $F \in \mathscr{R}$ satisfies $e^{\lambda s}F = Fe^{\lambda s}$ for all real λ iff F is defined by a single right-sided distribution through convolution. It would seem appropriate hereafter to say that such an F *is* a distribution, and so to be concise we shall. However, we shall often distinguish notationally between F considered as an operator transformation and the corresponding distribution f as a linear functional.

DEFINITION 2. An operator transformation F which satisfies $e^{\lambda s}F = Fe^{\lambda s}$ for all real λ will be called a *distribution*. It will be identified with the right-sided distribution f for which $F(\phi) = f * \phi$ holds for all $\phi \in \mathscr{C}$, according to Theorem 2.

Using (2) we now construct two examples of nontraditional operator transformations. Let $g \in \mathscr{D}'_L$ and have support in $(-\infty, -1)$. Then $\phi \in \mathscr{C}$ implies that $\psi \in \mathscr{C}$, where

$$\psi(\tau) = \langle g(t/(1+\tau^2)), \phi(t)\rangle = \langle g(t), (1+\tau^2)\phi((1+\tau^2)t)\rangle$$
$$= \langle g(t), \sigma(t)(1+\tau^2)\phi((1+\tau^2)t)\rangle,$$

and σ is infinitely differentiable with $\sigma(t) = 1$ for $t < -\frac{1}{2}$ and $\sigma(t) = 0$ for $t > -\frac{1}{4}$. Indeed, ψ is clearly infinitely differentiable and $(1+\tau^2)(-\frac{1}{4}) \to -\infty$ as $|\tau| \to \infty$ so that $\sigma(t)(1+\tau^2)\phi((1+\tau^2)t) = 0$ for all t provided $|\tau|$ is sufficiently large. Thus $\psi(\tau) = \langle g(t), 0\rangle = 0$ for all $|\tau|$ sufficiently large and ψ has, in fact, compact support. By Theorem 1, the one-parameter family $g_\tau(t) = g(t/(1+\tau^2))$ of left-sided distributions defines an operator transformation J through the representation (2), i.e., $J(\phi)(\tau) = \langle g(t/(1+\tau^2)), \phi(t)\rangle$ for all $\phi \in \mathscr{C}$ and real τ. For our second nontraditional example let $\sigma \in \mathscr{C}$, $g \in \mathscr{D}'_L$ and $f_\tau(t) = \sigma(\tau)g(t)$. Then $\phi \in \mathscr{C}$ implies that $\psi \in \mathscr{C}$, where

$$\psi(\tau) = \langle f_\tau(t), \phi(t)\rangle = \sigma(\tau)\langle g, \phi\rangle,$$

since $\langle g, \phi\rangle$ is simply a complex number and $\sigma \in \mathscr{C}$. In this case, the corresponding operator transformation is the mapping $G_1\colon \phi \mapsto \sigma\langle g, \phi\rangle = \psi$ for all $\phi \in \mathscr{C}$. Note that for a fixed ϕ, either $\psi = 0$ or ψ has the same support as does σ. If the support of ϕ is "sufficiently far to the right", then $G_1(\phi) = \psi = 0$ and the support of ψ can be considered to have "moved off" to $+\infty$.

As in this last example, operator transformations F invariably seem to have the property that the support of $\psi = F(\phi)$ tends to $+\infty$ as the support of ϕ tends to $+\infty$. It is conjectured that this is indeed always the case. In any event, if G is an operator transformation with this support property, then the mapping F defined by

$$(3)\quad \phi(t) \mapsto \langle f(-u), G(\phi(t+\tau))(u)\rangle = \langle f(-u), Ge^{\tau s}(\phi)(u)\rangle = \psi(\tau) = F(\phi(t))(\tau),$$

with $f \in \mathscr{D}'_R$ is an operator transformation. Indeed, ψ is infinitely differentiable and the support of $\phi(t+\tau)$ tends to $+\infty$ as τ tends to $-\infty$. Thus by the support property, the support of $\sigma(u) = G(\phi(t+\tau))(u)$ tends to $+\infty$ as τ tends to $-\infty$. But the support of $f(-u)$ is bounded on the right and so $\psi(\tau) = 0$ for all τ sufficiently negative and thus $\psi \in \mathscr{C}$. If G and $e^{\tau s}$ commute for all reall τ, i.e., G is a distribution, then

$$\langle f(-u), Ge^{\tau s}(\phi)(u)\rangle = \langle f(-u), e^{\tau s}G(\phi)(u)\rangle$$
$$= \langle e^{-\tau s}f(-u), G(\phi)(u)\rangle = \langle f(\tau - u), G(\phi)(u)\rangle$$
$$= (f * G(\phi))(\tau) = ((f * g) * (\phi))(\tau),$$

where g is the right-sided distribution corresponding to G. Thus (3) gives $F = f * g$, where the latter denotes ordinary convolution in $\mathscr{D}'_R$. When G and $e^{\tau s}$ do not commute for all real τ, then the operator transformation F defined by (3) may be considered as a type of "generalized" convolution of an element f of $\mathscr{D}'_R$

with an operator transformation G, i.e., $F = f * G$. Hence if the above conjecture is correct, we can extend convolution to all pairs of operator transformations provided one of them is a distribution.

For the particular nondistributional operator transformations in § 3, we obtain from (3) $f * U_k = -kU_{1/k}[f]$, $f * T^p = -T^{-p}[f]$, $f * D = D[f]$ and $f * \mu = -\hat{\mu} f$, where $kU_{1/k}[f(t)] = f(t/k)$, $T^{-p}[f(t)] = e^{-pt} f(t)$, $D[f(t)] = -tf(t)$ and $\hat{\mu}(t) = \mu(-t)$ as distributions. Here the bracket notation designates the transformations of the distributions indicated. For the above nonstandard operator transformations, we obtain from (3) $(f * J)(t) = \langle -f(-u), g(-t/(1+u^2))\rangle$ and $f * G_1 = \langle \hat{f}, \sigma \rangle \hat{g}$ as distributions. It is rather interesting to observe that all of these generalized convolutions result in distributions and that the first four amount to symbolic extensions of the corresponding mappings from functions to distributions in the traditional manner. Moreover, if f and g are both right-sided distributions and if F is an operator transformation with the above support property, then the composite mapping $g \circ (f * F)$, which is the product in $\mathcal{R}$ of g and $f * F$, turns out to be the same as the iterated generalized convolutions $g * (f * F)$. Indeed, $(g \circ (f * F))(\phi)(\tau) = \langle g(\tau - v), \langle f(-u), F(\phi(t+v))(u)\rangle\rangle = \langle g(-v), \langle f(-u), F(\phi(t+v+\tau))(u)\rangle\rangle = (g * (f * F))(\phi)(\tau)$. This further suggests that the generalized convolution $f * F$ always results in a distribution.

We note in passing that if $f \in \mathcal{D}'_R$ and if f corresponds to the operator transformation F, then the three distributions $D[f]$, $T^p[f]$ and $U_k[f]$ correspond to the operator transformations $DF - FD$, $T^p F T^{-p}$ and $U_k F U_{1/k}$, respectively. These expressions are readily verified by applying the bracket distributions as convolutors. Also it is easily verified that $U_k T^p = T^{kp} U_k$, $U_k D = kDU_k$, $U_k U_l = U_{kl}$, $T^p T^q = T^{p+q}$, $kU_k s = sU_k$ and $Ds = 1 + sD$.

5. Commutativity properties. According to Definition 2, all operator transformations which are distributions commute with all translation operators. Of course, all distributions commute with all other distributions and all multipliers commute with all other multipliers.

Perhaps the most significant commutativity properties of operator transformations are given in the following two theorems.

THEOREM 3. *An operator transformation F commutes with the algebraic derivative D, i.e., $DF = FD$ iff F is a multiplier μ.*

Proof. Clearly, if F is a multiplier μ, then F commutes with D. On the other hand, if F commutes with D, then $F(-t\phi(t))(\tau) = -\tau F(\phi(t))(\tau)$ holds for every ϕ, t and τ. Let $f_\tau(t)$ be defined by (2) for this F, and then

$$\langle f_\tau(t), -t\phi(t)\rangle = -\tau \langle f_\tau(t), \phi(t)\rangle,$$

or, what is the same,

$$\langle t f_\tau(t), \phi(t)\rangle = \langle \tau f_\tau(t), \phi(t)\rangle$$

holds for every ϕ, t and τ. Thus for a fixed τ, $(t-\tau) f_\tau(t) = 0$ in $\mathcal{D}'_L$, which implies [14] that $f_\tau(t) = \mu(\tau)\,\delta(\tau - t)$ for some number $\mu(\tau)$. Hence $F(\phi(t))(\tau) = \langle \mu(\tau)\,\delta(\tau - t), \phi(t)\rangle = \mu(\tau)\phi(\tau)$ holds for every ϕ and τ. Clearly, μ must be infinitely differentiable, since $\psi(\tau) = \mu(\tau)\phi(\tau)$ is so for every $\phi \in \mathcal{C}$, and thus F is the multiplier μ, which completes the proof.

THEOREM 4. *An operator transformation F commutes with the differentiation operator s, i.e., $sF = Fs$ iff F is a distribution.*

Before giving the proof of this theorem it is convenient to introduce some analytical concepts in $\mathcal{R}$. For example, if we have a one-parameter family $\{F_\lambda\} = F_\lambda$ of operator transformations, then any limits, continuity or differentiability or integrability properties of this family with respect to λ will be understood in the usual weak sense; i.e., for every $\phi \in \mathcal{C}$, $F_\lambda(\phi)$ is to possess the corresponding analytical property where, of course, convergence always means in $\mathcal{C}$. With this understanding it is easy to verify that most of the standard theorems of elementary calculus hold for $\mathcal{R}$-valued functions of a real variable λ.

Proof of Theorem 4. The one-parameter family of translations $e^{\lambda s}$ is differentiable with respect to λ and satisfies $de^{\lambda s}/d\lambda = se^{\lambda s}$. Now suppose an operator transformation F commutes with s and let $G_\lambda = e^{-\lambda s}Fe^{\lambda s}$. Then G_λ is differentiable with respect to λ and satisfies $dG_\lambda/d\lambda = -se^{-\lambda s}Fe^{\lambda s} + e^{-\lambda s}Fse^{\lambda s} = 0$, since s, $e^{-\lambda s}$ and s, F commute with each other. This implies that G_λ is, in fact, independent of λ; i.e., there exists a fixed operator transformation G such that $G = e^{-\lambda s}Fe^{\lambda s}$ holds for all λ. For $\lambda = 0$, this implies that necessarily $G = F$ and so F commutes with all translation operators and is, therefore by Definition 2, a distribution. This completes the proof of Theorem 4.

In [4] Gesztelyi obtained an integral representation of Mikusiński operator transformations reminiscent of the spectral representation for self-adjoint operators in Hilbert space. The following theorem gives the counterpart in $\mathcal{R}$ of Gesztelyi's representation theorem.

THEOREM 5. *For any $F \in \mathcal{R}$ and $\phi \in \mathcal{C}$ we have*

$$F \circ \phi = \int_{-\infty}^{\infty} \phi(\lambda) Fe^{-\lambda s}\, d\lambda, \tag{4}$$

where $F \circ \phi = F\phi$ is the product (composition) in $\mathcal{R}$ of F and ϕ, with the latter interpreted as a convolutor.

Proof. As a convolutor, ϕ satisfies $\phi(t) = \int_{-\infty}^{\infty} \phi(\lambda)\, e^{-\lambda s}\, d\lambda$, since for any $\psi \in \mathcal{C}$ we have

$$\begin{aligned}
(\phi * \psi)(t) &= \int_{-\infty}^{\infty} \phi(\lambda)\psi(t-\lambda)\, d\lambda \\
&= \int_{-\infty}^{\infty} \phi(\lambda)\, e^{-\lambda s}(\psi(t))\, d\lambda \\
&= \lim_{a\to\infty} \int_{-a}^{a} \phi(\lambda)\, e^{-\lambda s}(\psi(t))\, d\lambda \\
&= \lim_{a\to\infty} \left(\int_{-a}^{a} \phi(\lambda)\, e^{-\lambda s}\, d\lambda\right)(\psi(t)) \\
&= \left(\int_{-\infty}^{\infty} \phi(\lambda)\, e^{-\lambda s}\, d\lambda\right)(\psi(t))
\end{aligned}$$

for all t. This is essentially Mikusiński's formula [6]. Thus (by the continuity of F)

$$\begin{aligned} F \circ \phi &= F \circ \int_{-\infty}^{\infty} \phi(\lambda)\, e^{-\lambda s}\, d\lambda \\ &= F \circ \lim_{a \to \infty} \int_{-a}^{a} \phi(\lambda)\, e^{-\lambda s}\, d\lambda \\ &= \lim_{a \to \infty} \int_{-a}^{a} \phi(\lambda) F \circ e^{-\lambda s}\, d\lambda \\ &= \int_{-\infty}^{\infty} \phi(\lambda) F e^{-\lambda s}\, d\lambda, \end{aligned}$$

which proves the theorem.

Since the collection $\phi * \mathscr{C} = \{\phi * \psi : \psi \in \mathscr{C}\}$ is dense in $\mathscr{C}$ [3] if $\phi \neq 0$, the mapping $F \circ \phi$ in (4) completely characterizes F as an operator transformation (on $\mathscr{C}$) provided only that $\phi \neq 0$. Gesztelyi found the representation (4) very useful in studying Mikusiński operator transformations; however, in $\mathscr{R}$ it seems to be mainly a notational device and the representation (2) is the one of primary importance.

We now introduce a further algebraic concept in $\mathscr{R}$.

DEFINITION 3. An operator transformation F is said to be *multiplicative* if

$$F(\phi * \psi) = F(\phi) * F(\psi) \tag{5}$$

holds for every ϕ and ψ in $\mathscr{C}$.

If F and G are multiplicative, then their product FG satisfies

$$FG(\phi * \psi) = F(G(\phi) * G(\psi)) = F(G(\phi)) * F(G(\psi)) = FG(\phi) * FG(\psi)$$

for all ϕ and ψ of $\mathscr{C}$, and thus is also multiplicative. This establishes the following theorem.

THEOREM 6. *The subcollection $\mathscr{R}^*$ of all multiplicative operator transformations form a multiplicative semigroup in $\mathscr{R}$.*

The principal types of multiplicative operator transformations are the dilatations U_k and the exponential shifts T^p. It is unknown whether or not the group generated by these two types constitute all of $\mathscr{R}^* - \{0\}$ in this, or in the Mikusiński setting [13]. The next theorem tells us that the exponential shifts are the only (nonzero) multipliers in $\mathscr{R}$ which are multiplicative operator transformations.

THEOREM 7. *If μ is a (nonzero) multiplier in $\mathscr{R}$ and is a multiplicative operator transformation, then $\mu = T^p$ for some complex number p.*

Proof. If μ is multiplicative, then $\mu(\phi * \psi) = (\mu\phi) * (\mu\psi)$ holds for all ϕ and ψ of $\mathscr{C}$. Let $\{\psi_n\}$ be a delta function sequence [5] in $\mathscr{C}$; i.e., $\psi_n \in \mathscr{C}$ and $\psi_n \to \delta$ as $n \to \infty$ in $\mathscr{D}'_R$. Then for any $\phi \in \mathscr{C}$, we have

$$\mu(\phi * \psi_n) = (\mu\phi) * (\mu\psi_n) \quad \text{for all } n,$$

while $\mu(\phi * \psi_n) \to \mu\phi$ and $(\mu\phi) * (\mu\psi_n) \to \mu\phi\mu(0)$ as $n \to \infty$. Hence $\mu(0) = 1$ since μ is not identically zero. Now if in the above argument we translate each ψ_n by a

fixed amount τ, then we conclude that

$$\mu(t)\phi(t+\tau)=\mu(t+\tau)\phi(t+\tau)\mu(-\tau)$$

holds for all ϕ, t and τ. Thus μ satisfies the functional equation $\mu(t)=\mu(t+\tau)\mu(-\tau)$; i.e., $\mu(t+\tau)=\mu(t)\mu(\tau)$ for all t and τ. The only nonzero continuous (infinitely differentiable) solutions of this functional equation are the exponential functions. Hence there exists a complex number p such that $\mu(t)=e^{pt}$ holds for all t; i.e., $\mu = T^p$, and the proof is complete.

In view of Theorem 3, we have immediately the following corollary.

COROLLARY 1. *If F is a multiplicative operator transformation and if $FD = DF$, then there exists a complex number p such that $F = T^p$.*

By similar arguments, we can show that if F is multiplicative and satisfies $FD = kDF$ for some positive k, then there exists a complex number p such that $F = U_k T^p$.

Next we consider a special result for distributions.

THEOREM 8. *If $F = f$ is a distribution and if $FD = DF$, then F is a numerical operator transformation.*

Proof. For every $\phi \in \mathscr{C}$, we have

$$FD(\phi)=DF(\phi)=D(f*\phi)=(D[f])*\phi+f*D(\phi)$$

$$=(D[f])*\phi+FD(\phi) \quad \text{(since } D \text{ acts as a derivation)},$$

and so $(D[f])*\phi=0$ for every $\phi \in \mathscr{C}$. This implies that $(D[f])(t)=-tf(t)=0$ in $\mathscr{D}'_R$ so that $f(t)=a\delta(t)$ for some complex number a. Hence F is the numerical operator "a" and this proves the theorem.

Again in view of Theorem 3, we have immediately the following corollary.

COROLLARY 2. *Numerical operator transformations are the only multipliers which are distributions.*

This may be seen directly also by noting that if $\mu(\tau)\phi(\tau)=\langle f(t), \phi(\tau-t)\rangle$ holds for all ϕ, t and τ, then $\mu(\tau+\theta)\phi(\tau+\theta)=\langle f(t), \phi(\tau+\theta-t)\rangle=\mu(\tau)\phi(\tau+\theta)$ holds for all ϕ, t, τ and θ, so that $\mu(\tau+\theta)=\mu(\tau)$ holds for all τ and θ.

Another special result for distributions is the following.

THEOREM 9. *If $F = f$ is a distribution and if $U_k F = FU_k$ holds for all positive numbers k, then F is a numerical operator transformation.*

Proof. For every $\phi \in \mathscr{C}$ and k, we have $U_k F(\phi)=U_k(f*\phi)=U_k[f]*U_k(\phi)$ (since U_k is multiplicative), while $FU_k(\phi)=f*U_k(\phi)$. Hence $U_k[f]=f$ for every k. In [7] it is shown that if $U_k[f]=f$ for every k, then necessarily $f(t)=c_1\text{p.v.}(1/t)+c_2\,\delta(t)$ for suitable constants c_1 and c_2. Since p.v. $(1/t)$ is not a right-sided distribution, it must be that $c_2=0$ and so $f(t)=c_2\,\delta(t)$. Thus F is the numerical operator transformation "c_2", and this proves the theorem.

By using the multiplicative property of the dilatations, i.e., $U_{kl}=U_k U_l$, we can prove the following two corollaries.

COROLLARY 3. *If $F = f$ is a distribution and if $U_n F = FU_n$ holds for $n = 1, 2, \cdots$, then F is a numerical operator transformation.*

Proof. For any natural number m, we have $F=U_{1/m}U_m F=U_{1/m}FU_m$, so that $FU_{1/m}=U_{1/m}F$. Hence for any positive rational number $r=n/m$, we have $U_r F=U_{1/m}U_n F=U_{1/m}FU_n=FU_{1/m}U_n=FU_r$. Clearly then $U_k F=FU_k$ holds for all positive real numbers k and the result follows directly from Theorem 9.

COROLLARY 4. *If $F = f$ is a distribution and if $\lim_{n\to\infty} U_n F U_{1/n} = L$ exists in $\mathcal{R}$, (or, equivalently if $\lim_{n\to\infty} U_n[f] = l$ exists in $\mathcal{D}'_R$), then the limit L is a numerical operator transformation (or, equivalently l is a multiple of δ).*

Proof. It is easy to show that if the limit L exists, then $U_m L = L U_m$ holds for $m = 1, 2, \cdots$, and the result then follows from Corollary 3.

In connection with Theorem 9, it would be of interest to give examples of nondistributional operator transformations which commute with all dilatations. Such an F must satisfy $F(\phi(t))(k\tau) = F(\phi(kt))(\tau)$ for all $\phi \in \mathcal{C}$, real τ and $k > 0$. If $\tau > 0$ and $k = 1/\tau$, then $F(\phi(t))(1) = F(\phi(t/\tau))(\tau)$, where the mapping $\phi(t) \mapsto F(\phi(t))(1)$ defines a left-sided distribution, say f. Then $F(\phi(t/\tau))(\tau) = \langle f(t), \phi(t)\rangle$ holds for all $\phi \in \mathcal{C}$ and $\tau > 0$, or what is the same, $F(\phi(t))(\tau) = \langle f(t), \phi(\tau t)\rangle$ holds for all $\phi \in \mathcal{C}$ and $\tau > 0$. If this holds also for $\tau \leqq 0$ as well, then it is necessary for $f(t)$ to vanish for $t \leqq 0$, since $\psi(\tau) = \langle f(t), \phi(\tau t)\rangle$ must be right-sided for each right-sided ϕ. Thus f must have compact support in the open half-line $t > 0$. Conversely, if f is such a distribution, then the mapping $\phi \mapsto \langle f(t), \phi(\tau t)\rangle$ defines an operator transformation F which satisfies $U_k F = F U_k$ for all $k > 0$, as is readily verified. Note that such an F is represented (according to Theorem 1) by the family of distributions $U_{1/\tau}[f] = (1/\tau) f(t/\tau)$ for $\tau \neq 0$ and $\langle f, 1\rangle\, \delta(t)$ for $\tau = 0$. In particular, if $f(t) = a l^{j+1}\, \delta^{(j)}(t - l)$ for $l > 0$, then $F = a U_l D^j s^j$, and Theorem 9 corresponds to the case $l = 1$, $j = 0$. Using other distributions with compact support in the half-line $t > 0$, we obtain interesting examples of nonstandard operator transformations in this fashion.

6. Laplace transforms. Corollary 4 is the basis for a development of the Schwartz–Laplace transform theory for (not necessarily right-sided) distributions in [7]. In [4] an analogous development of a Laplace transform theory for Mikusiński operators has been given. Both of these developments hinge mainly on the facts that for any complex number p, if a limit like $\lim_{n\to\infty} U_n T^{-p} F T^p U_{1/n} = L(p)$ exists, then $L(p)$ is necessarily a number (numerical operator) and that if F is a function f with a Laplace transform, then $U_n T^{-p} F T^p U_{1/n} = U_n T^{-p}[f] = n e^{-pnt} f(nt)$ formally passes over to the Laplace transform of f as $n \to \infty$, i.e.,

$$\begin{aligned}
\lim_{n\to\infty} \langle n e^{-pnt} f(nt), \phi(\tau - t)\rangle &= \lim_{n\to\infty} \int_{-\infty}^{\infty} n e^{-pnt} f(nt) \phi(\tau - t)\, dt \\
&= \lim_{n\to\infty} \int_{-\infty}^{\infty} e^{-pu} f(u) \phi\left(\tau - \frac{u}{n}\right) du \\
&= \left(\int_{-\infty}^{\infty} e^{-pu} f(u)\, du\right) \phi(\tau) = L(p)\phi(\tau) \\
&= L(p)(\delta * \phi)(\tau)
\end{aligned}$$

for any test function ϕ and real τ. Thus the Laplace transform $L(p)$ of F is defined more generally here as $\lim_{n\to\infty} U_n T^{-p} F T^p U_{1/n} = L(p)$ if this limit exists in $\mathcal{R}$, (equivalently for distributions, as $L(p) = (1/\phi(0)) \lim_{n\to\infty} \langle U_n T^{-p}[f], \phi\rangle$ with $\phi(0) \neq 0$ if this limit exists in $\mathcal{D}'_R$). Under this definition, the relevant (rather complete) results in [7] carry over immediately to the distributions in $\mathcal{R}$ (which, of course, here, are right-sided). In particular, if f is a right-sided distribution whose Laplace transform $L(p)$ exists for $\mathrm{Re}\,(p) > c$, in the sense defined by Schwartz [8],

then $L(p)$ is an analytic function for $\operatorname{Re}(p) > c$ and is also given by the above limit for $\operatorname{Re}(p) > c$. This is easily verified directly for each of the distributions (a) listed in § 3. All of the Laplace transform theory or right-sided distributions can be developed directly from this limit definition.

On the other hand, with one exceptional case, the limit does not exist for any of the nondistributional operator transformations encountered in this paper. The exceptional case occurs for those F which commute with all U_k and for $p = 0$, where then $U_n T^{-p} F T^p U_{1/n} = F$ holds for all n, so that the limit is just F.

7. Inversion in $\mathscr{R}$ and in $\mathscr{M}$. It is easy to see that (a) the nonzero numerical operators, the differentiation, integration and translation operators, (b) the dilatations, (c) the exponential shifts and more generally, (f) the nonvanishing multipliers are all invertible in the ring $\mathscr{R}$. Of course, as in any ring, the collection $\mathscr{I}$ of invertible elements of $\mathscr{R}$ form a multiplicative subgroup of $\mathscr{R}$ so that the products of any of the above are also invertible in $\mathscr{R}$. We shall now show that $\mathscr{I}$ contains all of the distributions of $\mathscr{R}$ which are bijections on $\mathscr{C}$, such as those illustrated in (a) above.

THEOREM 10. *If F is a distribution and is a bijection on $\mathscr{C}$,* then F has an inverse F^{-1} *in $\mathscr{R}$ which is also a distribution.*

Proof. Clearly, since F is a bijection on $\mathscr{C}$, the algebraic inverse F^{-1} of F exists. Moreover, since F is a distribution, it commutes with convolution on $\mathscr{C}$ (see (1)) and since it is a bijection on $\mathscr{C}$, then its inverse F^{-1} also commutes with convolution on $\mathscr{C}$. Hence F^{-1} is an operator homomorphism [10] and thus corresponds to a distribution, which completes the proof.

The elementary type of example G_1 in § 4, where $G_1(\phi) = \sigma\langle g, \phi\rangle$, permits us to construct linear mappings on $\mathscr{C}$ which are not continuous, for we need only select a linear functional g on $\mathscr{C}$ which is not continuous. It would be of considerable interest to know if there are linear *bijections* on $\mathscr{C}$ which are not continuous since otherwise Theorem 10 could be extended from distributions to all of $\mathscr{R}$.

According to Theorem 3, if an operator transformation is *not* a multiplier, then the commutator $[D, F] = DF - FD \neq 0$. For distributions $F = f$, this becomes $D[f] \neq 0$ and suggests the following result.

THEOREM 11. *The commutator equation $[D, F] = DF - FD = 1$ has the solution $F = s$ in $\mathscr{R}$, which is unique to within an arbitrary (additive) multiplier. More generally, for any distribution G, the commutator equation $[D, F] = G$ has a (distributional) solution in $\mathscr{R}$, which is unique to within an arbitrary multiplier.*

Proof. It is readily verified that the differentiation operator s satisfies $Ds - sD = 1$ in $\mathscr{R}$. More generally, it is well known [9] that for any distribution g, the divisor problem $D[f](t) = -tf(t) = g(t)$ has a distributional solution $f(t) = -g(t)/t$ (actually, unique to within an arbitrary multiple of δ). If g has left bounded support, then so does f, and thus f corresponds to a solution F of $[D, F] = G$, where g corresponds to G. The difference of any two solutions of $[D, F] = G$ commutes with D and, hence by Theorem 3, is a multiplier. This completes the proof.

There is a sense in which all right-sided distributions are invertible, namely, as Mikusiński operators [5]. In the present context, a Mikusiński operator

(M-operator) x is merely a linear mapping in $\mathscr{C}$ which commutes with convolution, but may have for its domain only a proper ideal J of the ring $\mathscr{C}$. That is, $x: J \to \mathscr{C}$ satisfies $x(a\phi + b\psi) = ax(\phi) + bx(\psi)$ and $x(\phi * \sigma) = x(\phi) * \sigma$ for all $\phi, \psi \in J$, $\sigma \in \mathscr{C}$ and complex a, b. Because the ring $\mathscr{C}$ has no zero divisors, any such x can be uniquely extended to a maximal one and is uniquely determined by any restriction of itself to a nonzero ideal of $\mathscr{C}$. The right-sided distributions then are simply those M-operators which have extensions to all of $\mathscr{C}$ and turn out to be precisely the continuous M-operators. Two M-operators x and y are added $x + y$ and composed (multiplied) xy as mappings on suitably small nonzero ideals, and it is not difficult to show [5] that the collection $\mathscr{M}$ of all M-operators under this type of addition and composition becomes a field. This field is isomorphic to the one originally defined by Mikusiński [6]. The inverse $1/x$ in $\mathscr{M}$ of an M-operator x becomes simply the inverse mapping so that, in particular, every right-sided distribution f is invertible in $\mathscr{M}$, and its inverse there is the usual inverse mapping, f^{-1}.

Mikusiński's definition of convergence in $\mathscr{M}$ is of a very weak variety and, in this context, becomes pointwise convergence of functions *at a single* (nonzero) *point.* For example, an $\mathscr{M}$-valued function $\{x_\lambda\} = x_\lambda$ of a real variable λ is said to be continuous if for some nonzero ϕ in $\mathscr{C}$, the mapping $\lambda \mapsto x_\lambda(\phi)$ from say, some subset of real numbers into $\mathscr{C}$ is continuous in the standard topological sense. Of course, this requires that ϕ belong to the domain of the operator x_λ for each λ considered. Gesztelyi's requirement of continuity for a linear transformation F on $\mathscr{M}$ is essentially that every such $\mathscr{M}$-valued continuous function x_λ be mapped by F to another one, $y_\lambda = F(x_\lambda)$.

REFERENCES

[1] V. DOLEZAL, *A representation of linear continuous operators on testing functions and distributions*, this Journal, 1 (1970), pp. 491–506.

[2] V. DOLEZAL AND J. SANBORN, *Extendability of operators defined on testing function spaces*, Rev. Roumaine Math. Pures Appl., 20 (1975), pp. 33–53.

[3] C. FOIAS, *Approximation des opérateurs de J. Mikusiński par des fonctions continues*, Studia Math., 21 (1961), pp. 73–74.

[4] E. GESZTELYI, *Über lineare Operatortransformationen*, Publ. Math. Debrecen, 14 (1967), pp. 169–206.

[5] C. C. HUGHES AND R. A. STRUBLE, *Neocontinuous Mikusiński operators*, Trans. Amer. Math. Soc., 185 (1973), pp. 383–400.

[6] J. MIKUSIŃSKI, *Operational Calculus*, Pergamon Press, New York, 1959.

[7] D. B. PRICE, *On the Laplace transform for distributions*, this Journal, 6 (1975), pp. 49–80.

[8] L. SCHWARTZ, *Théorie des Distributions*, Hermann, Paris, 1966.

[9] ———, *Some applications of the theory of distributions*, Lectures on Modern Mathematics, T. Saaty, ed., John Wiley, New York, 1963.

[10] R. A. STRUBLE, *Operator homomorphisms*, Math. Z., 130 (1973), pp. 275–285.

[11] ———, *An algebraic view of distributions and operators*, Studia Math., 37 (1971), pp. 103–109.

[12] E. C. TITCHMARSH, *The zeros of certain integral functions*, Proc. London Math. Soc., 25 (1926), pp. 283–302.

[13] C. M. WALTERS, *Continuous linear transformations on the field of Mikusiński operators*, Ph.D. thesis, North Carolina State Univer., Raleigh, 1971.

[14] A. H. ZEMANIAN, *Distribution Theory and Transform Analysis*, McGraw-Hill, New York, 1965.

SIAM J. MATH. ANAL.
Vol. 8, No. 2, April 1977

ANALYTIC PROPERTIES OF GENERALIZED COMPLETELY CONVEX FUNCTIONS*

J. K. SHAW†

Abstract. The analytic character of real functions f in $C^\infty[a, b]$ which satisfy a certain positivity condition is studied. The condition is of the form $L^k f(x) \geqq 0$, $a \leqq x \leqq b$, $k = 0, 1, 2, \cdots$, where L is a Sturm-Liouville operator and L^k is its kth iterate. It is shown, for a special class of operators, that a function with this property is necessarily the restriction to $[a, b]$ of an analytic function in some complex neighborhood of $[a, b]$. The proof is based on a series representation associated with Sturm-Liouville boundary value problems.

1. Introduction. This paper is concerned with the analytic character of real infinitely differentiable functions which satisfy certain "positivity" properties. We shall investigate a type of condition, involving successive iterates of a linear differential operator, under which a real function f of class $C^\infty[a, b]$ is necessarily the restriction to $[a, b]$ of a complex function analytic in some complex neighborhood of the interval $[a, b]$.

The most familiar result in this direction is the well-known theorem of D. V. Widder [5] which asserts that a function $f \in C^\infty[0, 1]$ having the property

$$(-1)^k f^{(2k)}(x) \geqq 0, \qquad 0 \leqq x \leqq 1, \quad k = 0, 1, 2, \cdots, \tag{1.1}$$

is necessarily the restriction to $[0, 1]$ of an *entire* function, that is, a function analytic in the whole complex plane. Functions satisfying (1.1), the simplest being $\sin \pi x$, were termed *completely convex* by Widder. For a discussion of these and related classes of analytic functions, we refer the reader to the survey article on this subject by R. P. Boas [1].

The positivity condition we shall study was introduced by this author and J. D. Buckholtz [2] in a recent paper on a generalization of completely convex functions. Let L be the Sturm-Liouville operator given by

$$Ly = -(Py')' + Qy, \tag{*}$$

where P is a positive, continuously differentiable function on the interval $[a, b]$, and where Q is a real continuous function on $[a, b]$. Let

$$B_a y = \alpha y(a) + \alpha' y'(a), \qquad B_b y = \beta y(b) + \beta' y'(b)$$

be linearly independent boundary forms such that the eigenvalue problem

$$Ly = \lambda y, \qquad B_a y = B_b y = 0 \tag{1.2}$$

is self-adjoint. We say that a function $f \in C^\infty[a, b]$ is *LB-positive* if

$$(L^k f)(x) \geqq 0, \qquad a \leqq x \leqq b, \quad k = 0, 1, 2, \cdots, \tag{1.3a}$$

and

$$B_a L^k f \geqq 0, \quad B_b L^k f \geqq 0, \qquad k = 0, 1, 2, \cdots, \tag{1.3b}$$

* Received by the editors September 4, 1975, and in revised form January 29, 1976.

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

where L^k is the kth iterate of L. These are generalized completely convex functions in that (1.3) reduces to (1.1) in case $Ly = -y''$, $B_0y = y(0)$ and $B_1y = y(1)$. Similarly, for the system $Ly = -y''$, $B_0y = -y'(0)$ and $B_1y = y(1)$, the LB-positive functions coincide with a subclass of completely convex functions studied by S. Pethe and A. Sharma [4].

The object of the present paper is to prove, for suitably restricted operators L, that *LB-positive functions are analytic.* Our results applied to the system $Ly = -y''$, $B_0y = y(0)$, $B_1y = y(1)$ yield Widder's theorem as a special case. In general, though, the region of analyticity will vary with the operator L.

2. Hypotheses and statements of results. The methods we use presently are based on certain representation theorems in [2]. Thus, the hypotheses required in [2] must also be required here. We shall suppose that the eigenvalues of (1.2) are all *positive*, and that the signs on the constants in B_a and B_b are normalized so that

$$\alpha' \leqq 0, \quad \text{and if} \quad \alpha' = 0 \quad \text{then } \alpha > 0,$$

and

$$\beta' \geqq 0, \quad \text{and if} \quad \beta' = 0 \quad \text{then } \beta > 0.$$

This normalization either holds or can be brought about, without affecting the eigenvalues or eigenfunctions of (1.2), by multiplying one or both of the equations $B_ay = 0$ and $B_by = 0$ by -1. These hypotheses imply (see [2]) that solutions of the homogeneous equation $Ly = 0$ are *nonnegative* in $[a, b]$ if $B_ay \geqq 0$ and $B_by \geqq 0$, and either $B_ay > 0$ or $B_by > 0$. Thus we are assured that such functions are LB-positive.

Finally, it is clear that we must make some assumption concerning analyticity of the coefficient functions P and Q of the operator L, as otherwise the solutions of $Ly = 0$ would in general fail to be analytic. Our last hypothesis, then, is that $P(x)$ and $Q(x)$ are restrictions to $[a, b]$ of functions which are defined and analytic in some complex neighborhood of $[a, b]$. We denote the extensions of $P(x)$ and $Q(x)$ by $P(z)$ and $Q(z)$, respectively, and follow this same notational convention throughout when referring to analytic continuations of real functions.

Now let Ω be a simply connected region of the complex plane which contains the interval $[a, b]$, and in which $P(z)$ and $Q(z)$ are analytic and $P(z) \neq 0$. Then we have the following basic result [3].

THEOREM A. *Let $\psi(z)$, $P(z)$ and $Q(z)$ be analytic in a simply connected region Ω containing $[a, b]$, μ a complex number, and L defined by (*). Then every solution of*

$$Ly = \mu y + \psi$$

on $[a, b]$ is continuable analytically throughout Ω.

Our principal result is

THEOREM 1. *Let L and Ω be defined as in Theorem* A, *and let f be LB-positive. Then f is analytic in Ω; that is, f is the restriction of a complex function analytic in Ω.*

Once Ω has been established, then analytic continuations into Ω are unique by simple connectivity. Thus multiple-valued functions are avoided. However,

there may not be an optimal, or "largest possible," simply connected Ω. To illustrate the theorem, let us consider the functions x^ν on $[a, b]$, $0<a<b<\infty$, where ν ranges over the real numbers. Let $L_0y=-(x^2y')'$, $B_ay=y(a)$ and $B_by=y(b)$. One easily verifies that $L_0(x^\nu)=-(\nu^2+\nu)x^\nu$, and therefore $L_0^k(x^\nu)=(-1)^k(\nu^2+\nu)^kx^\nu$, $k=0, 1, 2, \cdots$. Hence x^ν is L_0B-positive if $\nu^2+\nu\leqq 0$. If $\nu^2+\nu>0$ we consider instead the operator $L_\nu y=-(x^2y')'+2(\nu^2+\nu)y$, and the same boundary forms. We have $L_\nu(x^\nu)=(\nu^2+\nu)x^\nu$, and therefore $L_\nu^k(x^\nu)=(\nu^2+\nu)^kx^\nu$, $k=0, 1, 2, \cdots$. Then x^ν is $L_\nu B$-positive if $\nu^2+\nu\geqq 0$.

Consequently, each of the functions x^ν, $-\infty<\nu<\infty$, is LB-positive for an appropriate choice of L. For either L_0 or L_ν, the region Ω may be taken as the complex plane with a ray from $z=0$ removed. The theorem concludes, as expected, that x^ν extends to the complex function z^ν having the ray as its (possible) branch line.

To prove Theorem 1, we shall actually establish a much stronger result having to do with series representation of LB-positive functions. The series, termed an *LB-series* in [2], has the form

$$f(x)=\sum_{k=0}^{\infty}\{(B_bL^kf)p_{2k}(x)+(B_aL^kf)p_{2k+1}(x)\}, \tag{2.1}$$

where the functions $\{p_k\}_0^\infty$ are defined recursively by

$$\begin{aligned} &Lp_0=Lp_1=0,\\ &B_ap_0=0, \qquad B_bp_0=1,\\ &B_ap_1=1, \qquad B_bp_1=0,\\ &Lp_m=p_{m-2}, \qquad B_ap_m=B_bp_m=0, \qquad m=2, 3, 4, \cdots. \end{aligned} \tag{2.2}$$

It can be shown [2] that each $p_k(x)$ is nonnegative in $[a, b]$, and is therefore LB-positive. Moreover, a simple induction argument based on Theorem A shows that $p_k(x)$ extends to an analytic function $p_k(z)$, $z\in\Omega$, $k=0, 1, 2, \cdots$.

The series representation we shall employ (Theorem 1 of [2]) may be stated as follows: if f is LB-positive, then f admits on $[a, b]$ the uniformly convergent representation

$$f(x)=Cy_0(x)+\sum_{k=0}^{\infty}\{(B_bL^kf)p_{2k}(x)+(B_aL^kf)p_{2k+1}(x)\}, \tag{2.3}$$

where C is a nonnegative constant dependent on f, and where y_0 is a nonnegative eigenfunction corresponding to the smallest eigenvalue λ_0 of (1.2).

In connection with LB-series, we prove the following.

THEOREM 2. *Let Ω be defined as in Theorem* A. *Suppose that for a sequence of nonnegative numbers $\{h_k\}_0^\infty$ the series*

$$S(x)=\sum_{k=0}^{\infty}\{h_{2k}p_{2k}(x)+h_{2k+1}p_{2k+1}(x)\} \tag{2.4}$$

converges in $[a, b]$. Then the complex series

$$S(z)=\sum_{k=0}^{\infty}\{h_{2k}p_{2k}(z)+h_{2k+1}p_{2k+1}(z)\}$$

converges absolutely for all $z \in \Omega$, and uniformly on compact subsets of Ω. Thus $S(z)$ is analytic in Ω.

Assuming for the moment that this result is true, consider its application to the representation (2.3). The LB-series in (2.3) is continuable to an analytic function in Ω by Theorem 2. The eigenfunction y_0, which satisfies $Ly_0 = \lambda_0 y_0$ on $[a, b]$, also extends to an analytic function $y_0(z)$, $z \in \Omega$, by Theorem A. Consequently, the function

$$f(z) = Cy_0(z) + \sum_{k=0}^{\infty} \{(B_b L^k f)p_{2k}(z) + (B_a L^k f)p_{2k+1}(z)\}, \qquad z \in \Omega,$$

is analytic in Ω, and this proves Theorem 1.

Thus there remains only to prove Theorem 2. For this we shall need a few basic properties of LB-series from [2]. These are listed in the following section.

3. Preliminaries. The following result (Theorem 3.2 of [2]) shows that coefficients in LB-series are uniquely determined and satisfy a summability condition.

THEOREM B. *Let $\{h_k\}_0^\infty$ be a real or complex sequence and suppose that the series*

$$S(x) = \sum_{k=0}^{\infty} h_k p_k(x) \tag{3.1}$$

converges in $[a, b]$. Then (3.1) converges uniformly in $[a, b]$, and

$$(L^n S)(x) = \sum_{k=2n}^{\infty} h_k p_{k-2n}(x)$$

for $a \leqq x \leqq b$ and $n = 0, 1, 2, \cdots$. Moreover

$$B_a L^n S = h_{2n+1}, \qquad B_b L^n S = h_{2n}, \qquad n = 0, 1, 2, \cdots,$$

and the series $\sum_{k=0}^{\infty} \alpha_k$, where

$$\alpha_{2k} = (p_0, y_0)\frac{h_{2k}}{\lambda_0^k}, \qquad \alpha_{2k+1} = (p_1, y_0)\frac{h_{2k+1}}{\lambda_0^k}, \qquad k = 0, 1, 2, \cdots,$$

is convergent.

Remark 1. *The symbol (u, v) is the usual inner product $(u, v) = \int_a^b u(t)\overline{v(t)}\, dt$.* Since p_0, p_1 and y_0 are nonnegative, we have $(p_0, y_0) > 0$ and $(p_1, y_0) > 0$.

Remark 2. The statement of Theorem B in [2] is for *real* sequences $\{h_k\}$, but the proof remains valid even if the h_k are complex.

Remark 3. If the h_k are nonnegative, then (3.1) converges absolutely. In particular, the terms in the series can be grouped as in (2.4) without affecting the sum.

We denote the eigenvalues of (1.2) by $\{\lambda_k\}_0^\infty$, with $0 < \lambda_0 < \lambda_1 < \cdots$. These are *simple* eigenvalues so that we may associate with each λ_k a single eigenfunction y_k, $Ly_k = \lambda_k y_k$, $B_a y_k = B_b y_k = 0$, normalized so that $\|y_k\|_2 = 1$.

The functions p_k have eigenfunction expansions

$$p_{2k+j}(x) = \sum_{n=0}^{\infty} \lambda_n^{-k}(p_j, y_n)y_n(x), \qquad j = 1, 2, \tag{3.2}$$

which converge uniformly in $[a, b]$ for $k = 1, 2, 3, \cdots$. Taking the terms corresponding to $n = 0$ to the left side of (3.2), multiplying by λ_0^k, and using some simple inequalities, one obtains

$$\lim_{k\to\infty} \lambda_0^k p_{2k+j}(x) = (p_j, y_0)y_0(x), \qquad j = 1, 2,$$

with uniform convergence in $[a, b]$. Thus we can find constants M_0 and M_1 such that

$$\begin{aligned} 0 &\leqq p_{2k}(x) \leqq M_0\lambda_0^{-k}, \\ 0 &\leqq p_{2k+1}(x) \leqq M_1\lambda_0^{-k}, \end{aligned} \tag{3.3}$$

for $a \leqq x \leqq b$ and $k = 0, 1, 2, \cdots$.

Finally, we require corresponding asymptotic estimates on the sequence of analytic functions $\{\varphi_k\}$ defined by

$$\varphi_{2k+j}(z) = p_{2k+j}(z) - \lambda_0^{-k}(p_j, y_0)y_0(z), \qquad z \in \Omega.$$

Starting from the eigenfunction expansion

$$\varphi_{2k+j}(x) = \sum_{n=1}^{\infty} \lambda_n^{-k}(p_j, y_n)y_n(x), \qquad a \leqq x \leqq b,$$

and proceeding as before, we find

$$\lim_{k\to\infty} \lambda_1^k \varphi_{2k+j}(x) = (p_j, y_1)y_1(x), \qquad j = 1, 2,$$

with uniform convergence in $[a, b]$. Thus we are led to the bounds

$$\begin{aligned} |\varphi_{2k}(x)| &\leqq M_0'\lambda_1^{-k}, \\ |\varphi_{2k+1}(x)| &\leqq M_1'\lambda_1^{-k}, \end{aligned} \tag{3.4}$$

for $a \leqq x \leqq b$, $k = 0, 1, 2, \cdots$, and for suitable constants M_0' and M_1'.

4. Proof of Theorem 2. Our proof will be based on extending the bounds (3.3) over into the complex domain Ω. We start by defining the generating functions

$$K_j(x, w) = \sum_{k=0}^{\infty} p_{2k+j}(x)w^{2k+j}, \qquad j = 1, 2. \tag{4.1}$$

Using (3.3), we have for complex w

$$\sum_{k=0}^{\infty} |p_{2k+j}(x)w^{2k+j}| = \sum_{k=0}^{\infty} \lambda_0^k p_{2k+j}(x)\left|\frac{w^2}{\lambda_0}\right|^k |w| \leqq M_j|w| \sum_{k=0}^{\infty} \left|\frac{w^2}{\lambda_0}\right|^k, \tag{4.2}$$

for $j = 1, 2$ and all x. Therefore, series (4.1) converges absolutely and uniformly on closed subsets of the complex disc $|w| < \sqrt{\lambda_0}$ for each fixed x, $a \leqq x \leqq b$. Then $K_j(x, w)$ is analytic in $|w| < \sqrt{\lambda_0}$, as a function of w, for each fixed x.

Now for fixed w, $|w|<\sqrt{\lambda_0}$, (4.1) is an LB-series in x on $[a, b]$. Applying Theorem B with $h_{2k+j} = w^{2k+j}$, one has

$$\begin{aligned} LK_j(x, w) &= \sum_{k=1}^{\infty} p_{2(k-1)+j} w^{2k+j} \\ &= w^2 \sum_{k=1}^{\infty} p_{2(k-1)+j}(x) w^{2(k-1)+j} \\ &= w^2 \sum_{k=0}^{\infty} p_{2k+j}(x) w^{2k+j} \\ &= w^2 K_j(x, w), \end{aligned}$$

that is,

$$LK_j(x, w) = w^2 K_j(x, w).$$

By Theorem A, then, $K_j(x, w)$ is continuable to an analytic function $K_j(z, w)$, $z \in \Omega$, for fixed $|w|<\sqrt{\lambda_0}$.

Hence $K_1(z, w)$ and $K_2(z, w)$ are analytic in the complex variables z and w separately, for $z \in \Omega$ and $|w|<\sqrt{\lambda_0}$.

Now consider the auxiliary generating functions

$$\hat{K}_j(x, w) = \sum_{k=0}^{\infty} \varphi_{2k+j}(x) w^{2k+j}, \qquad j = 1, 2. \tag{4.3}$$

Using (3.4) and proceeding as in (4.2), one sees that for fixed x, (4.3) converges absolutely and uniformly on closed subsets of the *larger* disc $|w|<\sqrt{\lambda_1}$. Then $\hat{K}_1(x, w)$ and $\hat{K}_2(x, w)$ are analytic there for fixed x.

Now suppose $|w|<\sqrt{\lambda_0}$ and write (4.1) in the form

$$\begin{aligned} K_j(x, w) &= \sum_{k=0}^{\infty} p_{2k+j}(x) w^{2k+j} \\ &= \sum_{k=0}^{\infty} \{\varphi_{2k+j}(x) + \lambda_0^{-k}(p_j, y_0) y_0(x)\} w^{2k+j} \\ &= (p_j, y_0) y_0(x) w^j \sum_{k=0}^{\infty} \left(\frac{w^2}{\lambda_0}\right)^k + \hat{K}_j(x, w) \\ &= (p_j, y_0) y_0(x) \frac{\lambda_0 w^j}{\lambda_0 - w^2} + \hat{K}_j(x, w) \\ &= Y_j(x, w) + \hat{K}_j(x, w), \end{aligned} \tag{4.4}$$

where

$$Y_j(x, w) = (p_j, y_0) y_0(x) \frac{\lambda_0 w^j}{\lambda_0 - w^2}, \qquad j = 1, 2.$$

Note that

$$LY_j(x, w) = (p_j, y_0)y_0(x)\frac{\lambda_0^2 w^j}{\lambda_0 - w^2}$$
$$= \lambda_0 w^j (p_j, y_0)y_0(x) + w^2(p_j, y_0)y_0(x)\frac{\lambda_0 w^j}{\lambda_0 - w^2}$$
$$= \lambda_0 w^j (p_j, y_0)y_0(x) + w^2 Y_j(x, w).$$

Then applying the operator L to $\hat{K}_j(x, w)$ in (4.4) yields

$$L\hat{K}_j(x, w) = LK_j(x, w) - LY_j(x, w)$$
$$= w^2 K_j(x, w) - w^2 Y_j(x, w) - \lambda_0 w^j (p_j, y_0)y_0(x),$$

and so

$$L\hat{K}_j(x, w) = w^2 \hat{K}_j(x, w) - \lambda_0 w^j (p_j, y_0)y_0(x). \tag{4.5}$$

Now this equation, as does (4.4), holds for $a \leqq x \leqq b$ and $|w| < \sqrt{\lambda_0}$. But $\hat{K}_j(x, w)$ is analytic in the larger disc $|w| < \sqrt{\lambda_1}$. Therefore, each side of (4.5) is analytic in $|w| < \sqrt{\lambda_1}$. Since the separate sides agree on $|w| < \sqrt{\lambda_0}$, they must agree on $|w| < \sqrt{\lambda_1}$ by the identity theorem. This shows that (4.5) is valid for $a \leqq x \leqq b$ and $|w| < \sqrt{\lambda_1}$. By Theorem A, $\hat{K}_j(x, w)$ extends to an analytic function $\hat{K}_j(z, w)$, $z \in \Omega$, and it therefore follows that $\hat{K}_1(z, w)$ and $\hat{K}_2(z, w)$ are analytic in the variables z and w separately for $z \in \Omega$ and $|w| < \sqrt{\lambda_1}$.

LEMMA 1. *If $z \in \Omega$, then*

$$\lim_{k \to \infty} \lambda_0^k p_{2k+j}(z) = (p_j, y_0)y_0(z), \qquad j = 1, 2.$$

Furthermore, the convergence is uniform on compact subsets of Ω.

Proof. By uniqueness of the Taylor series coefficients in (4.3), we have for each k,

$$\varphi_{2k+j}(x) = \frac{1}{2\pi i}\int_{|w|=R} \frac{\hat{K}_j(x, w)}{w^{2k+j+1}}\, dw \tag{4.6}$$

for $a \leqq x \leqq b$, $j = 1, 2$, and where R may be chosen to satisfy $\lambda_0 < R^2 < \lambda_1$. The separate sides of (4.6) extend to analytic functions in Ω which agree on $[a, b]$. Thus

$$\varphi_{2k+j}(z) = \frac{1}{2\pi i}\int_{|w|=R} \frac{\hat{K}_j(z, w)}{w^{2k+j+1}}\, dw \tag{4.7}$$

for $z \in \Omega$ and $j = 1, 2$. For each compact subset $A \subset \Omega$, let $m(A)$ denote the maximum of $|\hat{K}_j(z, w)|$ for $z \in A$ and $|w| = R$. Then by (4.7),

$$|\varphi_{2k+j}(z)| \leqq \frac{m(A)}{R^{2k+j}}, \qquad z \in A,$$

and so

$$|\lambda_0^k \varphi_{2k+j}(z)| \leqq \frac{m(A)}{R^j}\left(\frac{\lambda_0}{R^2}\right)^k, \qquad z \in A.$$

Recalling the definition of $\varphi_{2k+j}(z)$, we now have

$$|\lambda_0^k p_{2k+j}(z) - (p_j, y_0)y_0(z)| \leqq \frac{m(A)}{R^j}\left(\frac{\lambda_0}{R^2}\right)^k, \qquad z \in A.$$

Noting that the right side of this inequality is independent of z, and that $\lambda_0 < R^2$, we obtain the desired result by letting $k \to \infty$.

The following is a trivial consequence of Lemma 1.

LEMMA 2. *For each compact subset $A \subset \Omega$, there exists a constant $M(A)$ such that*

$$|\lambda_0^k p_{2k+j}(z)| \leqq M(A)$$

for $z \in A$, $k = 0, 1, 2, \cdots$, and $j = 1, 2$.

We now proceed with the proof of Theorem 2. We are given a sequence $\{h_k\}_0^\infty$ of nonnegative numbers such that the series

$$S(x) = \sum_{k=0}^{\infty} \{h_{2k}p_{2k}(x) + h_{2k+1}p_{2k+1}(x)\}$$

converges everywhere in $[a, b]$. By Theorem B and the remarks following it, the series

$$\sum_{k=0}^{\infty} \frac{h_{2k} + h_{2k+1}}{\lambda_0^k}$$

is convergent. Now let A be a compact subset of Ω and consider the series

$$S(z) = \sum_{k=0}^{\infty} \{h_{2k}p_{2k}(z) + h_{2k+1}p_{2k+1}(z)\} \tag{4.8}$$

for $z \in A$. By Lemma 2,

$$\begin{aligned}\sum_{k=0}^{\infty} \{|h_{2k}p_{2k}(z)| + |h_{2k+1}p_{2k+1}(z)|\} &= \sum_{k=0}^{\infty} \left\{\frac{h_{2k}}{\lambda_0^k}|\lambda_0^k p_{2k}(z)| + \frac{h_{2k+1}}{\lambda_0^k}|\lambda_0^k p_{2k+1}(z)|\right\} \\ &\leqq M(A) \sum_{k=0}^{\infty} \frac{h_{2k} + h_{2k+1}}{\lambda_0^k}.\end{aligned}$$

Therefore, series (4.8) converges absolutely and uniformly on A, and this completes the proof.

REFERENCES

[1] R. P. BOAS, *Signs of derivatives and analytic behavior*, Amer. Math. Monthly, 78 (1971), pp. 1085–1093.

[2] J. D. BUCKHOLTZ AND J. K. SHAW, *Generalized completely convex functions and Sturm–Liouville operators*, this Journal, 6 (1975), pp. 812–828.

[3] E. A. CODDINGTON AND N. LEVINSON, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.

[4] S. PETHE AND A. SHARMA, *Modified Abel expansion and a subclass of completely convex functions*, this Journal, 3 (1972), pp. 546–558.

[5] D. V. WIDDER, *Functions whose even derivatives have a prescribed sign*, Proc. Nat. Acad. Sci. U.S.A., 26 (1940), pp. 657–659.

SIAM J. MATH. ANAL.
Vol. 8, No. 2, April 1977

ON AN ISOPERIMETRIC INEQUALITY FOR THE FIRST EIGENVALUE OF A BOUNDARY VALUE PROBLEM*

MIRIAM BAREKET†

Abstract. Let D be a two-dimensional simply connected bounded domain whose boundary ∂D consists of a finite number of regular arcs. This paper suggests that for all such domains D of the same area A, the circle yields the maximum value for the first eigenvalue λ_1 of the problem:

$$\Delta u + \lambda u = 0 \quad \text{in } D,$$

$$\frac{\partial u}{\partial n} = Zu \qquad \text{on } \partial D.$$

Here $\partial/\partial n$ denotes differentiation with respect to the exterior normal to D and Z is a *positive* constant.

This isoperimetric property of λ_1 is proved for any $Z>0$ under certain assumptions on the circumference, and the local extremum property is shown for certain values of the parameter.

1. Introduction. It was first conjectured by Lord Rayleigh [13] that of all fixed membranes of a given area, the circle yields the lowest fundamental tone; that is: "For all domains D of the same area A the circle yields the minimum value for the first eigenvalue λ_1 of

$$\Delta u + \lambda u = 0 \quad \text{in } D,$$

$$u = 0 \qquad \text{on } \partial D."$$

This conjecture was formally proved by Faber [7] and Krahn [10], and later a different proof was given by Pólya and Szegö [12]. In [14] Szegö proved an isoperimetric property for the first nonzero eigenvalue μ_2 of the free membrane problem:

$$\Delta u + \mu u = 0 \quad \text{in } D$$

$$\frac{\partial u}{\partial n} = 0 \qquad \text{on } \partial D.$$

Here it is assumed that the boundary ∂D is an analytic curve and it is proved that: "For all domains D of a given area $A(D)$ the circle yields the maximum value of μ_2." In the case of

$$\begin{aligned} &\Delta u + \lambda u = 0 \quad \text{in } D \\ &\frac{\partial u}{\partial n} = Zu \qquad \text{on } \partial D, \quad Z>0, \end{aligned} \tag{1}$$

the first eigenvalue λ_1 is always negative, since the Rayleigh quotient related to (1) is

$$Q(u, D) = \frac{\iint_D (\nabla u)^2 \, d\sigma - Z \int_{\partial D} u^2 \, ds}{\iint_D u^2 \, d\sigma} \tag{2}$$

* Received by the editors February 13, 1975, and in revised form October 1, 1975.

† Tel-Aviv University, Israel, and Department of Mathematics, University of Maryland, College Park, Maryland 20742.

and $\lambda_1 = \min_u Q(u, D)$, where the minimum is sought over all functions u that are continuous and have piecewise continuous first derivatives in D. Setting $u = \text{const.}$ in (2), we get

$$\lambda_1 \leqq -Z\frac{\int_{\partial D} ds}{\iint_D d\sigma} = -Z\frac{L(D)}{A(D)} < 0, \tag{3}$$

where $L(D)$ and $A(D)$ denote the circumference and the area of D respectively. Inequality (3) also furnishes an upper bound for λ_1, and by the fundamental isoperimetric property of the circle, for all domains D of the same area A, the circle yields the maximum upper bound of the form $-ZL(D)/(A(D))$. This, and other supporting arguments lead to the conjecture that: "For all domains D of a given area $A(D)$ the circle yields the maximum value of λ_1 of (1)." This conjecture is the subject of this paper.

It is worth mentioning that although problems similar to (1) are usually connected with membranes, (1) is not. It appears in acoustics in connection with propagation of waves through elastic cylinders [11].

2. The main theorem. In the proofs given in this section, we shall use the following bound for λ_1 for a circular domain R of radius a:

$$Z^2 + \frac{Z}{a} \leqq x^2 \leqq Z^2 + \frac{2Z}{a}, \tag{4}$$

where $x^2 = -\lambda_1(R)$. Inequality (4) was derived in [4] using Barta's method [5]. Since the function used for obtaining (4) is not a solution of (1), the equality signs in (4) do not hold for any finite Z except for the case $Z = 0$, for which $\lambda_1(R) = 0$.

LEMMA. *Let $\lambda_1(D)$ be the first eigenvalue of* (1) *for a domain D of area A and circumference $L(D)$, and let $\lambda_1(R)$ be the first eigenvalue of* (1) *for a circle R of radius a with the same area A. If $L(D)/(2\pi a) \geqq 1 + Za/2$, then $\lambda_1(D) \leqq \lambda_1(R)$.*

Proof. From (4), we have $-Z^2 - 2Z/a \leqq \lambda_1(R)$. Since (3) holds for any domain D, then as long as $-ZL(D)/A \leqq -Z^2 - 2Z/a$ holds, (which is equivalent to the condition of the lemma), we have $\lambda_1(D) \leqq \lambda_1(R)$.

Remark. The larger $L(D)$ gets and the smaller Z is, the larger is the family of domains for which the lemma holds.

In the proof we used the inequality (3). In case of several star-shaped domains, better upper bounds for $\lambda_1(D)$ can be found [3], and the condition of the lemma can be weakened. When D is a circle or a rectangle, we can find solutions to (1) by separation of variables. The first eigenfunction for the circle R of radius a is given by

$$u_1(R) = A_0 I_0(xr) \qquad (x^2 = -\lambda_1(R)), \tag{5}$$

where A_0 is a constant and x is determined by solving the transcendental equation

$$\frac{I_1(xa)}{I_0(xa)} = \frac{Z}{x}. \tag{6}$$

I_0 and I_1 are the hyperbolic Bessel functions of order 0 and 1, respectively. The

TABLE 1

domain	circle $a=1$	square $a=\frac{\sqrt{\pi}}{2}$ $\cong 0.88625$	rectangle $a=1$ $b=\frac{\pi}{4}$ $\cong 0.7854$	rectangle $a=2$ $b=\frac{\pi}{8}$ $\cong 0.3927$	rectangle $a=10$ $b=\frac{\pi}{40}$ $\cong 0.07854$	rectangle $a=100$ $b=\frac{\pi}{400}$ $\cong 0.00785$	rectangle $a=1000$ $b=\frac{\pi}{4000}$
z/circumference	6.2832	7.0892	7.1416	9.5708	40.314	400.03	4000.01
1	−2.5921	−3.1090	−3.1266	−3.9857	−14.167	−128.69	−1274.70
2	−6.6822	−8.8141	−8.8560	−10.7525	−30.854	−260.93	−2559.60
5	−30.5698	−50.0280	−50.0432	−51.7746	−97.990	−670.22	−6398.17
10	−110.5232	−200.0000	−200.0000	−200.1530	−268.66	−1407.55	−12,864.69

first eigenfunction for a rectangle D' of sides $2a$ and $2b$ is given by

$$u_1(D') = A_1 \cosh \alpha x \cosh \beta y,$$

where A_1 is a constant, α is determined by $\tanh \alpha a = Z/\alpha$ and β by $\tanh \beta b = Z/\beta$ and $\lambda_1(D') = -(\alpha^2+\beta^2)$.

In order to get an idea of how the first eigenvalue of (1) changes with the circumference of D, we can look at Table 1 in which λ_1 was calculated for the unit circle and for a square and various rectangles all of whose areas are π. Since $\lambda_1 \leqq -ZL(D')/(A(D'))$, it follows that $\lambda_1 \to -\infty$ as $L(D') \to \infty$ and $A(D')$ is fixed.

Now we check how the first eigenvalue of (1) varies when an infinitesimal area preserving perturbation of D is performed. Let λ_N and u_N be the Nth eigenvalue and eigenfunction of (1) for the domain D. Let D be perturbed as in Courant–Hilbert [6, pp. 419–421] and let δn be the infinitesimal change in ∂D, in the direction of the outer normal to D. (δn may be positive or negative). It is easy to show that

$$\delta\lambda_N = \int_{\partial D} \left[\left(\frac{\partial u_N}{\partial s}\right)^2 - (Z^2+\kappa Z+\lambda_N)u_N^2\right]\delta n \, ds,$$

where $\delta\lambda_N$ denotes the change in λ_N up to terms of first order and κ denotes the curvature of ∂D. In the case of the circle, u_1 depends on r alone (5), and therefore $u_1(\partial R) = \text{const.}$ and $\partial u_1/\partial s(\partial R) = 0$. In this case, $\kappa = 1/a$ is constant and

$$\delta\lambda_1 = -\left(Z^2+\frac{Z}{a}+\lambda_1\right)u_1^2(\partial R)\int_0^{2\pi} a\delta n \, d\theta. \tag{7}$$

Here $\delta n = \delta n(\theta)$, and from the requirement for area preservation, we have

$$\begin{aligned} A(R) = \pi a^2 = A(R^*) &= \int_0^{2\pi}\int_0^{a+\delta n} r \, dr \, d\theta \\ &= \pi a^2 + a\int_0^{2\pi} \delta n \, d\theta + \frac{1}{2}\int_0^{2\pi} (\delta n)^2 \, d\theta, \end{aligned} \tag{8}$$

where $A(R^*)$ is the area of the perturbed domain. Hence, up to terms of higher than the first order, the area preservation requirement yields

$$\int_0^{2\pi} \delta n \, d\theta = 0,$$

and thus in case of the circle, $\delta\lambda_1 = 0$. Hence the circle is a "stationary point" which is a necessary but not a sufficient condition that the extremal domain should satisfy.

We shall next calculate the change of $\lambda_1(R)$ up to terms of second order, caused by an area preserving perturbation.

DEFINITION. We call a domain R' a *nearly circular domain* if it is obtained from the circle R (of radius a) by an infinitesimal perturbation as in [6], for which $\delta n = \delta n(\theta)$ is given by the periodic function $\delta(\theta)$. R' is given by $r \leqq a + \delta(\theta)$.

We call $\delta(\theta)$ an *area preserving perturbation* if $A(R') = A(R)$.

Using [6] and [9], it is easy to show that $\delta(\theta)$ and its derivative $\delta'(\theta)$ are of the same infinitesimal order.

Remark. A perturbation of this type clearly changes the circle into a star-shaped domain [8] with respect to the origin.

THEOREM. *Let R' be a nearly circular domain, obtained by an area preserving perturbation $\delta(\theta)$ of the unit circle. Then there exists a positive constant Z_0 such that $\lambda_1(R') \leqq \lambda_1(R)$ for all $Z < Z_0$.*

Proof. Let D be a domain star-shaped with respect to the origin, whose boundary curve is given by $r = \eta(\theta)$. Let $h = h(\theta)$ be the distance between the origin and the tangent to ∂D at the point $(r = \eta(\theta), \theta)$ [8, pp. 410–411]. By Appendix A,

$$\lambda_1(D) \leqq \frac{\lambda_1(R)}{2A(D)} \left\{ \int_{\partial D} \frac{ds}{h} + K\left[L(D) - \int_{\partial D} \frac{ds}{h} \right] \right\}, \tag{9}$$

where $\lambda_1(R) = -x^2$ is the first eigenvalue of (1) for the unit circle and

$$K = \frac{2Z}{x^2 - Z^2}. \tag{10}$$

By [8], we have

$$\frac{ds}{h(\theta)} = \left[1 + \frac{\eta'(\theta)}{\eta(\theta)} \right]^2 d\theta \quad \text{and} \quad ds = \sqrt{\eta^2 + \eta'^2}\, ds, \tag{11}$$

where $'$ indicates differentiation with respect to θ.

LEMMA 1. $\int_0^{2\pi} (\delta')^2\, d\theta > \int_0^{2\pi} \delta^2\, d\theta$.

Proof. Calculating $L(D)$ using (11) for $\eta(\theta) = 1 + \delta(\theta)$ yields

$$L(D) = \int_0^{2\pi} (1+\delta) \sqrt{1 + \left(\frac{\delta'}{1+\delta} \right)^2}\, d\theta, \tag{12}$$

which, up to second order terms, yields

$$L(D) = 2\pi + \int_0^{2\pi} \delta\, d\theta + \int_0^{2\pi} \frac{(\delta')^2}{2}\, d\theta.$$

Since $\delta = \delta(\theta)$ is an area preserving perturbation, $L(D) > L(R) = 2\pi$, and thus

$$\int_0^{2\pi} \frac{(\delta')^2}{2}\, d\theta > -\int_0^{2\pi} \delta\, d\theta.$$

Looking back to (8), we have by the area preserving requirement that $-\int_0^{2\pi} \delta\, d\theta = \int_0^{2\pi} \delta^2/2\, d\theta$ and therefore $\int_0^{2\pi} (\delta')^2\, d\theta > \int_0^{2\pi} \delta^2\, d\theta$.

LEMMA 2. *The function $K(Z)$ is a monotone increasing function of Z, $1 < K(Z) < 2$ for $0 < Z < \infty$ and $\lim_{Z \to 0} K(Z) = 1$.*

In order to prove the lemma, we first note:

(a) Since $x^2|_{Z=0} = 0$ and $I_1(0) = 0$ while $I_0(0) = 1$ [1, pp. 374–375, pp. 423–428], we have by (6) that $\lim_{Z \to 0} (Z/x)^2 = 0$.

(b) $x^2 = -\lambda_1(R)$ is an analytic function of Z [2]. Differentiation yields

$$\frac{dx^2}{dZ} = \frac{2x^2}{x^2 - Z^2} = \frac{2}{1 - \left(\frac{Z}{x} \right)^2}, \tag{13}$$

and therefore

$$\left.\frac{dx^2}{dZ}\right|_{Z=0}=2.$$

(c) $(Z/x)^2$ is a monotone increasing function in Z. This can be seen by checking the first derivative with respect to Z.

Proof of Lemma 2. First, by using (4), we get

$$1\leqq K(Z)=\frac{2Z}{x^2-Z^2}\leqq 2. \tag{14}$$

By l'Hôpital's rule and by note (b) we have

$$\lim_{Z\to 0}K(Z)=\lim_{Z\to 0}\frac{2}{dx^2/dZ-2Z}=1.$$

We now wish to show that $K(Z)$ is a monotone increasing function of Z. In view of the preceding remarks, $K(Z)$ is an analytic function of Z, and by differentiating it, we get

$$\dot{K}(Z)=\frac{dK}{dZ}=\frac{2(x^2-Z^2)-2Z(dx^2/dZ-2Z)}{(x^2-Z^2)^2}.$$

By using (13) and the definition of K, we can write this last expression in the form

$$\dot{K}(Z)=\frac{2x^2}{(x^2-Z^2)^2}\left[1+\left(\frac{Z}{x}\right)^2-K(Z)\right]. \tag{15}$$

It is sufficient to show that $\dot{K}(Z)>0$. By expanding $K(Z)$ in power series of Z in the neighborhood of the origin, we find that $K=1+Z/4+o(Z^2)$, and therefore K is increasing there. For any $Z>0$ it is sufficient to show that $1+(Z/x)^2>K(Z)$. Suppose the inequality does not hold for all $0<Z<\infty$; then there exists some $Z=Z'$ for which $1+(Z/x)^2=K(Z)$ and $1+(Z/x)^2<K(Z)$ for some $Z>Z'$. $K(Z)$ is a decreasing function there since $\dot{K}(Z)<0$ in this interval. By note (c), $1+(Z/x)^2$ is an increasing function of Z, and if for $Z>Z'$, $K(Z)$ is decreasing, then $1+(Z/x)^2>K(Z)$ there, and that is a contradiction. Hence either $1+(Z/x)^2>K(Z)$ for all $0<Z<\infty$ or $1+(Z/x)^2=K(Z)$ for some Z's. In both cases, $K(Z)$ is increasing in Z, and that completes the proof of Lemma 2.

Now we return to the proof of the Theorem. By using (11) for $\eta(\theta)=1+\delta(\theta)$ and inserting the result into (9) for $A(D)=A(R)=\pi$, we get

$$\lambda_1(R')\leqq\lambda_1(R)\left[1+\frac{1-K}{2\pi}\int_0^{2\pi}\left(\frac{\delta'}{1+\delta}\right)^2 d\theta+K\left(\frac{L(D)}{2\pi}-1\right)\right]. \tag{16}$$

Since $\lambda_1(R)$ is negative, as long as its multiplier is greater than 1, $\lambda_1(R')\leqq\lambda_1(R)$. Therefore we wish to exhibit conditions under which

$$\frac{1-K}{2\pi}\int_0^{2\pi}\left(\frac{\delta'}{1+\delta}\right)^2 d\theta+K\left(\frac{L(D)}{2\pi}-1\right)>0. \tag{17}$$

Calculating (17) up to terms of second order, we find that (17) will hold as long as

$$K \leqq \frac{2\int_0^{2\pi} (\delta')^2\, d\theta}{\int_0^{2\pi} (\delta')^2\, d\theta + \int_0^{2\pi} \delta^2\, d\theta}. \tag{18}$$

Denote the right-hand side of (18) by K'. $K' < 2$ since $\int_0^{2\pi} \delta^2\, d\theta > 0$. By Lemma 1, $K' > 1$ and therefore $1 < K' < 2$. By Lemma 2, there exists a Z_0 such that $K' = K(Z_0)$, and for every $Z < Z_0$, $K(Z) \leqq K(Z_0) = K'$. So there exists a $Z_0 > 0$ such that for any $Z < Z_0$, $\lambda_1(R') \leqq \lambda_1(R)$. Q.E.D.

The question of the existence of a similar isoperimetric inequality for the elastic supported membrane (i.e., equation (1) for negative values of Z), namely, that for all domains of a given area A, the circle yields the minimum value for λ_1, is still unanswered. Yet, a calculation shows that $\delta\lambda_1 = 0$ for a circular domain, and therefore the circle is a "stationary point" in this case as well.

Appendix A. Derivation of formula (9). By (2), we have

$$\lambda_1(D) \leqq \frac{\iint_D (\nabla u)^2\, d\sigma - \int_{\partial D} Z u^2\, ds}{\int_D u^2\, d\sigma}. \tag{A.1}$$

For a star-shaped domain D, whose boundary curve is given by $r = \eta(\theta)$, we introduce the coordinate transformation $r = \rho\eta$ [3], [8] and calculate (A.1) for functions u of the form $u(r, \theta) = v(r/(\eta(\theta))) = v(\rho)$, obtaining

$$\lambda_1(D) \leqq \frac{\int_{\partial D} ds/h \int_0^1 [v'(\rho)]^2 \rho\, d\rho - Z[v(1)]^2 L(D)}{2A(D)\int_0^1 [v(\rho)]^2 \rho\, d\rho}, \tag{A.2}$$

where h is as in the proof of the main theorem.

Different upper bounds for $\lambda_1(D)$ can be established by inserting different functions $v(\rho)$ into (A.2). In particular, if we choose $v(\rho)$ to be $I_0(x\rho)$, the first eigenfunction of (1) for the unit circle R, we get

$$\lambda_1(D) \leqq \frac{\lambda_1(R)}{2A(D)}\left[\int_{\partial D} \frac{ds}{h} + \frac{2}{x}\frac{I_0(x)I_1(x)}{I_0^2(x) - I_1^2(x)}\left[L(D) - \int_{\partial D} \frac{ds}{h}\right]\right], \tag{A.3}$$

where $\lambda_1(R) = -x^2$. Since $I_0(x\rho)$ satisfies the boundary condition (6), $I_1(x)/(I_0(x)) = z/x$, formula (9) is concluded.

REFERENCES

[1] M. ABRAMORITZ AND I. A. STEGNON, *Handbook of Mathematical Functions*, Dover, New York, 1965.

[2] C. BANDLE AND R. P. SPERB, *Application of Rellich's perturbation theory to a classical boundary and eigenvalue problem*, Z. Angew. Math. Phys., 24 (1973), pp. 710–720.

[3] M. BAREKET, *An eigenvalue problem for an elliptic operator with special boundry conditions*, Ph.D. dissertation, Tel-Aviv Univ., 1973. (In Hebrew, with English summary.)

[4] M. BAREKET AND B. RULF, *An eigenvalue problem related to sound propagation in elastic tubes*, J. of Sound Vibration, 38 (1975), pp. 437–449.

[5] J. BARTA, *Sur la vibration fondamentale d'une membrane*, C.R. Acad. Sci. Paris, 204 (1937), pp. 482–473.

[6] R. COURANT AND D. HILBERT, *Methods of Mathematical Physics*, vol. I, Interscience, New York, 1966.

[7] G. FABER, *Beweis dass unter allen homogenen membranen bon gleicher fläche und gleicher spannung die kreisformige den tiefsten grundton fibt*, Sitzungsberichte der Bayrischen Akademie der Wissenschaften (1923), pp. 169–172.

[8] P. GARABEDIAN, *Partial Differential Equations*, John Wiley, New York, 1964.

[9] P. R. GARABEDIAN AND M. SCHIFFER, *Convexity of domain functionals*, J. Anal. Math., 2 (1952–3), Chap. II.

[10] E. KRAHN, *Über eine non Rayleigh formulierte Minimal-eigenschaft des Kreises*, Math. Ann., 94 (1924), pp. 97–100.

[11] P. M. MORSE, *Vibration and Sound*, 2nd. ed., McGraw-Hill, New York, 1948.

[12] G. PÓLYA AND G. SZEGÖ, *Isoperimetric Inequalities in Mathematical Physics*, Princeton University Press, Princeton, N.J., 1951.

[13] LORD RAYLEIGH, *The Theory of Sound*, 2nd ed., London, 1894. Reprinted: Dover, New York, 1945.

[14] G. SZEGÖ, *Inequalities for certain eigenvalues of a membrane of given area*, J. Rational Mech. Anal., 3 (1954), pp. 343–356.

SIAM J. MATH. ANAL.
Vol. 8, No. 2, April 1977

ON CHAPLYGIN'S PROBLEM*

M. S. KLAMKIN†

Abstract. Chaplygin's problem is to determine the closed path of an airplane flying in a horizontal plane with a constant speed with respect to a constant windfield if it is to fly around the greatest area in a given time. A complete solution is obtained by using Wulff's construction.

In a recent note "*On extreme length flight paths*" [1], the author gave some elementary solutions of several extremal problems. A related but more difficult one is Chaplygin's problem [2, pp. 206–208]. Here an airplane is flying horizontally at a constant speed v with respect to a constant wind field given by $\mathbf{W}$ ($w=|\mathbf{W}|<v$) and we want to determine the closed path one should fly, with respect to ground in a given time, such that the area enclosed by the path is a maximum. Using the calculus of variations [2, pp. 206–208], it has been shown formally that the path is an ellipse whose major axis is perpendicular to $\mathbf{W}$ and whose eccentricity is w/v. By using Wulff's construction [3] for the equilibrium shape of crystals, we can give a simpler and complete proof. To apply Wulff's construction, we consider the dual problem, i.e., minimizing the time to traverse the boundary of a region of given area or minimize $T=\oint(ds/v_g)$ subject to $\oint[x(dy/ds)-y(dx/ds)]\,ds=\text{const.}$

To determine the speed v_g of the airplane with respect to ground when its path with respect to ground makes an angle θ with $\mathbf{W}$, we resolve $\mathbf{W}$ and $\mathbf{V}$ (the plane's velocity with respect to $\mathbf{W}$) along and perpendicular to the path. See Fig. 1.

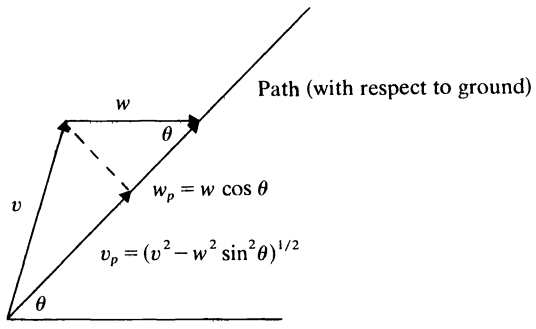


FIG. 1

The component of $\mathbf{V}$ along the path is $(v^2-w^2\sin^2\theta)^{1/2}$ and then

$$v_g = w\cos\theta+(v^2-w^2\sin^2\theta)^{1/2}.$$

For Wulff's construction, we first plot r versus θ, where

$$r=(v^2-w^2)/v_g=\sqrt{v^2-w^2\sin^2\theta}-w\cos\theta. \tag{1}$$

* Received by the editors September 29, 1975, and in revised form December 30, 1975.

† Department of Mathematics, University of Alberta, Edmonton, Alberta, Canada T6G 2G1.

Then aside from a simple scale transformation, the desired extremal path is found by finding the envelope of the family of lines normal to the radius vector (from the origin of the (r, θ) coordinate system) at each point of (1). A justification of this construction using the Brunn–Minkowski theorem is given by J. A. Taylor [4] and C. A. Johnson and G. D. Chakerian [5]. The family of normal lines is given by

$$x \cos\theta + y \sin\theta = \sqrt{v^2 - w^2 \sin^2\theta} - w \cos\theta.$$

To find the envelope, we first differentiate partially with respect to θ and then solve parametrically for x and y:

$$-x \sin\theta + y \cos\theta = \frac{-w^2 \sin\theta \cos\theta}{\sqrt{v^2 - w^2 \sin^2\theta}} + w \sin\theta;$$

$$x + w = \frac{v^2 \cos\theta}{\sqrt{v^2 - w^2 \sin^2\theta}}, \qquad y = \frac{(v^2 - w^2)\sin\theta}{\sqrt{v^2 - w^2 \sin\theta}}.$$

It now follows that

$$\frac{(x+w)^2}{v^2} + \frac{y^2}{v^2 - w^2} = 1,$$

giving the desired result.

It is to be noted that this result contains the isoperimetric theorem for circles (just set $\mathbf{W} = 0$).

Coincidentally, we also get the congruent path

$$\frac{(x-w)^2}{v^2} + \frac{y^2}{v^2 - w^2} = 1$$

if we minimize $\oint v_g \, ds$ instead of $\oint (v^2 - w^2)\, ds/v_g$.

For other wind field problems, see [6], [7], [8].

Acknowledgment. For the proof here, the author is indebted to F. J. Almgren, Jr. who in his excellent series of three Hedrick Lectures [9] illustrated Wulff's construction geometrically with respect to the same problem for a sailboat in which the vector velocity of the boat is given graphically.

REFERENCES

[1] M. S. KLAMKIN, *On extreme length flight paths*, Classroom Note, SIAM Rev., 18 (1976), pp. 486–488.

[2] N. I. AKHIEZER, *The Calculus of Variations*, Blaisdell, New York, 1962.

[3] G. WULFF, *Zur Frage der Geschwindigkeit des Wachsthums und der Auflösung der Krystallflachen*, Z. Krist., 34 (1901), pp. 449–530.

[4] J. E. TAYLOR, *Existence and structure of solutions to a class of nonelliptic variational problems*, 1st. Naz. Alta Matematica, Symp. Math., 14 (1974), pp. 499–508.

[5] C. A. JOHNSON AND G. D. CHAKERIAN, *On the proof and uniqueness of Wulff's construction of the shape of minimum surface free energy*, J. Mathematical Phys., 6 (1965), pp. 1403–1404.

[6] *Problem* 61-4, *Flight in an irrotational wind field*, SIAM Rev., 4 (1962), pp. 155–156.

[7,8] M. S. KLAMKIN AND D. J. NEWMAN, *Flying in a Wind Field. I, II*, Amer. Math. Monthly, 76 (1969), pp. 16–23, 1013–1019.

[9] F. G. ALMGREN, JR., *Geometric measure theory and the calculus of variations*, M. A. A. Summer Meeting, Western Michigan Univ., Kalamazoo, 1975.

SIAM J. MATH. ANAL.
Vol. 8, No. 2, April 1977

DOUBLY ORTHOGONAL CONCENTRATED POLYNOMIALS*

E. N. GILBERT AND D. SLEPIAN†

Abstract. We seek an nth degree polynomial $f_0^{(n)}(x)$ which maximizes the ratio

$$R(f)=\int_{I_a}|f(x)|^2\,dx\Big/\int_{I_b}|f(x)|^2\,dx,$$

where I_a and I_b are two intervals on the real line. $R(f)$ may be interpreted as an energy ratio and $f_0^{(n)}(x)$ as the polynomial having its energy most concentrated into I_a at the expense of its energy in I_b. Maximizing $R(f)$ is equivalent to finding the largest eigenvalue $\lambda_0^{(n)}$ and corresponding eigenfunction $f_0^{(n)}(x)$ of an eigenvalue problem. The other eigenfunctions, which are also polynomials of degree n, have interest because the eigenfunctions $f_j^{(n)}(x)$, $j=0,\cdots,n$, are orthogonal both on I_a and on I_b simultaneously.

For small n the eigenvalue problem can be solved numerically by standard matrix methods. We give special attention to asymptotic results for n large. When I_a and I_b are disjoint, $\lambda_0^{(n)}$ grows as $C_1 n^{-1} C_2^n$. We give C_1 and C_2 as functions of I_a and I_b. We also solve the problem when I_a is centrally positioned inside I_b, say, $I_a=[-a,a]$, $I_b=[-1,1]$, with $a<1$. Then, for large n, $\lambda_0^{(n)}$ has the behavior $1-C_3 n^{1/2} C_4^n$ and we obtain C_3 and C_4. In both these cases the eigenvalue problem can be put into differential equation form.

When I_a and I_b are disjoint we maximize other ratios, related to $R(f)$, to obtain maximizing polynomials which are simple expressions involving Chebyshev or Legendre polynomials. These polynomials have $R(f)$ growing with the same exponential term C_2^n as $\lambda_0^{(n)}$ but with constant factors different from C_1.

1. Introduction.

Consider two intervals I_a and I_b on the real line. The ratio

$$R(f)=\int_{I_a}|f(x)|^2\,dx\Big/\int_{I_b}|f(x)|^2\,dx \tag{1}$$

is a simple index of how much larger a complex-valued function $f(x)$ is when $x\in I_a$ than when $x\in I_b$. $R(f)$ will be called the *concentration* of f because it may be interpreted physically as measuring the extent to which the energy of $f(x)$ is concentrated in the (time) interval I_a and away from I_b. When $f(x)$ is restricted to the set F_n of polynomials of degree not greater than n, there is a largest possible ratio

$$\lambda_0^{(n)}=\max_{f\in F_n}R(f). \tag{2}$$

Much of this paper is concerned with determining how rapidly $\lambda_0^{(n)}$ grows with n.

In connection with this problem we also study $n+1$ *concentrated polynomials* $f_0^{(n)}(x), f_1^{(n)}(x),\cdots,f_n^{(n)}(x)$ each contained in F_n. Here $f_0^{(n)}(x)$ is a *most concentrated* polynomial, i.e., one in F_n for which $R(f_0^{(n)})=\lambda_0^{(n)}$. For $j=1,2,\cdots,n$, the polynomial $f_j^{(n)}(x)$ has greatest possible concentration among all polynomials of degree n orthogonal in the usual Hermitian sense on I_b to $f_0^{(n)}, f_1^{(n)},\cdots,f_{j-1}^{(n)}$. We denote this extremal concentration by

$$\lambda_j^{(n)}\equiv R(f_j^{(n)}),\qquad j=0,1,\cdots,n. \tag{3}$$

* Received by the editors June 16, 1975, and in revised form January 15, 1976.
† Bell Laboratories, Murray Hill, New Jersey 07974.

In the remainder of this Introduction we present without proof some of the more interesting properties of the polynomials $f_j^{(n)}(x)$ and of their concentrations $\lambda_j^{(n)}$.

The concentrations are nonincreasing,

$$\lambda_0^{(n)} \geqq \lambda_1^{(n)} \geqq \cdots \geqq \lambda_n^{(n)} > 0 \tag{4}$$

and interleaving,

$$\lambda_0^{(n)} \geqq \lambda_0^{(n-1)} \geqq \lambda_1^{(n)} \geqq \lambda_1^{(n-1)} \cdots \geqq \lambda_{n-1}^{(n-1)} \geqq \lambda_n^{(n)}. \tag{5}$$

The polynomials $f_j^{(n)}$ are real and are orthogonal both on I_a and on I_b. If we scale them so that

$$\int_{I_b} f_j^{(n)}(x) f_k^{(n)}(x)\, dx = \delta_{jk}, \tag{6}$$

which we will henceforth assume done, then

$$\int_{I_a} f_j^{(n)}(x) f_k^{(n)}(x)\, dx = \lambda_j^{(n)} \delta_{jk} \tag{7}$$

for $j, k = 0, 1, \cdots, n$, and for $n = 0, 1, 2, \cdots$. This double orthogonality property, (6) and (7), can also be used to define the $f_j^{(n)}$ and the $\lambda_j^{(n)}$.

Double orthogonality is useful in certain least-squares approximation problems. Suppose that one seeks an nth degree polynomial $f(x)$ to approximate a given function $g(x)$ and that the approximation error is measured by

$$E = w_a \int_{I_a} |f(x) - g(x)|^2\, dx + w_b \int_{I_b} |f(x) - g(x)|^2\, dx,$$

where the constants w_a and w_b are real positive weights. To minimize E, write $f(x)$ as

$$f(x) = \sum_j c_j f_j^{(n)}(x)$$

and obtain the minimizing coefficients

$$c_j = \frac{w_a \int_{I_a} f_j^{(n)} g\, dx + w_b \int_{I_b} f_j^{(n)} g\, dx}{w_a \lambda_j^{(n)} + w_b},$$

$$j = 0, 1, \cdots, n.$$

Without loss of generality, we henceforth choose the interval I_b to be $[-1, 1]$.

The nature of the polynomials $f_j^{(n)}(x)$ depends markedly on whether I_a is disjoint from $I_b = [-1, 1]$ or is contained in it. When I_a is disjoint from I_b, we speak of an *exterior* problem: when $I_a \in I_b$, we speak of an *interior* problem. Figures 1, 2 and 3 show some typical $f_j^{(n)}$ for these two cases for $n = 5$. Note the change of scale necessary within I_a in Figs. 1 and 2 to show detail of $f_4^{(5)}$ and $f_5^{(5)}$. For the case shown in Fig. 3, one clearly has $f_{5-i}^{(5)}(x) = f_i^{(5)}(1-x)$, $\lambda_i^{(5)} = 1/\lambda_{5-i}^{(5)}$ for $i = 0, 1, \cdots, 5$.

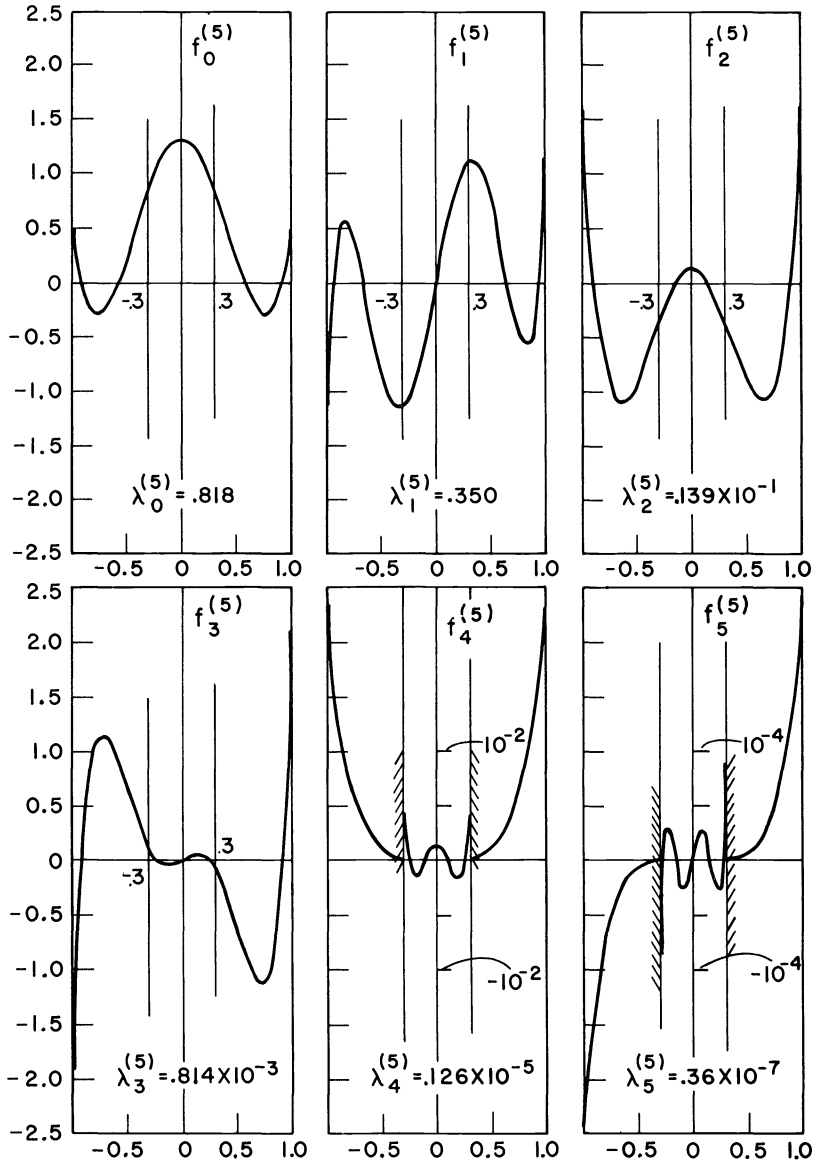


FIG. 1. *Concentrated doubly orthogonal polynomials for symmetric interior case* $I_b = [-1, 1]$, $I_a = [-.3, .3]$, $n = 5$

For the exterior problem with $I_a = [a_1, a]$, where $1 \leqq a_1 < a$, we find the asymptotic result

$$\lambda_0^{(n)} = \frac{[a + \sqrt{a^2 - 1}]^{2n+2}}{8\pi n \sqrt{a^2 - 1}} \left[1 + O\left(\frac{1}{n}\right)\right] \tag{8}$$

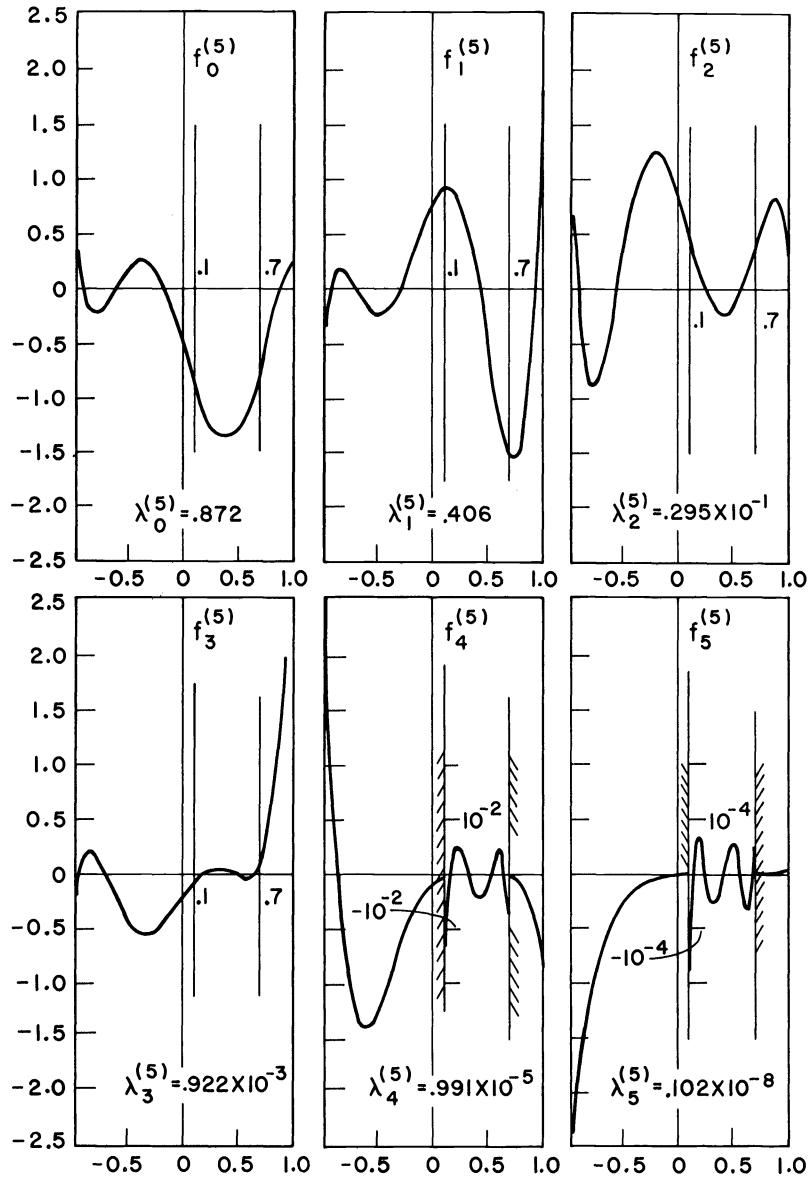


FIG. 2. *Concentrated doubly orthogonal polynomials for asymmetric interior case* $I_b = [-1, 1]$, $I_a = [.1, .7]$, $n = 5$

which does not depend on a_1. For the interior case, $\lambda_0^{(n)} \to 1$ as $n \to \infty$. When I_a is centered in I_b, so that $I_a = [-a, a]$ with $0 < a < 1$, we find the asymptotic result

$$1 - \lambda_j^{(n)} = \frac{4\sqrt{\pi a}}{1+a} \frac{1}{j!} \left(\frac{8a}{1-a^2} \right)^j n^{j+1/2} \left(\frac{1-a}{1+a} \right)^{n+1} \left[1 + O\left(\frac{1}{n} \right) \right] \tag{9}$$

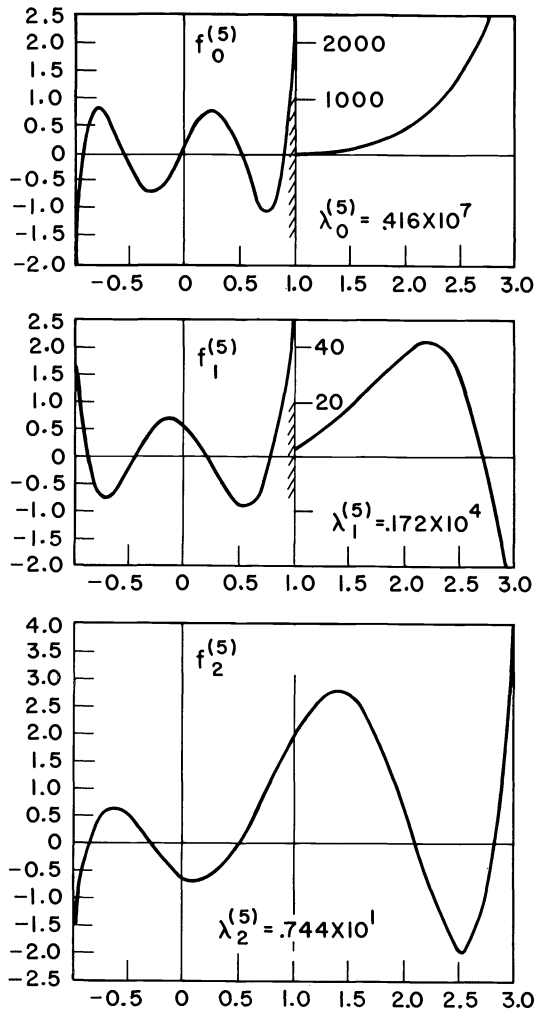


FIG. 3. *Concentrated doubly orthogonal polynomials for exterior case* $I_b = [-1, 1]$, $I_a = [1, 3]$, $n = 5$

for fixed j. The asymptotic form of $f_j^{(n)}$ is given by the complicated expressions (97)–(99). When $I_a \subset I_b$ but is not centered, we have not been able to determine the rate at which $\lambda_j^{(n)} \to 1$.

The determination of the $f_j^{(n)}$ and their concentrations is equivalent to the solution of a matrix eigenvalue problem of the form $A\mathbf{x} = \lambda B\mathbf{x}$ with A and B $(n+1)\times(n+1)$ real symmetric positive definite matrices and $\mathbf{x}$ an $(n+1)$-vector. The polynomials are also the solutions of the integral equation

$$\int_{I_a} K_n(x, y) f(y)\, dy = \lambda f(x), \qquad x \in I_a, \tag{10}$$

where

$$K_n(x, y) = \frac{n+1}{2} \frac{P_{n+1}(x)P_n(y) - P_{n+1}(y)P_n(x)}{x - y} \tag{11}$$

with $P_n(x)$ the usual Legendre polynomial. The corresponding eigenvalues of (10) are the concentrations $\lambda_j^{(n)}$.

In two special cases we have found an equivalent differential equation eigenvalue problem for the concentrated polynomials. For the centered interior case with $I_a = [-a, a]$, the $f_j^{(n)}(x)$ are all the polynomial solutions of

$$\frac{d}{dx}\left[(1-x^2)(a^2-x^2)\frac{df}{dx}\right]+[\chi - n(n+3)x^2]f=0 \tag{12}$$

and the same equation with n replaced by $n-1$. Such polynomial solutions exist only for special values χ_j of the parameter χ. These eigenvalues of (12) do not seem to be related in a simple way to the concentrations $\lambda_j^{(n)}$. For the adjacent exterior problem, where $I_a = [1, a]$, $a > 1$, there is also a formulation as a differential equation:

$$\frac{d}{dx}\left[(x^2-1)(x-a)\frac{df}{dx}\right]+[\chi - n(n+2)x]f=0. \tag{13}$$

We have not been able to find a corresponding differential equation formulation for either the general interior or general exterior problem.

The problem under consideration is a special case of one mentioned by Szegö [1]. Other somewhat similar problems are discussed in the literature, but we have found no reference that treats the concentrated polynomials considered here. Further properties of the polynomials will be found in the sections that follow.

2. Eigenvalues. Let $\varphi_0(x), \varphi_1(x), \cdots, \varphi_n(x)$ be a basis for the space F_n of nth degree polynomials. Each polynomial in F_n has an expression

$$f(x) = \sum_k f_k \varphi_k(x) \tag{14}$$

and so a representation as an $(n+1)$-tuple of coefficients $\mathbf{f} = (f_0, f_1, \cdots, f_n)$. Then $R(f)$ in (1) becomes a ratio of quadratic forms

$$R(f) = (\mathbf{f}^t A \mathbf{f}^*)/(\mathbf{f}^t B \mathbf{f}^*). \tag{15}$$

In (15), $\mathbf{f}$ is regarded as a column vector, $\mathbf{f}^t$ is its transpose, and A and B are $(n+1)\times(n+1)$ matrices having elements $\int \varphi_i(x)\varphi_j^*(x)\,dx$ taken over I_a and I_b, respectively. To maximize (15) one may solve a matrix eigenvalue problem

$$A\mathbf{f} = \lambda B\mathbf{f}. \tag{16}$$

The largest eigenvalue λ of (16) is the maximum ratio $R(f)$, i.e., $\lambda_0^{(n)}$. The coordinates f_k of the corresponding eigenvector $\mathbf{f}$ are the coefficients in (14) of the polynomial $f(x)$ having $R(f) = \lambda_0^{(n)}$, i.e., the coefficients of $f_0^{(n)}(x)$.

The eigenvalue problem (16) is well known [2], [3]. In the present problem, both A and B are positive definite Hermitian matrices since the quadratic forms in (15) are integrals of squared magnitudes of polynomials. From [3] one can find that the eigenvalues $\lambda_j^{(n)}$ of (16) are real, that they satisfy (4) and (5), and that there are real eigenvectors $\mathbf{f}_j^{(n)}$ which can be chosen to satisfy

$$(\mathbf{f}_j^{(n)})^t B(\mathbf{f}_k^{(n)}) = \delta_{jk}, \qquad (\mathbf{f}_j^{(n)})^t A(\mathbf{f}_k^{(n)}) = \lambda_j^{(n)}\delta_{jk}, \tag{17}$$

$$j, k = 0, 1, \cdots, n, \quad n = 0, 1, 2, \cdots.$$

Equations (6) and (7) follow from these relations combined with (14).

Since the eigenfunctions and eigenvalues are real, we restrict attention to real functions from now on. In particular, basis functions $\varphi_k(x)$ in (14) will always be real. Two bases suggest themselves immediately. First, one may use $\varphi_k(x) = x^k$. Then, with I_a of the form $[a_1, a_2]$ and $I_b = [-1, 1]$, the elements of A and B are

$$a_{ij} = \frac{a_2^{i+j+1} - a_1^{i+j+1}}{i+j+1}, \qquad b_{ij} = \frac{1-(-1)^{i+j+1}}{i+j+1}, \quad i, j = 0, 1, \cdots, n. \tag{18}$$

When cast in this form, the problem makes connection with the theory of the Hilbert matrix (with elements $1/(i+j+1)$), which has been widely studied as an example of an ill-conditioned matrix (see [4, p. 233 and p. 236, Prob. 4d] and [5, pp. 22–23]).

A second convenient basis is $\varphi_k(x) = (k+\frac{1}{2})^{1/2} P_k(x)$ where $P_k(x)$ is the kth Legendre polynomial. This basis is orthonormal on I_b and so B in (16) simplifies to the unit matrix. However, A becomes more complicated, having elements which are integrals of products of Legendre polynomials.

For small values of n the eigenvalue problem can be solved numerically by finding the roots of the characteristic equation $\det |A - \lambda B| = 0$. The basis $\varphi_i(x) = x^i$ is as convenient as any. Table 1 gives values of $\lambda_0^{(n)}$ for several choices of I_a and n. Figures 1–3 show some eigenfunctions.

TABLE 1

$\lambda_0^{(n)} = \textit{largest root of } \det(A - \lambda B) = 0$. $I_b = [-1, 1]$.

I_a	[1, 3]	[−.2, .2]	[−.5, .5]	[−.8, .8]	[−.2, .8]
$n = 1$	13.93	.20000	.50000	.80000	.66618
2	254.0	.43057	.86474	.99339	.78712
3	5,875	.43057	.86474	.99339	.96014
4	152,000	.62401	.97576	.999883	.96299
5	4,160,000	.62401	.97576	.999883	.99116
6	1.18×10^8		.99651		

To maximize the ratio $R(f)$ one may also maximize the numerator of (1) subject to the side condition that the denominator have the value 1. That leads to the problem of maximizing

$$\int_{I_a} f^2(x)\,dx - \lambda \int_{I_b} f^2(x)\,dx = \int \{\chi_a(x) - \lambda \chi_b(x)\} f^2(x)\,dx, \tag{19}$$

where λ is a Lagrange multiplier and $\chi_a(x)$ and $\chi_b(x)$ are the characteristic functions of I_a and I_b. The condition that an nth degree polynomial $f(x)$ be an extremal of (19) is just that

$$\int p(x)\{\chi_a(x) - \lambda \chi_b(x)\} f(x)\,dx = 0 \tag{20}$$

must hold for all nth degree polynomials $p(x)$. Again, the condition (20) can hold only for certain eigenvalues λ and corresponding eigenfunctions $f(x)$. Indeed, let

$\varphi_0(x), \cdots, \varphi_n(x)$ be any basis of F_n, orthonormal over I_b, and define

$$K(x, y) = \sum_{j=0}^{n} \varphi_j(x)\varphi_j(y). \tag{21}$$

The substitution $p(x) = K(x, y)$ in (20) puts the problem in the form of an integral equation

$$\int_{I_a} K(y, x)f(x)\,dx = \lambda f(y). \tag{22}$$

The matrix equation (16) is obtained again as the system for the coefficients f_k of (14) when (22) is multiplied by $\varphi_k(x)$ and integrated over I_b.

The integral equation (22) has an interesting interpretation for interior problems. For any function $F(x)$, the integral

$$\int_{I_b} K(x, y)F(y)\,dy = \sum \varphi_j(x) \int_{I_b} \varphi_j(y)F(y)\,dy$$

represents the orthogonal projection of $F(x)$ onto the space F_n of nth degree polynomials defined on I_b. The integral operator on the left side of (22) first truncates $f(x)$ to the interval I_a (equivalent to multiplication by $\chi_a(x)$) and then projects the truncated $f(x)$ back onto the space of polynomials.

Exterior problems can also be interpreted in terms of projections, although not directly from (22). An exterior problem is equivalent to the interior problem of maximizing

$$\int_{I_a} f^2(x)\,dx \Big/ \int_{I_a \cup I_b} f^2(x)\,dx = R(f)/(1+R(f))$$

and hence it has another integral equation involving projections onto the space of polynomials over $I_a \cup I_b$.

When $I_b = [-1, 1]$ and the basis functions φ_k are chosen to be the normalized Legendre functions, (21) and (22) become (10) and (11). Here we have used the identity ([6, 8.9.1, p. 335])

$$(x-y)\sum_{0}^{n}(j+\tfrac{1}{2})P_j(x)P_j(y) = (n+1)[P_{n+1}(x)P_n(y) - P_{n+1}(y)P_n(x)].$$

The double orthogonality property (6)–(7) provides another way of posing the original problem. We seek nth degree polynomials $\psi_0(x), \cdots, \psi_n(x)$ and numbers $\mu_0, \cdots, \mu_n$ which satisfy

$$\int_{I_b} \psi_i(x)\psi_j(x)\,dx = \delta_{ij}, \qquad \int_{I_a} \psi_i(x)\psi_j(x)\,dx = \delta_{ij}\mu_i.$$

But then these ψ_i are suitable basis functions for the nth degree polynomials and so can be used as the φ_i of (14). With this basis, the double orthogonality property puts both A and B in diagonal form: $B = \text{diag}\ (1, 1, \cdots, 1)$, $A = \text{diag}\ (\mu_0, \mu_1, \cdots, \mu_n)$. Then μ_k and $\psi_k(x)$ must be the eigenvalues and eigenfunctions of the original problem.

3. Comparison of Chebyshev polynomials with $f_0^{(n)}$. A definition of a different kind of concentration of a polynomial $f(x)$ might single out a point y where $f(y)$ is to be large, and a set S where $f(x)$ is to remain small. Then measure the concentration by the ratio

$$|f(y)|^2/\max_{x\in S}|f(x)|^2. \tag{23}$$

In our exterior problem, with $I_b=[-1,1]$ and $I_a=[a_1,a_2]$, $y=a_2$ and $S=I_b$ would be appropriate choices. In the interior problem with $I_a=[-a,a]$, one might take $y=0$ and $S=I_b-I_a$.

The maximizations of (1) and (23) lead to totally different problems. We mention (23) here because the maximizing nth degree polynomial in (23) has a large value of $R(f)$ which makes an interesting comparison with $\lambda_0^{(n)}$. It is not hard to show [7, p. 304] that in maximizing (23) over F_n, $f(x)$ must have all its zeros in S and exhibit "equal ripple" behavior there. It then follows that in the exterior problem, the maximizing $f(x)$ is the Chebyshev polynomial

$$T_n(x)=\cos n\theta \qquad (x=\cos\theta).$$

For the interior problem with $I_a=[-a,a]$, the solution for even $n=2N$ is

$$t_n(x)=T_N((1+a^2-2x^2)/(1-a^2)).$$

We omit formal proof.

These polynomials are especially simple because of the identity

$$T_n^2(x)=\cos^2 n\theta=\tfrac{1}{2}+\tfrac{1}{2}T_{2n}(x). \tag{24}$$

To integrate $T_n^2(x)$ for the exterior problem one may use (24) together with

$$\begin{aligned}\int T_{2n}(x)\,dx &= -\int \cos 2n\theta \sin\theta\,d\theta \\ &= -\frac{1}{2}\int(\sin(2n+1)\theta-\sin(2n-1)\theta)\,d\theta \qquad (25)\\ &= \frac{1}{2}((2n+1)^{-1}T_{2n+1}(x)-(2n-1)^{-1}T_{2n-1}(x)).\end{aligned}$$

In (25) the right-hand side is to be evaluated between limits appropriate to the numerator or denominator of (1). For an exterior interval $I_a=[1,a]$, the final result is

$$\int_{-1}^{1} T_n^2(x)\,dx=(4n^2-2)/(4n^2-1), \tag{26}$$

$$\begin{aligned}\int_1^a T_n^2(x)\,dx &= \tfrac{1}{2}(a-1)+\tfrac{1}{4}\{(2n+1)^{-1}(T_{2n+1}(a)-1) \\ &\quad -(2n-1)^{-1}(T_{2n-1}(a)-1)\}.\end{aligned} \tag{27}$$

Using (26), (27) and the recurrence

$$T_{k+1}(x)+T_{k-1}(x)=\cos(k+1)\theta+\cos(k-1)\theta$$
$$=2\cos k\theta\cos\theta \tag{28}$$
$$=2xT_k(x),$$

one can easily tabulate $R(T_n)$.

For the interior problem with $I_a=[-a, a]$, (24) again shows

$$t_{2N}^2(x)=\tfrac{1}{2}+\tfrac{1}{2}t_{4N}(x) \tag{29}$$

while (28) becomes

$$t_{2(k+1)}(x)+t_{2(k-1)}(x)=2(1+a^2-x^2)t_{2k}(x)/(1-a^2). \tag{30}$$

The integral of $t_{4N}(x)$ is not as simple as in (25). Instead, (30) shows that the coefficients $c_{N,k}$ of the polynomial

$$t_{2N}(x)=\sum_{k=0}^{N} c_{N,k}x^{2k}$$

satisfy a recurrence

$$c_{N+1,k}+c_{N-1,k}=(2(1+a^2)c_{N,k}-2c_{N,k-1})/(1-a^2). \tag{31}$$

Using (31) one can easily evaluate coefficients for $t_{4N}(x)$ in (29) and integrate term by term to get

$$\int_{-x}^{x} t_{2N}^2(x)\,dx=x+\sum_{k=0}^{2N} c_{2N,k}x^{2k+1}/(2k+1). \tag{32}$$

Set $x=a$ and $x=1$ to evaluate the integrals needed for $R(t_{2N})$.

Table 2 gives numerical values of $R(T_n)$ and $R(t_{2N})$ which may be compared with Table 1.

TABLE 2
$R(T_n)$ and $R(t_n)$

	$R(T_n)$	$R(t_n)$		
$I_a=$	[1, 3]	[−.2, .2]	[−.5, .5]	[−.8, .8]
$n=1$	13			
2	172.4	.4152	.8593	.99334
3	4,028.1			
4	105,397	.4567	.9625	.999844
5	2,917,953			
	$.837\times10^8$	.5641	.99451	.999998

Asymptotic results for large n follow easily from the formula

$$T_n(x)=\tfrac{1}{2}\{\rho(x)^n+\rho(x)^{-n}\}, \tag{33}$$

where $\rho(x)=\exp i\theta=x+(x^2-1)^{1/2}$. For any $x>1$ the term $\rho(x)^{-n}$ becomes small exponentially with n, leaving $T_n(x)\sim\frac{1}{2}\rho(x)^n$. For the exterior problem with

$I_a = [1, a]$, the numerator of (1) is given by (27). Then (33) may be used to obtain

$$R(T_n) = (a^2-1)^{1/2}(8n)^{-1}\rho(a)^{2n}\{1+O(n^{-1})\}. \tag{34}$$

Note that $\lambda_0^{(n)}$, as given by (8), also grows exponentially and contains a term $\rho(a)^{2n}$. Indeed (34) differs from (8) only by a constant factor.

For the interior problem with $I_a = [-a, a]$, the integrals required for $R(t_{2N})$ may be estimated as follows. In the range $-a \leqq x \leqq a$, write $t_{2N}(x) = \cosh Nu$ where $\cosh u = (1+a^2-2x^2)/(1-a^2)$. Then

$$\int_{-a}^{a} t_{2N}^2(x)\,dx = 2\int_0^a \cosh^2 Nu\,dx \sim \frac{1}{2}\int_0^a \exp(2Nu)\,dx$$

asymptotically as $N \to \infty$. This integral can be handled by the method of steepest descent. Let U be the value of u at $x = 0$, i.e., $\cosh U = (1+a^2)/(1-a^2)$, $\sinh U = 2a/(1-a^2)$. The appropriate variable of integration is $v = U - u$;

$$\int_{-a}^{a} t_{2N}^2(x)\,dx \sim \frac{1}{2}\exp(2NU)\int_0^U \exp(-2Nv)\left|\frac{dx}{dv}\right| dv. \tag{35}$$

Now

$$\begin{aligned} 2x^2 &= 1+a^2-(1-a^2)\cosh u \\ &= 2a\sinh v - (1+a^2)(\cosh v - 1) \end{aligned}$$

so that dx/dv has a series expansion

$$dx/dv = \tfrac{1}{2}(a/v)^{1/2} - 3(a+a^{-1})v^{1/2}/16 + \cdots. \tag{36}$$

Now Watson's lemma [8, p. 218] applies to (35) and gives an asymptotic series when (35) is integrated term by term, after substituting (36). The leading term is

$$\int_{-a}^{a} t_{2N}^2(x)\,dx = 2^{-3}(2\pi a/N)^{1/2}((1+a)/(1-a))^{2N}\{1+O(n^{-1})\}.$$

The energy of $t_{2N}(x)$ outside the interval I_a is

$$\begin{aligned} 2\int_a^1 t_{2N}^2(x)\,dx &= \int_a^1 (1+\cos 2N\theta)\,dx \\ &= 1-a+\int_0^\pi |dx/d\theta|\cos 2N\theta\,d\theta, \end{aligned}$$

where θ is defined again by $(1-a^2)\cos\theta = 1+a^2-2x^2$. As $N \to \infty$ the integral on θ tends to 0 (Riemann–Lebesgue theorem). Finally

$$R(t_n) = 1-4(1+a)(\pi a)^{-1/2}n^{1/2}((1-a)/(1+a))^{n+1}\{1+o(1)\}. \tag{37}$$

Now $1-\lambda_0^{(n)}$ and $1-R(t_n)$, as given by (9) and (37), become small exponentially at the same rate, and differ only by a constant factor.

4. Legendre polynomials and the derivation of equation (8). Still another kind of concentration uses a set S and point y as in (23), but forms the ratio

$$E(y; f) = |f(y)|^2 \Big/ \int_S |f(x)|^2\,dx. \tag{38}$$

The nth degree polynomial $f(x)$ that maximizes $E(y;f)$ differs from the ones that maximize (1) or (23). However, for exterior problems at least, $E(y;f)$ is easy to maximize and the solution will be used to derive the asymptotic result (8). Indeed, the problem of maximizing $E(y;f)$ may be regarded as a limiting form of the one for maximizing $R(f)$ as the interval I_a shrinks to a point y.

To maximize $E(y;f)$, let $\varphi_0(x), \cdots, \varphi_n(x)$ be a basis of F_n, orthonormal over S. For any polynomial $f(x)$ with coefficients $f_0, \cdots, f_n$ in (14),

$$E(y;f) = \left|\sum f_k \varphi_k(y)\right|^2 \Big/ \sum f_k^2 \leqq \sum \varphi_k^2(y). \tag{39}$$

The inequality here follows from Schwarz's inequality, and equality holds only if $f_k = c\varphi_k(y)$ with c independent of k. Thus the function

$$f_y(x) = K(x, y) \tag{40}$$

in (21) is a maximizing function, and the maximum is

$$\max_{f \in F_n} E(y;f) = K(y, y). \tag{41}$$

For the exterior problem with $I_a = [a_1, a]$ $(1 \leqq a_1 < a)$, take $S = I_b = [-1, 1]$. The basis

$$\varphi_k(x) = (k + \tfrac{1}{2})^{1/2} P_k(x)$$

is orthonormal on S and so the maximizing function is

$$\begin{aligned} K(x, y) &= \sum_{k=0}^{n} (k + \tfrac{1}{2}) P_k(x) P_k(y) \\ &= \tfrac{1}{2}(n+1)\{P_{n+1}(x)P_n(y) - P_n(x)P_{n+1}(y)\}/(x-y), \end{aligned} \tag{42}$$

the last line following from Christoffel's identity (see [6, 8.9.1, p. 335]).

The solution (40), (41) can now be used to derive bounds on $\lambda_0^{(n)}$. Since $K(x, a) = f_a(x) \in F_n$,

$$R(f_a) \leqq \lambda_0^{(n)}. \tag{43}$$

Also, (38) with $S = I_b$ shows that

$$R(f) = \int_{I_a} E(y;f)\, dy$$

holds for every f. In particular, take f to be a polynomial maximizing $R(f)$ and use (41) to obtain

$$\lambda_0^{(n)} \leqq \int_{I_a} K(y, y)\, dy = \sum_{k=0}^{n} \int_{I_a} \varphi_k^2(y)\, dy = \sum_{k=0}^{n} R(\varphi_k). \tag{44}$$

The upper bound (44) can also be derived from the theory of the integral equation (22). The integral

$$\int_{I_a} K(y, y)\, dy$$

is actually the sum of the eigenvalues $\lambda_j^{(n)}$, and so is an upper bound on the largest eigenvalue $\lambda_0^{(n)}$.

When n is large the bounds (43) and (44) on $\lambda_0^{(n)}$ are close. This can be shown with the help of Darboux' asymptotic formula for $P_n(x)$ for large n. When $1 \leqq x \leqq a$, let $x = \cosh u$ so that $\exp u = \rho(x) = x + (x^2-1)^{1/2}$ as in (33). Darboux' formula (see [8, p. 285]) states that

$$\begin{aligned} P_n(x) &= (2n\pi \sinh u)^{-1/2} \exp[(n+\tfrac{1}{2})u]\{1+O(n^{-1})\} \\ &= [n\pi(1-\rho(x)^{-2})]^{-1/2}\rho(x)^n\{1+O(n^{-1})\}. \end{aligned} \tag{45}$$

To estimate the lower bound $R(f_a)$, let U and U_1 satisfy $\cosh U = a$, or $\exp U = \rho(a)$, and $\cosh U_1 = a_1$. Then write

$$\begin{aligned} \int_{-1}^{1} f_a^2(x)\,dx &= \sum_{k=0}^{n} (k+\tfrac{1}{2})P_k^2(a) \\ &= (\pi^{-1}\rho(a)^{2n}/(1-\rho(a)^{-2}))\{1+O(n^{-1})\}, \end{aligned}$$

$$\begin{aligned} \int_1^a f_a^2(x)\,dx &= \sum_{j,k} (j+\tfrac{1}{2})(k+\tfrac{1}{2})P_j(a)P_k(a)\int_{U_1}^{U} P_j(x)P_k(x)\sinh u\,du \\ &= \sum (2\pi^2(k+j+1))^{-1}\rho(a)^{2k+2j+1}\{1+O(k^{-1}+j^{-1})\} \\ &= (4n\pi^2)^{-1}\rho(a)^{4n+1}(1-\rho(a)^{-2})^{-2}\{1+O(n^{-1})\}. \end{aligned}$$

The final result is

$$R(f_a) = (8n\pi(a^2-1)^{1/2})^{-1}\rho(a)^{2n+2}\{1+O(n^{-1})\}, \tag{46}$$

in agreement with (8).

Likewise, in the upper bound (44) the term $R(\varphi_k) = R(P_k)$ is

$$\begin{aligned} R(P_k) &= (k+\tfrac{1}{2})\int_{U_1}^{U} P_k^2(x)\sinh u\,du \\ &= (k+\tfrac{1}{2})\int_{U_1}^{U} (2k\pi)^{-1}\exp(2k+1)u\,du\,\{1+O(k^{-1})\} \\ &= (\rho(a)^{2k+1}/(4\pi k))\{1+O(k^{-1})\}. \end{aligned}$$

The upper bound (44) on $\lambda_0^{(n)}$ is now

$$\sum_{k=0}^{n} R(P_k) = (\rho(a)^{2n+1}/(4n\pi(1-\rho(a)^{-2})))\{1+O(n^{-1})\} \tag{47}$$

which can be rewritten in the form (46) by using the identity $\rho(a)-\rho(a)^{-1} = 2(a^2-1)^{1/2}$. Since the bounds (46) and (47) agree to within a factor $1+O(n^{-1})$, the result (8) follows.

Bounds like (43) and (44) also hold for the interior problem. $S = I_b - I_a$ is the appropriate set in (38). Inequalities (43) and (44) may be rederived with two alterations. One is that these bounds now relate to

$$R'(f) = \int_{I_a} f^2(x)\,dx \Big/ \int_S f^2(x)\,dx = R(f)/(1-R(f)).$$

With $I_a = [-a, a]$ one would use $f_0(x) = K(x, 0)$ to derive $R'(f_0) \leqq \lambda_0^{(n)}/(1-\lambda_0^{(n)})$, in (43), or $R(f_0) \leqq \lambda_0^{(n)}$. The analogue to (44) is

$$\lambda_0^{(n)}/(1-\lambda_0^{(n)}) \leqq \sum_{k=0}^{n} R'(\varphi_k).$$

The other alteration requires a new basis $\varphi_0(x), \cdots, \varphi_n(x)$, which now must be orthonormal over $I_b - I_a$. Thus $K(x, y)$ no longer has a simple expression (42) in terms of Legendre functions. The derivations of (46) and (47) relied heavily on special properties of Legendre polynomials and do not generalize directly to the interior problem. It is not clear whether or not (9) for $j = 0$ can also be derived from the analogues of (43) and (44). Fortunately, the differential equation obtained in § 6 provides another way of deriving (9).

5. An alternate path to (8). We now outline briefly another derivation of (8) that is of interest in its own right.

When $I_b = [-1, 0]$, $I_a = [0, a]$, $a > 0$, and $\varphi_k = x^k$, the matrices in (16) have elements

$$a_{ij} = \frac{a^{i+j+1}}{i+j+1}, \qquad b_{ij} = \frac{(-1)^{i+j+1}}{i+j+1}, \tag{48}$$

$i, j = 0, 1, \cdots, n$. Now it is easy to invert B. The elements of the inverse are (see [4, p. 233, p. 263, prob. 4d] or [5, p. 23])

$$(B^{-1})_{ij} = \frac{g_i^{(n)} g_j^{(n)}}{i+j+1}, \qquad g_i^{(n)} = \frac{(n+i+1)!}{(n-i)![i!]^2}, \qquad i, j = 0, 1, \cdots, n. \tag{49}$$

Equation (16) in the form $B^{-1}A\mathbf{f} = \lambda\mathbf{f}$ now reads

$$\sum_{jk} \frac{g_i^{(n)} g_j^{(n)} a^{i+j+1}}{(i+j+1)(j+k+1)} f_k = \lambda f_i,$$

$i = 0, 1, \cdots, n$. By direct substitution one verifies that this is the same as

$$T^2\mathbf{x} = \lambda\mathbf{x}, \tag{50}$$

where the $(n+1)\times(n+1)$ matrix T has elements

$$t_{ij} = \sqrt{a}\,\frac{l_i^{(n)} l_j^{(n)}}{i+j+1}, \qquad i, j = 0, 1, \cdots, n, \tag{51}$$

and

$$l_i^{(n)} \equiv \sqrt{a^i g_i^{(n)}}, \qquad x_i \equiv f_i / l_i^{(n)}, \qquad i = 0, 1, \cdots, n. \tag{52}$$

The matrix T is real and symmetric. It is not hard to show that it is positive definite. Equation (50) is equivalent to the well-studied eigenvalue problem

$$T\mathbf{x} = \theta\mathbf{x}, \qquad \lambda = \theta^2, \tag{53}$$

with which we now work.

The key to finding an asymptotic expression for $\theta_0^{(n)}$, the largest eigenvalue of T, is the observation that for large n the elements, t_{ij}, of T are large only when i and j are near the value $n\sqrt{a}/(1+a)$. More precisely, by straightforward tech-

niques using Stirling's formula for the factorials in (49), one finds that

$$l_{\delta n}^{(n)} \sim \frac{(1+\delta_0)^{3/4}}{\sqrt{2\pi\delta_0}(1-\delta_0)^{1/4}} e^{cn} e^{-(\delta-\delta_0)^2 n/(2\sigma^2)} \tag{54}$$

provided

$$\delta - \delta_0 = o(n^{-1/3}).$$

Here

$$\delta_0 \equiv \sqrt{\frac{a}{1+a}}, \quad \sigma^2 \equiv \delta_0^3/a, \quad c \equiv \log[\sqrt{a}+\sqrt{1+a}]. \tag{55}$$

Thus $l_i^{(n)}$ will be large only for i in the range $\delta_0 n - \sigma\sqrt{2dn} \leqq i \leqq \delta_0 n + \sigma\sqrt{2dn}$ for some large $d > 0$.

With i and j restricted to the ranges where $l_i^{(n)}$ and $l_j^{(n)}$ are large, the factor $1/(i+j+1)$ in (51) has the value $(2\delta_0 n)^{-1}[1+O(n^{-1/2})]$. Thus asymptotically T behaves like a singular matrix $\hat{T}$ with elements

$$\hat{t}_{ij} = \frac{\sqrt{a}}{2\delta_0 n} l_i^{(n)} l_j^{(n)}.$$

The largest eigenvalue and the corresponding eigenvector of such a matrix can be found at once. One has

$$\hat{\theta}_0 = \frac{\sqrt{a}}{2\delta_0 n} \sum_0^n [l_i^{(n)}]^2, \tag{56}$$

$$\hat{x}_i = l_i^{(n)}, \qquad i = 0, 1, \cdots, n. \tag{57}$$

The validity of the heuristically derived asymptotic forms (56) and (57) can be established rigorously as follows. The largest eigenvalue of T, say $\theta_0^{(n)}$, is bounded by ([5, p. 10])

$$M_- \equiv \frac{\sum t_{ij} u_i u_j}{\sum u_i^2} \leqq \theta_0^{(n)} \leqq M_+ \equiv \sqrt{\sum t_{ij}^2} \tag{58}$$

for every real $(n+1)$-vector $\mathbf{u} = (u_0, u_1, \cdots, u_n)$. As suggested by (57), we take $u_i = l_i^{(n)}$ in (58). Then some detailed analysis, omitted here, shows that for large n,

$$\begin{aligned} M_+ \sim M_- \sim \theta_0^{(n)} &\sim \frac{\sqrt{a}}{2\delta_0 n} \sum_i [l_i^{(n)}]^2 \\ &\sim \frac{[\sqrt{a}+\sqrt{a+1}]^{n+1}}{4\sqrt{\pi n}[a(1+a)]^{1/4}}. \end{aligned} \tag{59}$$

Here to find the asymptotic forms one can make use of (54). Now $\lambda_0^n = [\theta_0^{(n)}]^2$ as shown by (53). To compare the result (59) with (8), we must also take account of the fact that for (59) $I_b = [-1, 0]$, $I_a = [0, a]$ whereas for (8), $I_b = [-1, 1]$ and $I_a = [a_1, a]$. Replacing a in (59) by $(a-1)/2$ and squaring, one obtains (8) again. It is not hard to see that $\lambda_0^{(n)}$ for the intervals $I_b = [-1, 0]$, $I_a = [a_1, a]$, $0 < a_1 < a$ will have the same asymptotic value as that just found; i.e., the asymptotic value is independent of a_1.

Perhaps the most interesting feature of this derivation is the asymptotic formula (54) for the coefficients $\{l_i^{(n)}\}^2$ of the maximizing polynomial

$$f(x) = \sum_{i=0}^{n} \{l_i^{(n)}\}^2 x^i$$

when $I_b = [-1, 0]$. When n is large the coefficients follow a Gaussian distribution with mean $[a/(1+a)]^{1/2}n$ and standard deviation $\sigma(n/2)^{1/2}$. Thus the largest coefficient in the polynomial does not belong to x^n as might have been supposed.

An attempt to carry out a similar analysis for the interior case is thwarted by the fact that the analog of the matrix T now has elements t_{ij} whose sign varies as $(-1)^j$ and so no asymptotic form for the largest eigenvector is evident.

6. Differential equations. The derivation of the differential equations (12) and (13) requires some properties of the eigenvalues $\lambda_j^{(n)}$. Here we give the details for the interior problem $I_b = [-1, 1]$, $I_a = [-a, a]$; the exterior problem is less complicated.

We first investigate conditions under which the same number λ can be both an eigenvalue $\lambda_j^{(n)}$ and $\lambda_k^{(m)}$ for two different degrees n, m. Table 1 shows instances of $\lambda_0^{(n)} = \lambda_0^{(n+1)}$.

In (20) with $I_a = [-a, a]$ and $I_b = [-1, 1]$ it is clear that $f(-x)$ is always an eigenfunction if $f(x)$ is. Then if $f(x)$ be decomposed into a sum $f(x) = e(x) + o(x)$ of an even function and an odd function, both $e(x)$ and $o(x)$ will be eigenfunctions belonging to the same eigenvalue as $f(x)$. Thus a complete basis of eigenfunctions can always be constructed from functions which are either even or odd.

If n is even and $e(x)$ is one of the even eigenfunctions $f_j^{(n)}(x)$, then (20) holds with $\lambda = \lambda_j^{(n)}$, $f(x) = e(x)$, and $p(x) \in F_n$. But also one can set $p(x) = x^{n+1}$, an odd function, and then it follows that (20) holds when $p(x) \in F_{n+1}$. Thus $\lambda_j^{(n)}$ and $e(x)$ reappear among the list of eigenvalues $\lambda_k^{(n+1)}$ and eigenfunctions $f_k^{(n+1)}(x)$. Likewise if n is odd, each odd eigenfunction $o(x)$ is both an $f_j^{(n)}(x)$ and an $f_k^{(n+1)}(x)$. When n is even, the eigenfunctions are the $n/2$ odd eigenfunctions $f_j^{(n-1)}(x)$ and $\frac{1}{2}n + 1$ new even eigenfunctions. When n is odd, the eigenfunctions are the $\frac{1}{2}(n+1)$ even eigenfunctions $f_j^{(n-1)}(x)$ and $\frac{1}{2}(n+1)$ new odd eigenfunctions. The separation of the eigenvalue problem into two smaller problems, one for even eigenfunctions and the other for odd, is also evident in the matrix formulation (16), (18). The matrices A, B are decomposable and (16) factors into separate eigenvalue problems for the even and odd eigenfunctions.

Whether n is even or odd, $\lambda_0^{(n)}$ is always an eigenvalue belonging to one of the even eigenfunctions. For, any odd eigenfunction $o(x)$ can be written as $xe(x)$

where $e(x)$ is even. But

$$R(o) = \frac{1}{1+[\int_a^1 x^2e^2\,dx]/[\int_0^a x^2e^2\,dx]}$$

$$< \frac{1}{1+[\int_a^1 a^2e^2\,dx]/[\int_0^a a^2e^2\,dx]} = R(e).$$

Thus an odd function never maximizes $R(f)$. Since the even eigenfunctions for n even are the same as those for $n+1$, it is now clear that $\lambda_0^{(n+1)} = \lambda_0^{(n)}$ as observed in Table 1.

To derive the differential equation we first need to show that an eigenvalue $\lambda = \lambda_j^{(n)}$ never reappears as an eigenvalue $\lambda_k^{(n+2)}$ where the corresponding eigenfunctions $f_j^{(n)}(x)$ and $f_k^{(n+2)}(x)$ have the same parity. To prove this by contradiction, suppose the contrary, say, $f_j^{(n)}(x) = e_n(x)$ and $f_k^{(n+2)}(x) = e_{n+2}(x)$ are even eigenfunctions belonging to the same eigenvalue λ. Because even eigenfunctions for odd degree are the same as the even eigenfunctions for the preceding degree, we may take n to be even.

If $e_{n+2}(x) \in F_n$, then $e_{n+2}(x)$ is an eigenfunction for degree n as well as $n+2$. If $e_{n+2}(x)$ does not belong to F_n, then the following argument shows that $e_n(x)$ is an eigenfunction for degree $n+2$. First apply (20) with $f(x) = e_{n+2}(x)$ and $p(x) = e_n(x)$,

$$\int e_{n+2}(x)\{\chi_a(x) - \lambda\chi_b(x)\}\, e_n(x)\, dx = 0. \tag{60}$$

Next

$$\int x^{n+1}\{\chi_a(x) - \lambda\chi_b(x)\} e_n(x)\, dx = 0 \tag{61}$$

because $e_n(x)$ is even and x^{n+1} is odd. Finally

$$\int p(x)\{\chi_a(x) - \lambda\chi_b(x)\} e_n(x)\, dx = 0 \tag{62}$$

for all polynomials $p(x) \in F_n$. But $e_{n+2}(x)$ does not belong to F_n; then every polynomial in F_{n+2} is a linear combination of $e_{n+2}(x)$, x^{n+1}, and a $p(x)$ in F_n. Thus (60), (61), (62) combine to prove that (20) holds with $f(x) = e_n(x)$ and $p(x) \in F_{n+2}$; i.e., $e_n(x)$ is an eigenfunction for degree $n+2$.

Now one eigenfunction $e(x)$, either $e_n(x)$ or $e_{n+2}(x)$, is proved to belong to F_n and be an eigenfunction for both degrees n and $n+2$. But consider (20) for degree $n+2$, with $f(x) = e(x)$ and $p(x) = (x^2 - a^2)e(x) \in F_{n+2}$. Since $\lambda = R(e)$, one has $0 < \lambda < 1$ and $(x^2 - a^2)e^2(x)(\chi_a(x) - \lambda\chi_b(x)) \leqq 0$. Then the integral in (20) is negative, which is the desired contradiction to the assumption that $e_n(x)$ and $e_{n+2}(x)$ have the same eigenvalue. A similar contradiction can be derived for odd eigenfunctions.

We may now derive the differential equation for the eigenfunctions. Consider an even eigenfunction $f(x)$ and take n to be the even degree for which $f(x)$ is an eigenfunction. Since $f(x)$ is even, the condition (20) holds more generally for

polynomials $p(x) \in F_{n+1}$. Let $L(\cdot)$ denote the differential operator

$$L(f) \equiv \frac{d}{dx}\left[(1-x^2)(a^2-x^2)\frac{df}{dx}\right] - n(n+3)x^2 f \tag{63}$$

and note that $L(q) \in F_{n+1}$ for all polynomials $q(x) \in F_{n-1}$. Thus

$$\int L(q)(\chi_a(x) - \lambda\chi_b(x))f(x)\,dx = 0.$$

After two integrations by parts, this equation assumes the form

$$\int q(x)\{\chi_a(x) - \lambda\chi_b(x)\}L(f)\,dx = 0. \tag{64}$$

Also

$$\int q(x)\{\chi_a(x) - \lambda\chi_b(x)\}f(x)\,dx = 0 \tag{65}$$

follows because $f(x)$ is an eigenfunction of degree n and $q(x) \in F_{n-1} \subset F_n$. Now $f(x) \in F_n$ but $f(x) \notin F_{n-2}$ (because $f(x)$ cannot be an eigenfunction for both degrees n and $n-2$). Also $L(f)$ is even and the term $n(n+3)x^2 f$ in $L(f)$ has been included to make $L(f) \in F_n$. Then some linear combination $L(f) - \chi f(x)$ has degree $n-2$. But (64) and (65) show that

$$\int q(x)\{\chi_a(x) - \lambda\chi_b(x)\}[L(f) - \chi f(x)]\,dx = 0$$

for all $q(x) \in F_{n-1}$. Either the differential equation

$$L(f) = \chi f \tag{66}$$

holds identically or else λ has even eigenfunctions $f(x) \in L_n$ and $L(f) - \chi f(x) \in L_{n-2}$. But the latter alternative has been proved to be impossible. A similar argument rederives the differential equation (66) for the odd eigenfunctions, in which case n in (63) is an odd number.

Solutions to the differential equation (66) can be found in the form of a power series

$$f(x) = \sum c_k x^k$$

containing only even or odd powers according to the parity of n. The recurrence

$$\begin{aligned} a^2(k+2)(k+1)c_{k+2} &= [(1+a^2)k(k+1) - \chi]c_k \\ &\quad + [n(n+3) - (k-2)(k+1)]c_{k-2}, \end{aligned} \tag{67}$$

where $c_j \equiv 0$ for $j < 0$ then determines the coefficients as polynomials in χ. In general the series does not terminate. However, for special values of χ, $c_{n+2} = 0$. Then (67) with $k = n+2$ shows also that $c_{n+4} = 0$ and all higher order coefficients must vanish. Thus χ appears in (66) as an eigenvalue which is determined by the condition $c_{n+2} = 0$ to make $f(x) \in F_n$.

A polynomial equation with the eigenvalues χ as roots can be found. If $c_{n+2} = c_{n+4} = \cdots = 0$, the recurrence (67) becomes a square system of linear

homogeneous equations in $\cdots, c_{n-4}, c_{n-2}, c_n$ whose determinant, a function of χ, must vanish. For example, if $n=2m$, the consistency of (67) for $c_0, c_2, \cdots, c_{2m}$ requires that

$$(68)\qquad \begin{vmatrix} \alpha_0-\chi & \beta_0 & 0 & 0 & \cdots & 0 & 0 \\ \gamma_1 & \alpha_1-\chi & \beta_1 & 0 & \cdots & 0 & 0 \\ 0 & \gamma_2 & \alpha_2-\chi & \beta_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \alpha_{m-1}-\chi & \beta_{m-1} \\ 0 & 0 & 0 & 0 & \cdots & \gamma_m & \alpha_m-\chi \end{vmatrix} = 0.$$

The tri-diagonal matrix here has elements $\alpha_j=(1+a^2)2j(2j+1)$, $\beta_j=-a^2(2j+1)(2j+2)$, $\gamma_j=2m(2m+3)-(2j-2)(2j+1)$. The $m+1$ roots of (68), $\chi_j^{(n)}, j=0, 1, \cdots, m$, give rise to $m+1$ even polynomial solutions of (66) of degree $n=2m$. The other m concentrated polynomials belonging to the family $f_j^{(2m)}$ are the odd polynomial solutions of (66) with n there replaced by $2m-1$.

The eigenvalue λ depends on χ only indirectly. Having determined an eigenfunction $f(x)$ for an eigenvalue χ of (66), one obtains λ from $\lambda=R(f)$.

The exterior problem with $I_a=[1, a]$ also has a differential equation as given in (13). The proof is simpler and requires showing that no λ can be an eigenvalue for both degrees n and $n+1$. We omit the details.

7. Asymptotics for the differential equation (12). We now investigate the solutions of the differential equation

$$(69)\qquad \frac{d}{dx}\left[A(x)\frac{df}{dx}\right]+(\chi-n(n+3)x^2)f=0,$$
$$A(x)\equiv(1-x^2)(a^2-x^2),$$

as n becomes large. We seek solutions that are continuous for all finite values of x. Our techniques follow closely those used in [9] to investigate the asymptotic behavior of the prolate spheroidal wave functions. Different asymptotic expressions will be found for (69) in different intervals on the x-axis. Much of the analysis will be concerned with joining these separate pieces so that they correspond to a single continuous asymptotic solution. We shall frequently omit details of straightforward calculations. We proceed in a purely formal manner without investigating the convergence of assumed series solutions.

LEMMA 1. *If*

$$\frac{d}{dx}\left[L\frac{d}{dx}f\right]+Mf=0$$

and $f=g/\sqrt{L}$, *then*

$$\frac{d^2g}{dx^2}+\left[\frac{M}{L}+\frac{1}{4}\left(\frac{L'}{L}\right)^2-\frac{1}{2}\frac{L''}{L}\right]g=0.$$

LEMMA 2. *If*

$$\frac{d^2g}{dx^2}-\left[n^2E^2(x)+nF(x)+\sum_0^\infty \frac{G_j(x)}{n^j}\right]g=0$$

with E, F and the G's independent of n, then

$$g=\frac{1}{\sqrt{E}}\left[c_1\exp\left(-n\int E\,dx\right)\exp\left(-\tfrac{1}{2}\int\frac{F}{E}\,dx\right)+c_2\exp\left(n\int E\,dx\right)\exp\left(\tfrac{1}{2}\int\frac{F}{E}\,dx\right)\right]\cdot\left[1+O\left(\frac{1}{n}\right)\right]$$

for fixed x, not a zero of E.

Lemma 1 is proved by direct substitution and differentiation. Lemma 2, the WKB approximation, is established formally by substituting $g=e^{nw}$ in the differential equation to find

$$n^2[w'^2-E^2]+n[w''-F]-\sum_0\frac{G_j}{n^j}=0.$$

Write $w=\sum_0 w_j(x)/n^j$ here with w_j independent of n. Equate to zero the coefficients of separate powers of n. The resulting equations can be solved readily to give explicit forms for w_0 and w_1. The Lemma follows directly.

We return now to investigate (69) near $x=0$. The substitution $x=\sqrt{a/2nt}$ and division by $2na$ yields

$$\frac{d^2f}{dt^2}+\left[\frac{\chi}{2na}-\frac{t^2}{4}\right]f+O\left(\frac{1}{n}\right)=0.$$

Assume the series solution $f=\sum_0 u_j/n^j$ and $\chi/2na=\sum_0 c_j/n^j$ with the u_j and c_j independent of n. One then finds $u_0=D_j(t)$, $c_0=j+\frac{1}{2}$ for $j=0, 1, 2, \cdots$. Here D_j is the Weber function. See [10, vol. II, chap. 8]. We thus have

$$f(t)=D_j(t)+O\left(\frac{1}{n}\right), \tag{70}$$

$$\chi\equiv\chi_j=(2j+1)na+O(1). \tag{71}$$

The remainder in (70) is $O(1/n)$ for fixed t, that is, for $x=O(1/\sqrt{n})$. Investigation of the higher order terms, much as is done in [9, pp. 111, 112 and 119] shows that for

$$0\leqq x\leqq\frac{1}{n^{1/3}},$$

$$f(x)\equiv f_0(x)=D_j\left(\sqrt{\frac{2n}{a}}x\right)+O\left(\frac{1}{n^{1/3}}\right). \tag{72}$$

The Weber function $D_j(t)$ has precisely j zeros.

We note that from the asymptotic expansion for $D_j(t)$ (see [10, vol. II, eqs. (8.4.1), p. 122]) we have

$$f_0\left(\frac{u}{n^{1/3}}\right)=\left[\sqrt{\frac{2}{a}}n^{1/6}\right]^j u^j e^{-(1/2)(u^2/a)n^{1/3}}\left[1+O\left(\frac{1}{n^{1/3}}\right)\right] \tag{73}$$

which gives the value of the asymptotic solution (72) at the edge of its range of validity.

Equation (71) gives the asymptotic form for the eigenvalue χ in (69). This in turn permits us to find asymptotic expressions for f in ranges other than the one (72) close to the origin. Indeed, the substitutions (71) and $f = v/\sqrt{A}$ into (69) reduce it to

$$\frac{d^2v}{dx^2} - \frac{P(x)}{(1-x^2)^2(a^2-x^2)^2}v = 0,$$

where $P(x)$ is a polynomial of degree 3 in x^2. For large n one finds

$$\frac{d^2v}{dx^2} - n^2\frac{(x^2-x_0^2)(x^2-x_a^2)(x^2-x_1^2)}{(1-x^2)^2(a^2-x^2)^2}v = 0, \tag{74}$$

where

$$\begin{aligned} x_0 &= \sqrt{\frac{(2j+1)a}{n}}\left[1+O\left(\frac{1}{n}\right)\right], \\ x_a &= a - \frac{A}{n^2} + o\left(\frac{1}{n^2}\right), \\ x_1 &= 1 + \frac{B}{n^2} + o\left(\frac{1}{n^2}\right), \end{aligned} \tag{75}$$

where A and B are positive and independent of n. For $x \geqq 0$, the coefficient of v in (74) is positive only in the intervals $0 \leqq x < x_0$ and $x_a < x < x_1$ and the solution can exhibit oscillations only in these ranges. As n increases, the turning points x_a and x_1 approach a and 1, respectively, which are singular points of (69). Thus we are led to make separate investigations of (69) between each pair of turning points and in small neighborhoods of a and 1.

To investigate the solution of (69) between x_0 and x_a, we substitute (71) for χ in (69) and write $f = g/\sqrt{A}$ to obtain

$$\frac{d^2g}{dx^2} - \left[n^2\frac{x^2}{A} + n\frac{[3x^2-(2j+1)a]}{A} + O(1)\right]g = 0$$

on using Lemma 1. From Lemma 2, with $E = x/\sqrt{A}$ and $F = [3x^2-(2j+1)a]/A$, we find $\int E\,dx = \frac{1}{2}\log[1+a^2-2x^2-2\sqrt{A}]$, $\int (F/E)\,dx = \frac{3}{2}\log[1+a^2-2x^2-2\sqrt{A}] + \frac{1}{2}(2j+1)\log[(2a^2-(1+a^2)x^2+2a\sqrt{A})/x^2]$, so that

$$f_1(x) = \frac{g}{\sqrt{A}} = k_1\frac{x^j[1+a^2-2x^2-2\sqrt{A}]^{-(2n+3)/4}}{A^{1/4}[2a^2-(1+a^2)x^2+2a\sqrt{A}]^{(2j+1)/4}}, \tag{76}$$

$$\frac{1}{n^{1/3}} \leqq x \leqq a - \frac{1}{n}.$$

Here, of the two terms in Lemma 2 we have chosen the one that decays with n, as we must if we are to have agreement with (73) at $x = u/n^{1/3}$. Indeed, setting $x = u/n^{1/3}$ in (76) and using the fact that $[1+a^2-2x^2-2\sqrt{A}] =$

$(1-a)^2[1+x^2/a+O(x^4)]$, one finds that

$$f_1\left(\frac{u}{n^{1/3}}\right)\sim k_1\frac{(1-a)^{-(2n+3)/2}}{n^{j/3}2^{(2j+1)/2}a^{(2j+2)/2}}u^j e^{-(1/2)(u^2/a)n^{1/3}}.$$

Comparison with (73) yields

$$k_1=2^{(3j+1)/2}a^{j/2+1}(1-a)^{(2n+3)/2}n^{j/2}. \tag{77}$$

For $x=a-v/n$, one finds from (76) after some straightforward asymptotics,

$$f_1\left(a-\frac{v}{n}\right)\sim\frac{k_1a^j(1-a^2)^{-(2n+3)/4}\exp\sqrt{2av/(1-a^2)}n^{1/2}}{[(1-a^2)2a(v/n)]^{1/4}[a^2(1-a^2)]^{(2j+1)/4}}. \tag{78}$$

To investigate (69) near $x=a$, it is convenient to set $y=\sqrt{a^2-x^2}$ to find

$$\begin{aligned}(a^2-y^2)(1-a^2+y^2)\frac{d^2f}{dz^2}-\frac{1}{y}[-a^2(1-a^2)+y^2(2-5a^2)+4y^4]\frac{df}{dy}\\ +[\chi-n(n+3)(a^2-y^2)]f=0.\end{aligned} \tag{79}$$

Now let $y=t/n$, use (71) and divide by $n^2a^2(1-a^2)$ to obtain

$$\frac{d^2f}{dt^2}+\frac{1}{t}\frac{df}{dt}-\frac{1}{1-a^2}f+O\left(\frac{1}{n}\right)=0. \tag{80}$$

We thus are led to $f(t)=I_0(t/\sqrt{1-a^2})+O(1/n)$ with I_0 the usual Bessel function notation. In terms of our original variables, then, we take

$$f_2(x)=k_2I_0\left(\sqrt{\frac{a^2-x^2}{1-a^2}}\right),\qquad a-\frac{1}{n}\leqq x\leqq a. \tag{81}$$

The asymptotic expression for I_0 ([10, vol. II, eq. (7.13.5), p. 86]) shows that

$$\begin{aligned}f_2\left(a-\frac{v}{n}\right)&\sim k_2I_0\left(\sqrt{n}\sqrt{\frac{2av}{1-a^2}}\right)\\ &\sim k_2\frac{1}{\sqrt{2\pi}}\left(\frac{1-a^2}{2a}\right)^{1/4}n^{-1/4}v^{-1/4}\exp\sqrt{\frac{2av}{1-a^2}}n^{1/2}.\end{aligned}$$

Comparison with (78) gives

$$k_2=\pi^{1/2}2^{1/2}a^{-1/2}(1-a^2)^{-(n+j+3)/2}n^{1/2}k_1. \tag{82}$$

For $x\geqq a$, we write $z=\sqrt{x^2-a^2}=iy$. When the obvious changes are then made in (79) and (80), one finds

$$f_3(x)=k_2J_0\left(n\sqrt{\frac{x^2-a^2}{1-a^2}}\right),\qquad a\leqq x\leqq a+\frac{1}{n}, \tag{83}$$

which joins with f_2 at $x = a$. The asymptotic formula for J_0 ([6, 9.2.1, p. 364]) gives

$$f_3\left(a+\frac{v}{n}\right) \sim k_2 J_0\left(\sqrt{n}\sqrt{\frac{2av}{1-a^2}}\right)$$

(84)

$$\sim k_2\sqrt{\frac{2}{\pi}}\left(\frac{1-a^2}{2av}\right)^{1/4} n^{-1/4} v^{-1/4} \cos\left[\sqrt{n}\sqrt{\frac{2av}{1-a^2}} - \frac{\pi}{4}\right].$$

For $a \leqq x \leqq 1$, we define

$$\bar{A}(x) = (1-x^2)(x^2-a^2) = -A(x)$$

and rewrite (69) in the form

$$\frac{d}{dx}\left(\bar{A}\frac{df}{dx}\right) + [n(n+3)x^2 - \chi]f = 0. \tag{85}$$

Set $f = g/\sqrt{\bar{A}}$. Lemma 1 yields

$$\frac{d^2g}{dx^2} + \left[\frac{n(n+3)x^2}{\bar{A}} - \frac{na(2j+1)}{\bar{A}} + O(1)\right]g = 0.$$

We next invoke Lemma 2 with

$$E = \frac{ix}{\sqrt{\bar{A}}}, \qquad F = \frac{(2j+1)a - 3x^2}{\bar{A}}.$$

Straightforward integration yields

$$f_4(x) = \frac{k_4 \cos\left[\frac{2n+3}{4}\arcsin\theta - \frac{2j+1}{4}\arcsin\varphi + \rho\right]}{\sqrt{x}\bar{A}^{1/4}} \tag{86}$$

where

$$\theta(x) = \frac{2x^2-(1+a^2)}{1-a^2}, \qquad \varphi(x) = \frac{(1+a^2)x^2 - 2a^2}{x^2(1-a^2)}.$$

Here ρ and k_4 are constants which we proceed to determine by making (86) agree with (84) at $x = a + v/n$. We find by straightforward expansion that

$$\theta\left(a+\frac{v}{n}\right) = -1 + \frac{4av}{(1-a^2)}\frac{1}{n} + O\left(\frac{1}{n^2}\right),$$

$$\varphi\left(a+\frac{v}{n}\right) = -I + \frac{8v}{a(1-a^2)}\frac{1}{n} + O\left(\frac{1}{n^2}\right).$$

Thus, using the formula [6, p. 81, 4.4.41]

$$\arcsin(1-z)=\frac{\pi}{2}-\sqrt{2z}\left[1+\sum a_k z^k\right],$$

we arrive at

$$f_4\left(a+\frac{v}{n}\right)\sim k_4\frac{\cos\left[\sqrt{n}\sqrt{\frac{2av}{1-a^2}}+\frac{2n-2j+2}{4}\left(-\frac{\pi}{2}\right)+\rho\right]}{\sqrt{a}(1-a^2)^{1/4}\left(\frac{2av}{n}\right)^{1/4}}.$$

Comparison with (84) gives

$$\begin{aligned} k_4&=\pi^{-1/2}2^{1/2}a^{1/2}(1-a^2)^{1/2}n^{-1/2}k_2,\\ \rho&=(n-j)\frac{\pi}{4}. \end{aligned}\tag{87}$$

We use the solution (86)–(87) for $a+1/n\leqq x\leqq 1-1/n$. The value near this latter point is readily computed to be

$$f_4\left(1-\frac{v}{n}\right)\sim\frac{k_4\cos\left[-\sqrt{n}\sqrt{\frac{2v}{1-a^2}}+\frac{2n+3-2j-1}{4}\frac{\pi}{2}+(n-j)\frac{\pi}{4}\right]}{\sqrt{2}(1-a^2)^{1/4}\left(\frac{2v}{n}\right)^{1/4}}.\tag{88}$$

Near $x=1$, we write $z=\sqrt{1-x^2}$ and (85) becomes

$$\begin{aligned}(1-z^2)(1-a^2-z^2)\frac{d^2f}{dz^2}+\left[\frac{1-a^2}{z}+(2a^2-5)z+4z^3\right]\frac{df}{dz}\\ +[n(n+3)(1-z^2)-\chi]f=0.\end{aligned}\tag{89}$$

Let $z=t/n$, use (71) for χ and divide by $n^2(1-a^2)$ to find

$$\frac{d^2f}{dt^2}+\frac{1}{t}\frac{df}{dt}+\frac{1}{1-a^2}f+O\left(\frac{1}{n}\right)=0.$$

Reasoning as before, we set

$$f_5(x)=k_5J_0\left(n\sqrt{\frac{1-x^2}{1-a^2}}\right),\qquad 1-\frac{1}{n}\leqq x\leqq 1.\tag{90}$$

The asymptotic formula for J_0 gives

$$f_5\left(1-\frac{v}{n}\right)\sim k_5\sqrt{\frac{2}{\pi}}n^{-1/4}2^{-1/4}(1-a^2)^{1/4}v^{-1/4}\cos\left(\sqrt{n}\sqrt{\frac{2v}{1-a^2}}-\frac{\pi}{4}\right).$$

Comparison with (88) shows that agreement is possible only if $n-j$ is even, say $n-j=2m$, and if

$$k_5=(-1)^m2^{-1}(1-a^2)^{-1/2}n^{1/2}\pi^{1/2}k_4, \qquad n-j=2m. \tag{91}$$

When $x>1$, the solution (90) passes into

$$f_6(x)=k_5I_0\left(n\sqrt{\frac{x^2-1}{1-a^2}}\right), \qquad 1\leqq x\leqq 1+\frac{1}{n},$$

which has the asymptotic value

$$f_6\left(1+\frac{v}{n}\right)\sim k_5\frac{n^{-1/4}(1-a^2)^{1/2}2^{-1/4}}{\sqrt{2\pi}v^{1/4}}\exp\sqrt{n}\sqrt{\frac{2v}{1-a^2}}. \tag{92}$$

When $x\geqq 1+1/n$, solution proceeds very much as in obtaining (76), but now we must take the solution that grows with n and so find

$$f_7(x)=k_7\frac{[2\sqrt{A}+2x^2-1-a^2]^{(2n+3)/4}[(1+a^2)x^2-2a^2-2a\sqrt{A}]^{(2j+1)/4}}{A^{1/4}x^{j+1}}, \qquad x\geqq 1+\frac{1}{n}. \tag{93}$$

Evaluation at $x=1+v/n$ gives

$$f_7\left(1+\frac{v}{n}\right)\sim k_7\frac{(1-a^2)^{(2n+3)/4}(1-a^2)^{(2j+1)/4}\exp\sqrt{n}\sqrt{\frac{2v}{1-a^2}}}{(1-a^2)^{1/4}\left(\frac{2v}{n}\right)^{1/4}}. \tag{94}$$

From (92), then, it follows that

$$k_7=\pi^{-1/2}2^{-1/2}(1-a^2)^{-(n+j+1)/2}n^{-1/2}k_5. \tag{95}$$

For large x, one finds from (93) that

$$f(x)\sim f_7(x)\sim k_7 2^{n+3/2}(1-a)^{j+1/2}x^n. \tag{96}$$

The preceding results are summarized by (97). With j and x fixed, and $n-j$ even, for large n,

$$f_j^{(n)}(x)\sim\begin{cases} D_j\left(\sqrt{\dfrac{2n}{a}}x\right), & 0\leqq x\leqq n^{-1/3},\\ k_1\dfrac{x^jB^{-(2n+3)/4}}{A^{1/4}C^{(2j+1)/4}}, & n^{-1/3}\leqq x\\ k_2I_0\left(n\sqrt{\dfrac{a^2-x^2}{1-a^2}}\right), & a-\dfrac{1}{n}\leqq x\leqq a,\\ k_2J_0\left(n\sqrt{\dfrac{x^2-a^2}{1-a^2}}\right), & a\leqq x\leqq a+\dfrac{1}{n},\\ k_4\dfrac{\cos\left(\dfrac{2n+3}{4}\arcsin\theta-\dfrac{2j+1}{4}\arcsin\varphi+(n-j)\dfrac{\pi}{4}\right)}{\sqrt{x}[-A]^{1/4}}, & a+\dfrac{1}{n}\leqq x\leqq 1-\dfrac{1}{n},\\ k_5J_0\left(n\sqrt{\dfrac{1-x^2}{1-a^2}}\right), & 1-\dfrac{1}{n}\leqq x\leqq 1,\\ k_5I_0\left(n\sqrt{\dfrac{x^2-1}{1-a^2}}\right), & 1\leqq x\leqq 1+\dfrac{1}{n},\\ k_7\dfrac{[-B]^{(2n+3)/4}[-C]^{(2j+1)/4}}{A^{1/4}x^{j+1}}, & 1+\dfrac{1}{n}\leqq x, \end{cases} \tag{97}$$

$$A\equiv(1-x^2)(a^2-x^2),$$

$$B\equiv 1+a^2-2x^2-2\sqrt{A},$$

$$C\equiv 2a^2-(1+a^2)x^2+2a\sqrt{A},$$

$$\theta\equiv[2x^2-(1+a^2)]/(1-a^2),$$

$$\varphi\equiv[(1+a^2)x^2-2a^2]/(x^2(1-a^2)),$$

$D_j(x)$, the Weber function, J_0 and I_0 Bessel functions. The constants are given by

$$k_i=\pi^{Y_1/2}2^{Y_2/2}a^{Y_3/2}(1-a)^{Y_4/2}(1+a)^{Y_5/2}n^{Y_6/2}(-1)^{Y_7/2} \tag{98}$$

with values in the table in (99).

(99)

i	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7
1	0	$3j+1$	$j+2$	$2n+3$	0	j	0
2	1	$3j+2$	$j+1$	$n-j$	$-n-j-3$	$j+1$	0
4	0	$3j+3$	$j+2$	$n-j+1$	$-n-j-2$	j	0
5	1	$3j+1$	$j+2$	$n-j$	$-n-j-3$	$j+1$	$n-j$
7	0	$3j$	$j+2$	$-2j-1$	$-2n-2j-4$	j	$n-j$

8. Derivation of (9). The asymptotic formula (9) for $\lambda_j^{(n)}$ for large n and fixed j for the symmetric interior problem can be derived by direct integration of the results given by (97), (98) and (99). One has

$$\begin{aligned} 1-\lambda_j^{(n)} &= \int_a^1 [f_j^{(n)}(x)]^2\,dx \Big/ \int_0^1 [f_j^{(n)}(x)]^2\,dx \\ &\sim \int_a^1 [f_j^{(n)}(x)]^2\,dx \Big/ \int_0^a [f_j^{(n)}(x)]^2\,dx. \end{aligned} \tag{100}$$

We indicate how these integrals can be evaluated giving a minimum of detail.

Proceeding in sections as given by (97)–(99), we have

$$\begin{aligned} L_0 \equiv \int_0^{n^{-1/3}} [f_j^{(n)}(x)]^2\,dx &\sim \int_0^{n^{-1/3}} D_j^2\left(\sqrt{\frac{2n}{a}}x\right) dx \\ &= \sqrt{\frac{a}{2n}}\left[\int_0^\infty D_j^2(t)\,dt - \int_{n^{1/6}\sqrt{2/a}}^\infty D_j^2(t)\,dt\right] \\ &= \sqrt{\frac{a}{2n}}\left[\sqrt{\frac{\pi}{2}}j! + O(n^{(j-1)/6}\,e^{-n^{1/3}/(2a)})\right] \\ &= \frac{1}{2}\sqrt{\frac{a\pi}{n}}j!\left[1+O\left(\frac{1}{n}\right)\right]. \end{aligned} \tag{101}$$

Here we use (8.3.23)[1] and (8.4.1) from [10, p. 122] to evaluate the integrals.

To bound

$$L_1 = \int_{n^{-1/2}}^{a-1/n} [f_j^{(n)}(x)]^2\,dx \sim k_1^2 \int_{n^{-1/2}}^{a-1/n} \frac{x^{2j}\,e^{-(2n+3)/2\log B}}{A^{1/2}C^{(2j+1)/2}}\,dx \tag{102}$$

we compare $y \equiv \log B(x) = \log[1+a^2-2x^2-2\sqrt{(1-x^2)(a^2-x^2)}]$ and $z(x) \equiv \log(1-a)^2 + \log[1+x^2/a]$. One finds $y(0) = z(0)$ and that for $x \geqq 0$,

$$y'(x) = 2x/\sqrt{(1-x^2)(a^2-x^2)} \geqq z'(x) = 2x/(a+x^2) \geqq 0,$$

so that for $x \geqq 0$, z is nondecreasing and is not greater than $\log B$. We thus overbound the integral on the right of (102) by replacing $\log B$ by $z(x)$. Each factor in the integrand is then monotone and evaluating each factor at the appropriate endpoint, we upper bound the integral by

$$\begin{aligned} \int_{n^{-1/3}}^{a-1/n} \frac{x^{2j}e^{-(n+3/2)\log B}}{A^{1/2}C^{(2j+1)/2}}\,dx &\leqq \int_{n^{-1/3}}^{a-1/n} \frac{x^{2j}\,e^{-(n+3/2)z(x)}}{A(x)^{1/2}C(x)^{(2j+1)/2}}\,dx \\ &\leqq a\frac{a^{2j}\,e^{-(n+3/2)z(n^{-1/3})}}{A(a-1/n)^{1/2}C(a)^{(2j+1)/2}} \\ &= O\left(\frac{\sqrt{n}\,e^{-n^{1/3}/a}}{(1-a)^{2n+3}}\right). \end{aligned}$$

[1] The lower limit in (8.3.23) should read $-\infty$ (cf. [11, p. 351]).

Since from (98)–(99), $k_1^2 = O((1-a)^{2n+3})$, we see that asymptotically $L_1 = O(\sqrt{n}\, e^{-n^{1/3}/a})$ and hence is negligible compared to L_0.

For the next region we have simply

$$L_2 \equiv \int_{a-1/n}^{a} [f_j^n(x)]^2\, dx \sim k_2^2 \int_{a-1/n}^{a} I_0^2\left(n\sqrt{\frac{a^2-x^2}{1-a^2}}\right) dx$$

$$\leqq k_2^2 \frac{1}{n} I_0^2\left(n\sqrt{\frac{2a}{n(1-a^2)}\left[1-\frac{1}{2an}\right]}\right)$$

$$\sim \frac{k_2^2}{2\pi n^{3/2}} \sqrt{\frac{1-a^2}{2a}} \exp \sqrt{n}\sqrt{\frac{8a}{1-a^2}},$$

on using first the fact that $I_0(x)$ is monotone increasing, and next using the asymptotic form for $I_0(x)$ for large x [10, (7.13.5), p. 86]. But from (98)–(99),

$$k_2^2 = O\left(\exp\left(-n \log \frac{1+a}{1-a}\right)\right),$$

so that L_2 is also much smaller than L_0. For the denominator of (100) we now have

$$\int_0^a [f_j^{(n)}(x)]^2\, dx = L_0 + L_1 + L_2 = L_1\left[1 + o\left(\frac{1}{n}\right)\right]$$

(103)

$$= \frac{1}{2}\sqrt{\frac{a\pi}{n}}\, j!\left[1 + o\left(\frac{1}{n}\right)\right].$$

We now proceed to estimate the numerator of (100).

One has

$$L_3 \equiv \int_a^{a+1/n} [f_j^{(n)}(x)]^2\, dx \sim k_2^2 \int_a^{a+1/n} J_0^2\left(n\sqrt{\frac{x^2-a^2}{1-a^2}}\right) dx \leqq k_2^2 \frac{1}{n},$$

since $0 \leqq J_0^2(x) \leqq 1$ for all x. Thus

$$L_3 = O\left(\frac{1}{n} k_2^2\right) = O\left(\frac{1}{n} \exp\left(-n \log \frac{1+a}{1-a}\right)\right).$$

In like manner

$$L_5 \equiv \int_{1-1/n}^{1} [f_j^{(n)}(x)]^2\, dx \sim k_5^2 \int_{1-1/n}^{1} J_0^2\left(n\sqrt{\frac{1-x^2}{1-a^2}}\right) dx$$

$$= O\left(\frac{1}{n} k_5^2\right) = O\left(\frac{1}{n} \exp\left(-n \log \frac{1+a}{1-a}\right)\right).$$

The main contribution to the numerator of (100) comes from

$$
\begin{aligned}
L_4 &\equiv \int_{a+1/n}^{1-1/n} [f_j^{(n)}(x)]^2\,dx \\
&\sim k_4^2 \int_{a+1/n}^{1-1/n} \frac{\cos^2\left(\frac{2n+3}{4}\arcsin\theta - \frac{2j+1}{4}\arcsin\varphi + (n-j)\frac{\pi}{4}\right)}{x[-A]^{1/2}}\,dx \\
&\sim k_4^2 \frac{1}{2}\int_{a+1/n}^{1-1/n} \frac{dx}{x\sqrt{(1-x^2)(x^2-a^2)}} \\
&= \frac{1}{2}k_4^2\frac{1}{2a}\arcsin\frac{(1+a^2)x^2-2a^2}{x^2(1-a^2)}\Bigg|_{a+1/n}^{1-1/n} \\
&\sim \frac{\pi}{4a}k_4^2 = O\left(n^j \exp\left(-n\log\left(\frac{1+a}{1-a}\right)\right)\right).
\end{aligned}
$$

Using the results just obtained for L_3 and L_5, we find that

$$
\begin{aligned}
\int_a^1 [f_j^{(n)}(x)]^2\,dx = L_3+L_4+L_5 &\sim L_5\left[1+O\left(\frac{1}{n}\right)\right] \\
&\sim \frac{\pi}{4a}k_4^2\left[1+O\left(\frac{1}{n}\right)\right] \\
&\sim \frac{\pi 2^{3j+1}a^{j+1}}{(1+a)(1-a^2)^j}n^j\left(\frac{1-a}{1+a}\right)^{n+1}\left[1+O\left(\frac{1}{n}\right)\right].
\end{aligned} \tag{104}
$$

Finally, combining this result with (100) and (103), we have

$$
1-\lambda_j^{(n)} \sim \frac{\sqrt{\pi}2^{3j+2}a^{j+1/2}}{(1+a)(1-a^2)^j j!}n^{j+1/2}\left(\frac{1-a}{1+a}\right)^{n+1}\left[1+O\left(\frac{1}{n}\right)\right],
$$

which is (9). We note that for the function $f_j^{(n)}$ given asympotically by (97)–(99)

$$
\int_{-1}^{1} [f_j^{(n)}(x)]^2\,dx \sim \sqrt{\frac{a\pi}{n}}j!\left[1+O\left(\frac{1}{n}\right)\right] \tag{105}
$$

from (103) and (104).

Acknowledgment. We are grateful to our colleagues, B. F. Logan and H. J. Landau for many helpful discussions during the course of this work and for many direct contributions to it. Among other contributions, Mr. Landau suggested the problem and Mr. Logan discovered the differential equation (12) (by a route quite different from that taken here).

REFERENCES

[1] G. SZEGÖ, *On some Hermitian forms associated with two given curves of the complex plane*, Trans. Amer. Math. Soc., 40 (1936), pp. 450–461.

[2] M. BÔCHER, *Introduction to Higher Algebra*, Macmillan, New York, 1936.

[3] F. P. GANTMACHER AND M. G. KREIN, *Oscillation matrices and kernels and small vibrations of mechanical systems*, State Publishing House for Technical Literature, Moscow, 1950; English trans., AEC TR-4481, Office of Tech. Services, Dept. of Commerce, Washington, D.C., pp. 60–90.

[4] A. RALSTON, *A First Course in Numerical Analysis*, McGraw-Hill, New York, 1965.

[5] M. MARCUS, *Basic Theorems in Matrix Theory*, National Bureau of Standards Applied Math. Series 57, National Bureau of Standards, Washington, D.C., 1960.

[6] M. ABRAMOWITZ AND I. A. STEGUN, *Handbook of Mathematical Functions*, Dover, New York, 1965.

[7] S. KARLIN AND W. J. STUDDEN, *Tchebycheff Systems: With Applications in Analysis and Statistics*, Interscience, New York, 1966.

[8] E. J. COPSON, *Theory of Functions of a Complex Variable*, Oxford University Press, London, 1935.

[9] D. SLEPIAN, *Some asymptotic expansions for prolate spheroidal wave functions*, J. Math. and Physics, 44 (1965), pp. 99–140.

[10] A. ERDÉLYI, *Higher Transcendental Functions*, vol. 2, McGraw-Hill, New York, 1953.

[11] E. T. WHITTAKER AND G. N. WATSON, *Modern Analysis*, Macmillan, New York, 1947.

SIAM J. MATH. ANAL.
Vol. 8, No. 2, April 1977

SOME EXPANSIONS AND CONVOLUTION FORMULAS RELATED TO MacMAHON'S MASTER THEOREM*

L. CARLITZ†

Abstract. The writer has previously applied MacMahon's master theorem to obtain the expansion

$$(*)\qquad \sum_{m_i=0}^{\infty} (\bar{m}_1+\alpha_1)^{m_1}\cdots(\bar{m}_n+\alpha_n)^{m_n}\frac{u_1^{m_1}\cdots u_n^{m_n}}{m_1!\cdots m_n!} = e^{\alpha_1 x_1+\cdots+\alpha_n x_n}(\Delta(x_1, x_2, \cdots, x_n))^{-1},$$

where $\bar{m}_j=\sum_{i=1}^n m_i a_{ij}$ $(j=1, 2, \cdots, n)$, $\Delta(x_1, x_2, \cdots, x_n)=\det(\delta_{ij}-x_i a_{ij})$ and

$$(**)\qquad u_i = x_i \exp\left\{-\sum_{j=1}^n a_{ij}x_j\right\} \qquad (i=1, 2, \cdots, n).$$

In the present paper a number of related results are obtained. These include the inverse of (**) and certain convolution formulas, one of which can be viewed as an n-dimensional extension of Abel's generalization of the binomial theorem. In addition "factorial" analogs of these results are also derived.

1. Introduction.

Let

$$(1.1)\qquad A=(a_{ij}) \qquad (i, j=1, 2, \cdots, n)$$

denote an $n\times n$ array of real or complex numbers and put

$$(1.2)\qquad X_i=\sum_{j=1}^n a_{ij}x_j \qquad (i=1, 2, \cdots, n).$$

MacMahon's master theorem [6, pp. 93–123] asserts that if $m_1, m_2, \cdots, m_n$ are arbitrary nonnegative integers, the coefficient of $x_1^{m_1}x_2^{m_2}\cdots x_n^{m_n}$ in $X_1^{m_1}X_2^{m_2}\cdots X_n^{m_n}$ is equal to the coefficient of $x_1^{m_1}x_2^{m_2}\cdots x_n^{m_n}$ in the expansion of $(\Delta(x_1, x_2, \cdots, x_n))^{-1}$, where

$$(1.3)\qquad \Delta(x_1, x_2, \cdots, x_n)=\det(\delta_{ij}-x_i a_{ij}).$$

The writer has applied this theorem to prove the following result [1]:

$$(1.4)\qquad \sum_{m_1,\cdots,m_n=0}^{\infty} (\bar{m}_1+\alpha_1)^{m_1}\cdots(\bar{m}_n+\alpha_n)^{m_n}\frac{u_1^{m_1}\cdots u_n^{m_n}}{m_1!\cdots m_n!} = e^{\alpha_1 x_1+\cdots+\alpha_n x_n}(\Delta(x_1, x_2, \cdots, x_n))^{-1},$$

where $\alpha_1, \alpha_2, \cdots, \alpha_n$ are arbitrary and

$$(1.5)\qquad \bar{m}_j=\sum_{i=1}^n m_i a_{ij} \qquad (j=1, 2, \cdots, n),$$

$$(1.6)\qquad u_i = x_i \exp\left\{-\sum_{j=1}^n a_{ij}x_j\right\} \qquad (i=1, 2, \cdots, n).$$

* Received by the editors December 29, 1975.

† Department of Mathematics, Duke University, Durham, North Carolina 27706. This work was supported in part by the National Science Foundation under Grant GP 37924X.

In a letter to the writer, I. J. Good has indicated a proof of (1.4) using the n-dimensional extension of the Lagrange expansion [2]. Good had proved the master theorem in this way in [3].

In the present paper we consider a number of related results. In the first place, we shall obtain the inverse of (1.6). Since the results for arbitrary n are rather complicated it seems advisable to first treat the case $n = 2$ and then give a brief statement of the general situation. Thus we shall show in particular that if

$$u = x\,e^{-ax-by}, \qquad v = y\,e^{-cx-dy}, \tag{1.7}$$

where a, b, c, d are arbitrary, then

$$\begin{aligned} x &= \sum_{m=1}^{\infty}\sum_{n=0}^{\infty} bm(am+cn)^{m-1}(bm+dn)^{n-1}\frac{u^m v^n}{m!n!}, \\ y &= \sum_{m=0}^{\infty}\sum_{n=1}^{\infty} cn(am+cn)^{m-1}(bm+dn)^{n-1}\frac{u^m v^n}{m!n!}. \end{aligned} \tag{1.8}$$

We also obtain certain convolution formulas, in particular,

$$\begin{aligned} &\sum_{r=0}^{m}\sum_{s=0}^{n}\binom{m}{r}\binom{n}{s}(br\alpha+cs\beta+\alpha\beta)(ar+cs+\alpha)^{r-1}(br+ds+\beta)^{s-1} \\ &\quad\cdot(a(m-r)+c(n-s)+\alpha')^{m-r}(bm+dn+\beta')^{n-s} \\ &\qquad = (am+cn+\alpha+\alpha')^{m}(bm+dn+\beta+\beta')^{n}. \end{aligned} \tag{1.9}$$

For the corresponding results involving a single summation see [4], [5], [8, Chap. 4]. For the general situation see Theorem 3 below.

We may evidently view (1.9) as a two-dimensional extension of Abel's generalization of the binomial theorem [8, Chap. 1].

In the next place we consider "factorial" analogs of the above. To begin with, we have the following analog of (1.4):

$$\begin{aligned} &\sum_{m_1,\cdots,m_n=0}^{\infty}(\bar{m}_1+\alpha_1)_{m_1}\cdots(\bar{m}_n+\alpha_n)_{m_n}\frac{u_1^{m_1}\cdots u_n^{m_n}}{m_1!\cdots m_n!} \\ &\qquad = (1+x_1)^{\alpha_1}\cdots(1+x_n)^{\alpha_n}(\Delta(x_1, x_2, \cdots, x_n))^{-1}, \end{aligned} \tag{1.10}$$

where $\bar{m}_j$, $\Delta(x_1, x_2, \cdots, x_n)$ have the same meaning as above,

$$(a)_m = a(a+1)\cdots(a+m-1)$$

and

$$u_j = x_j\prod_{i=1}^{n}(1+x_i)^{a_{ij}+\delta_{ij}\alpha_j} \qquad (j = 1, 2, \cdots, n). \tag{1.11}$$

Corresponding to (1.7) and (1.8), we have the following result. Let

$$\begin{aligned} u &= x(1+x)^{-a}(1+y)^{-b}, \\ v &= y(1+x)^{-c}(1+y)^{-d}. \end{aligned} \tag{1.12}$$

Then

$$x = \sum_{m=1}^{\infty}\sum_{n=0}^{\infty} \frac{b}{bm+dn}\binom{am+cn}{m-1}\binom{bm+dn}{n}u^m v^n,$$
$$y = \sum_{m=0}^{\infty}\sum_{n=1}^{\infty} \frac{c}{am+cn}\binom{am+cn}{m}\binom{bm+dn}{n-1}u^m v^{n-1}. \tag{1.13}$$

The factorial analog of (1.9) is given by

$$\sum_{r=0}^{m}\sum_{s=0}^{n}\binom{ar+cs+d-1}{r}\binom{br+ds+\beta-1}{s}A_{m-r,n-s}(\alpha',\beta') = \binom{am+cn+\alpha+\alpha'-1}{m}\binom{bm+dn+\beta+\beta'-1}{n}; \tag{1.14}$$

we have also

$$\sum_{r=0}^{m}\sum_{s=0}^{n} A_{r,s}(\alpha,\beta)A_{m-r,n-s}(\alpha',\beta') = A_{m,n}(\alpha+\alpha',\beta+\beta'), \tag{1.15}$$

where

$$A_{m,n}(\alpha,\beta) = \frac{bm\alpha+cn\beta+\alpha\beta}{(am+cn+\alpha)(bm+dn+\beta)}\binom{am+cn+\alpha}{m}\binom{bm+dn+\beta}{n}. \tag{1.16}$$

For the general case see Theorems 7 and 8 below.

A curious result implied by (1.7) and (1.8) may be noted. If we take $a = d = 0$, (1.7) becomes

$$u = xe^{-by}, \qquad v = ye^{-cx}, \tag{1.17}$$

while (1.8) becomes

$$x = \sum_{m=1}^{\infty}\sum_{n=0}^{\infty}(cn)^{m-1}(bm)^n\frac{u^m v^n}{m!n!},$$
$$y = \sum_{m=0}^{\infty}\sum_{n=1}^{\infty}(cn)^m(bm)^{n-1}\frac{u^m v^n}{m!n!}. \tag{1.18}$$

Similarly (1.12) and (1.13) reduce to

$$u = x(1+y)^{-b}, \qquad v = y(1+x)^{-c} \tag{1.19}$$

and

$$x = \sum_{m=1}^{\infty}\sum_{n=0}^{\infty}\frac{1}{m}\binom{cn}{m-1}\binom{bm}{n}u^m v^n,$$
$$y = \sum_{m=0}^{\infty}\sum_{n=1}^{\infty}\frac{1}{n}\binom{cn}{m}\binom{bm}{n-1}u^m v^n, \tag{1.20}$$

respectively.

2. The case $n = 2$. We have

$$\sum_{m,n=0}^{\infty} (am+cn+\alpha)^m (bm+dn+\beta)^n \frac{u^m v^n}{m!n!} = e^{\alpha x+\beta y}/\Delta, \tag{2.1}$$

where

$$\Delta = \Delta(x, y) = 1 - ax - dy + (ad-bc)xy \tag{2.2}$$

and

$$u = x\, e^{-ax-by}, \qquad v = y\, e^{-cx-dy}. \tag{2.3}$$

Differentiation of (2.1) with respect to α and β gives

$$\sum_{m=r}^{\infty}\sum_{n=s}^{\infty} (am+cn+\alpha)^{m-r}(bm+dn+\beta)^{n-s}\frac{u^m v^n}{(m-r)!(n-s)!} = e^{\alpha x+\beta y}\frac{x^r y^s}{\Delta}.$$

Since, by (2.2),

$$x^r y^s = \frac{1}{\Delta}\{x^r y^s - a x^{r+1} y^s - d x^r y^{s+1} + (ad-bc)x^{r+1}y^{s+1}\},$$

it follows that

$$\begin{aligned} x^r y^s e^{\alpha x+\beta y} = \sum_{m=r}^{\infty}\sum_{n=s}^{\infty} &(am+cn+\alpha)^{m-r-1}(bm+dn+\beta)^{n-s-1} \\ &\cdot \frac{u^m v^n}{(m-r)!(n-s)!}\{(am+cn+\alpha)(bm+dn+\beta) - a(m-r) \\ &\qquad \cdot (bm+dn+\beta) \\ &\qquad - d(n-s)(am+cn+\alpha) + (ad-bc)(m-r)(n-s)\}. \end{aligned}$$

The quantity within braces $\{\cdots\}$ is equal to

$$(ar+cn+\alpha)(bm+ds+\beta) - bc(m-r)(n-s),$$

so that

$$\begin{aligned} x^r y^s e^{\alpha x+\beta y} = \sum_{m=r}^{\infty}\sum_{n=s}^{\infty} &\{(ar+cn+\alpha)(bm+ds+\beta) - bc(m-r)(n-s)\} \\ &\cdot (am+cn+\alpha)^{m-r-1}(bm+dn+\beta)^{n-s-1} \\ &\cdot \frac{u^m v^n}{(m-r)!(n-s)!} \qquad (r, s \geqq 0). \end{aligned} \tag{2.4}$$

In particular, for $(r, s) = (0, 0), (1, 0), (0, 1)$, (2.4) *yields*

$$\begin{aligned} e^{\alpha x+\beta y} = \sum_{m,n=0}^{\infty} &(bm\alpha + cn\beta + \alpha\beta)(am+cn+\alpha)^{m-1} \\ &\cdot (bm+dn+\beta)^{n-1}\frac{u^m v^n}{m!n!}, \end{aligned} \tag{2.5}$$

$$xe^{\alpha x+\beta y}=\sum_{m=1}^{\infty}\sum_{n=0}^{\infty}\{(a+\alpha)(bm+\beta)+(b+\beta)cn\}$$

$$\cdot m(am+cn+\alpha)^{m-2}(bm+dn+\beta)^{n-1}\frac{u^m v^n}{m!n!}, \tag{2.6}$$

$$ye^{\alpha x+\beta y}=\sum_{m=0}^{\infty}\sum_{n=1}^{\infty}\{(d+\beta)(cn+\alpha)+(c+\alpha)bm\}$$

$$\cdot n(am+cn+\alpha)^{m-1}(bm+dn+\beta)^{n-2}\frac{u^m v^n}{m!n!}. \tag{2.7}$$

Specializing further we take $\alpha=\beta=0$. Then (2.6) and (2.7) reduce to

$$x=\sum_{m=1}^{\infty}\sum_{n=0}^{\infty}bm(am+cn)^{m-1}(bm+dn)^{n-1}\frac{u^m v^n}{m!n!} \tag{2.8}$$

and

$$y=\sum_{m=0}^{\infty}\sum_{n=1}^{\infty}cn(am+cn)^{m-1}(bm+dn)^{n-1}\frac{u^m v^n}{m!n!}, \tag{2.9}$$

respectively. To sum up, we state

THEOREM 1. *Given*

$$u=xe^{-ax-by},\qquad v=ye^{-cx-dy}, \tag{2.10}$$

then (2.4), (2.5), (2.6), (2.7) *hold, where* $a, b, c, d, \alpha, \beta$ *are arbitrary. Moreover* (2.8) *and* (2.9) *furnish the inverse of* (2.10).

By means of (2.1) and (2.5) we obtain certain convolution formulas. In the first place, if we replace α, β in (2.5) by α', β' and then multiply each side of the resulting identity by the corresponding side of (2.5) we get

$$\sum_{r=0}^{m}\sum_{s=0}^{n}\binom{m}{r}\binom{n}{s}(br\alpha+cs\beta+\alpha\beta)(ar+cs+\alpha)^{r-1}$$

$$\cdot(br+ds+\beta)^{s-1}\cdot(b(m-r)\alpha'+c(n-s)\beta'+\alpha'\beta') \tag{2.11}$$

$$\cdot(a(m-r)+c(n-s)+\alpha)^{m-r-1}(b(m-r)+d(n-s)+\beta)^{n-s-1}$$

$$=(bm(\alpha+\alpha')+cn(\beta+\beta')+(\alpha+\alpha')(\beta+\beta')(am+cn+\alpha+\alpha')^{m-1}$$

$$\cdot(bm+dn+\beta+\beta')^{n-1}.$$

It follows similarly from (2.1) and (2.5) that

$$\sum_{r=0}^{m}\sum_{s=0}^{n}\binom{m}{r}\binom{n}{s}(br\alpha+cs\beta+\alpha\beta)(ar+cs+\alpha)^{r-1}(br+ds+\beta)^{s-1}$$

$$\cdot(a(m-r)+c(n-s)+\alpha')^{m-r}(b(m-r)+d(n-s)+\beta')^{n-s} \tag{2.12}$$

$$=(am+cn+\alpha+\alpha')^{m}(bm+dn+\beta+\beta')^{n}.$$

Note that when $b=c=0$, (2.10) becomes

$$u=xe^{-ax},\qquad v=ye^{-dy}, \tag{2.13}$$

while (2.8) and (2.9) reduce to

$$x = \sum_{m=1}^{\infty} \frac{(am)^{m-1}}{m!} u^m, \qquad y = \sum_{n=1}^{\infty} \frac{(dn)^{n-1}}{n!} v^n, \tag{2.14}$$

in agreement with [7, p. 125, no. 209].

If we assume only that $b = 0$, (2.10) becomes

$$u = x e^{-ax}, \qquad v = y e^{-cx-dy}, \tag{2.15}$$

while (2.8) and (2.9) become

$$\begin{aligned} x &= \sum_{m=1}^{\infty} \frac{(am)^{m-1}}{m!} u^m, \\ y &= \sum_{m=0}^{\infty} \sum_{n=1}^{\infty} cn(am+cn)^{m-1}(dn)^{n-1} \frac{u^m v^n}{m!n!}. \end{aligned} \tag{2.16}$$

3. The general case. We have, by (1.4),

$$\sum_{m_1,\cdots,m_n=0}^{\infty} (\bar{m}_1+\alpha_1)^{m_1} \cdots (\bar{m}_n+\alpha_n)^{m_n} \frac{u_1^{m_1} \cdots u_n^{m_n}}{m_1! \cdots m_n!} = e^{\alpha_1 x_1 + \cdots + \alpha_n x_n} \Delta^{-1}, \tag{3.1}$$

with $\Delta = \Delta(x_1, x_2, \cdots, x_n)$. Differentiation with respect to the α_i gives

$$\sum_{m_i=r_i}^{\infty} (\bar{m}_1+\alpha_1)^{m_1-r_1} \cdots (\bar{m}_n+\alpha_n)^{m_n-r_n} \frac{u_1^{m_1} \cdots u_n^{m_n}}{(m_1-r_1)! \cdots (m_n-r_n)!} = x_1^{r_1} \cdots x_n^{r_n} e^{\alpha_1 x_1 + \cdots + \alpha_n x_n} \Delta^{-1} \qquad (r_1, r_2, \cdots, r_n \geqq 0), \tag{3.2}$$

where

$$\sum_{m_i=r_i}^{\infty} \equiv \sum_{m_1=r_1}^{\infty} \cdots \sum_{m_n=r_n}^{\infty}.$$

Consider the operator

$$\Omega = \det\left(\delta_{ij} - a_{ij} \frac{\partial}{\partial_i}\right) \qquad (i, j = 1, 2, \cdots, n) \tag{3.3}$$

which is obtained from $\Delta(x_1, x_2, \cdots, x_n)$ by replacing x_i by $\partial/\partial\alpha_i$. Also put

$$\Delta(x_1, x_2, \cdots, x_n) = \sum A(r_1, \cdots, r_n) x_1^{r_1} \cdots x_n^{r_n},$$

where on the right each $r_i = 0$ or 1. Thus, by (1.3) and (3.3),

$$\Omega = \sum A(r_1, \cdots, r_n) \frac{\partial^{r_1+\cdots+r_n}}{\partial\alpha_1^{r_1} \cdots \partial\alpha_n^{r_n}}.$$

Thus (3.2) gives

$$e^{\alpha_1 x_1+\cdots+\alpha_n x_n}=\Omega \sum_{m_1,\cdots,m_n=0}^{\infty}(\bar{m}_1+\alpha_1)^{m_1}\cdots(\bar{m}_n+\alpha_n)^{m_n}\frac{u_1^{m_1}\cdots u_n^{m_n}}{m_1!\cdots m_n!}$$

$$=\sum_{m_1,\cdots,m_n=0}^{\infty}C(m_1,\cdots,m_n)$$

$$\cdot(\bar{m}_1+\alpha_1)^{m_1}\cdots(\bar{m}_n+\alpha_n)^{m_n}\frac{u_1^{m_1}\cdots u_n^{m_n}}{m_1!\cdots m_n!},$$

where

$$C(m_1,\cdots,m_n)=\sum A(r_1,\cdots,r_n)\left(\frac{m_1}{\bar{m}_1+\alpha_1}\right)^{r_1}\cdots\left(\frac{m_n}{\bar{m}_n+\alpha_n}\right)^{r_n}$$

$$=\det\left(\delta_{ij}-\frac{m_i}{\bar{m}_i+\alpha_i}a_{ij}\right)$$

$$=\prod_{i=1}^{n}(\bar{m}_i+\alpha_i)^{-1}\cdot\det\left((\bar{m}_i+\alpha_i)\delta_{ij}-m_i a_{ij}\right).$$

If we put

$$D(m_1,\cdots,m_n)=\det\left((\bar{m}_i+\alpha_i)\delta_{ij}-m_i a_{ij}\right), \tag{3.4}$$

it therefore follows that

$$e^{\alpha_1 x_1+\cdots+\alpha_n x_n}=\sum_{m_1,\cdots,m_n=0}^{\infty}D(m_1,\cdots,m_n)(\bar{m}_1+\alpha_1)^{m_1-1}\cdots(\bar{m}_n+\alpha_n)^{m_n-1}$$

$$\cdot\frac{u_1^{m_1}\cdots u_n^{m_n}}{m_1!\cdots m_n!}. \tag{3.5}$$

We now differentiate (3.5) with respect to α_i and then put $\alpha_1=\cdots=\alpha_n=0$. Since

$$D(m_1,\cdots,m_n)=0 \qquad (m_1,\cdots,m_n\geqq 0)$$

when $\alpha_1=\cdots=\alpha_n=0$, it accordingly follows that

$$x_i=\sum_{m_1,\cdots,m_n=0}^{\infty}C_i(m_1,\cdots,m_n)(\bar{m}_1+\alpha_1)^{m_1-1}\cdots(\bar{m}_n+\alpha_n)^{m_n-1}$$

$$\cdot\frac{u_1^{m_1}\cdots u_n^{m_n}}{m_1!\cdots m_n!} \qquad (i=1,2,\cdots,n), \tag{3.6}$$

where $C_i(m_1,\cdots,m_n)$ denotes the cofactor of the element in the (i,i) position of the determinant

$$\det\left(\bar{m}_j\delta_{jk}-m_j a_{jk}\right) \qquad (j,k=1,2,\cdots,n).$$

It is easily verified that, when $n=2$, (3.5) reduces to (2.5), while (3.6) reduces to (2.8) and (2.9).

We may state

THEOREM 2. *Given*:

$$u_i = x_i \exp\left\{-\sum_{j=1}^{n} a_{ij}x_j\right\} \qquad (i = 1, 2, \cdots, n), \tag{3.7}$$

where

$$A = (a_{ij}) \qquad (i, j = 1, 2, \cdots, n)$$

is a square array of real or complex numbers. Then (3.5) *and* (3.6) *hold*; *equation* (3.6) *furnishes the inverse of* (3.7).

Combining (3.1) and (3.5) we obtain the following theorem.

THEOREM 3. *We have*

$$\begin{aligned}&\sum_{r_i=0}^{m_i} D(r_1, \cdots, r_n)(\bar{r}_1+\alpha_1)^{r_1-1} \cdots (\bar{r}_n+\alpha_n)^{r_n-1}\\ &\quad\cdot(\bar{m}_1-\bar{r}_1+\beta_1)^{m_1-r_1} \cdots (\bar{m}_n-\bar{r}_n+\beta_n)^{m_n-r_n}\\ &\quad=(\bar{m}_1+\alpha_1+\beta_1)^{m_1} \cdots (\bar{m}_n+\alpha_n+\beta_n)^{m_n} \qquad (m_1, \cdots, m_n \geq 0),\end{aligned} \tag{3.8}$$

$$\begin{aligned}&\sum_{r_i=0}^{m_i} D_\alpha(r_1, \cdots, r_n)(\bar{r}_1+\alpha_1)^{r_1-1} \cdots (\bar{r}_n+\alpha_n)^{r_n-1}\\ &\quad\cdot D_\beta(m_1-r_1, \cdots, m_n-r_n)\\ &\quad\cdot(\bar{m}_1-\bar{r}_1+\beta_1)^{m_1-r_1-1} \cdots (\bar{m}_n-\bar{r}_n+\beta_n)^{m_n-r_n-1}\\ &\quad=D_{\alpha+\beta}(m_1, \cdots, m_n)(\bar{m}_1+\alpha_1+\beta_1)^{m_1-1} \cdots (\bar{m}_n+\alpha_n+\beta_n)^{m_n-1}\\ &\qquad\qquad (m_1, \cdots, m_n \geqq 0),\end{aligned} \tag{3.9}$$

$$\bar{m}_j = \sum_{i=1}^{n} m_i a_{ij}, \qquad \bar{r}_j = \sum_{i=1}^{n} r_i a_{ij},$$

$$D_\alpha(m_1, \cdots, m_n) = \det\left((\bar{m}_i+\alpha_i)\delta_{ij} - m_i a_{ij}\right) \qquad (i, j = 1, 2, \cdots, n)$$

and

$$\sum_{r_i=0}^{m_i} \equiv \sum_{r_1=0}^{m_1} \cdots \sum_{r_n=0}^{m_n}.$$

4. Factorial analog of (1.4). It is convenient to first treat the case $n = 2$. We consider the sum

$$\begin{aligned}S = \sum_{m,n=0}^{\infty} &(am+cn+\alpha)_m(bm+dn+\beta)_n \frac{x^m y^n}{m!n!}\\ &\cdot(1+x)^{-am-cn-m-\alpha}(1+y)^{-bm-dn-n-\beta}.\end{aligned} \tag{4.1}$$

Thus

$$S=\sum_{m,n=0}^{\infty}(am+cn+\alpha)_m(bm+dn+\beta)_n\frac{x^m y^n}{m!n!}$$

$$\cdot\sum_{j,k=0}^{\infty}(-1)^{j+k}\frac{(am+cn+m+\alpha)_j(bm+dn+n+\beta)_k}{j!k!}x^j y^k$$

$$=\sum_{m,n=0}^{\infty}\frac{x^m y^n}{m!n!}\sum_{j=0}^{m}\sum_{k=0}^{n}(-1)^{j+k}\binom{m}{j}\binom{n}{k}(a(m-j)+c(n-k)+\alpha)_m$$

$$\cdot(b(m-j)+d(n-k)+\beta)_n \tag{4.2}$$

$$=\sum_{m,n=0}^{\infty}\frac{x^m y^n}{m!n!}\sum_{j=0}^{m}\sum_{k=0}^{n}(-1)^{m+n-j-k}\binom{m}{j}\binom{n}{j}(aj+ck+\alpha)_m(bj+dk+\beta)_n$$

$$=\sum_{m,n=0}^{\infty}\frac{x^m y^n}{m!n!}S_{m,n},$$

say. Since

$$(a+b)_m=\sum_{r=0}^{m}\binom{m}{r}(a)_r(b)_{m-r},$$

it follows that

$$S_{m,n}=\sum_{j,k}(-1)^{m+n-j-k}\sum_{r=0}^{m}\binom{m}{r}(aj)_r(ck+\alpha)_{m-r}$$

$$\cdot\sum_{s=0}^{n}\binom{n}{s}(bj)_s(dk+\beta)_{n-s}$$

$$=\sum_{r=0}^{m}\sum_{s=0}^{n}\binom{m}{r}\binom{n}{s}\sum_{j=0}^{m}(-1)^{m-j}\binom{m}{j}(aj)_r(bj)_s$$

$$\cdot\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}(ck+\alpha)_{m-r}(dk+\beta)_{n-s}.$$

Since

$$\sum_{j=0}^{m}(-1)^{m-j}\binom{m}{j}(aj)_r(bj)_s=0\qquad(m>r+s)$$

and

$$\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}(ck+\alpha)_{m-r}(dk+\beta)_{n-s}=0\qquad(n>m+n-r-s),$$

it is clear that we need only consider those terms for which $r+s=m$. Moreover, when $r+s=m$, we have

$$\sum_{j=0}^{m}(-1)^{m-j}\binom{m}{j}(aj)_r(bj)_{m-r}=m!a^r b^{m-r},$$

$$\sum_{k=0}^{n}(-1)^{n-k}\binom{n}{j}(ck+\alpha)_s(dk+\beta)_{n-s}=n!c^s d^{n-s}.$$

Hence

$$S_{m,n} = m!n! \sum_{r+s=m} \binom{m}{r}\binom{n}{s} a^r b^{m-r} c^s d^{n-s}$$

$$= m!n! \sum_{s=0}^{\min(m,n)} \binom{m}{s}\binom{n}{s} a^{m-s} b^s c^s d^{n-s}.$$

Substituting in (4.2), we get

$$\sum_{m,n=0}^{\infty} (am+cn+\alpha)_m (bm+dn+\beta)_n \frac{x^m y^n}{m!n!} (1+x)^{-am-cn-m-\alpha} (1+y)^{-bm-dn-n-\beta}$$

$$= \sum_{m,n=0}^{\infty} x^m y^n \sum_{s=0}^{\min(m,n)} \binom{m}{s}\binom{n}{s} a^{m-s} (bc)^s d^{n-s}$$

$$= \sum_{s=0}^{\infty} (bcxy)^s \sum_{m=0}^{\infty} \binom{m+s}{s} (ax)^m \sum_{n=0}^{\infty} \binom{n+s}{s} (dy)^n$$

$$= \sum_{s=0}^{\infty} (bcxy)^s (1-ax)^{-s-1} (1-dy)^{-s-1}$$

$$= (1-ax)^{-1}(1-dy)^{-1}\left\{1 - \frac{bcxy}{(1-ax)(1-dy)}\right\}^{-1}$$

$$= \{1-ax-dy+(ad-bc)xy\}^{-1}.$$

We have therefore proved the identity

$$\sum_{m,n=0}^{\infty} (am+cn+\alpha)_m (bm+dn+\beta)_n \frac{x^m y^n}{m!n!} (1+x)^{-am-cn-m} \cdot (1+y)^{-bm-dn-n} = \frac{1}{\Delta}(1+x)^{\alpha}(1+y)^{\beta}, \tag{4.3}$$

where

$$\Delta = \Delta(x, y) = 1 - ax - dy + (ad-bc)xy. \tag{4.4}$$

If we replace a, b, c, d, α, β, x, y by their negatives, (3.4) becomes

$$\sum_{m,n=0}^{\infty} \binom{am+cn+\alpha}{m}\binom{bm+dn+\beta}{n} x^m y^n (1-x)^{am+cn-m} (1-y)^{bm+dn-n} = \frac{1}{\Delta}(1-x)^{-\alpha}(1-y)^{-\beta}. \tag{4.5}$$

Turning now to the general case, we take

$$S \equiv \sum_{m_j=0}^{\infty} \prod_{j=1}^{n} (\bar{m}_j+\alpha_j)_{m_j} \frac{x_j^{m_j}}{m_j!}(1+x_j)^{-\bar{m}_j-m_j-\alpha_j} \tag{4.6}$$

$$= \sum_{m_j=0}^{\infty} \prod_{j=1}^{n} (\bar{m}_j+\alpha_j)_{m_j} \frac{x_j^{m_j}}{m_j!} \sum_{k_j=0}^{\infty} (-1)^{k_j} \frac{(\bar{m}_j+m_j+\alpha_j)_{k_j} x_j^{k_j}}{k_j!}$$

$$= \sum_{m_j=0}^{\infty} \prod_{j=1}^{n} \frac{x_j^{m_j}}{m_j!} \sum_{k_j=0}^{m_j} (-1)^{m_j-k_j} \binom{m_j}{k_j} (\bar{k}_j+\alpha_j)_{m_j}$$

$$= \sum_{m_i=0}^{\infty} \frac{x_1^{m_1} \cdots x_n^{m_n}}{m_1! \cdots m_n!} S_{m_1, \cdots, m_n}, \tag{4.7}$$

say. Here of course

$$\bar{m}_j = \sum_{i=1}^{n} m_i a_{ij}, \qquad \bar{k}_j = \sum_{i=1}^{n} k_i a_{ij} \qquad (j = 1, 2, \cdots, n).$$

Now

$$(\bar{k}_j+\alpha_i)_{m_j} = \sum_{\sum_i m_{ij}=m_j} (m_{1j}, \cdots, m_{nj})(k_1 a_{1j})_{m_{1j}} \cdots (k_{n-1}a_{n-1,j})_{m_{n-1,j}} \cdot (k_n a_{nj}+\alpha_j)_{m_{nj}}, \tag{4.8}$$

where

$$(r_1, r_2, \cdots, r_n) = \frac{(r_1+r_2+\cdots+r_n)!}{r_1! r_2! \cdots r_n!}.$$

It follows that

$$\prod_{j=1}^{n} \sum_{k_j=0}^{m_j} (-1)^{m_j-k_j} \binom{m_j}{k_j} (\bar{k}_j+\alpha_j)_{m_{ij}} = 0$$

unless

$$\sum_{j=1}^{n} m_{ij} \geqq m_i \qquad (i = 1, 2, \cdots, n).$$

Since, by (4.8)

$$\sum_{i=1}^{n} m_{ij} = m_j \qquad (j = 1, 2, \cdots, n),$$

it is clear that

$$\sum_{j=1}^{n} m_{1j} = m_i \qquad (i = 1, 2, \cdots, n).$$

The remainder of the proof is like that of § 3 in [1] and will be omitted. Thus we finally get

THEOREM 4. *We have*

$$\sum_{m_i=0}^{\infty} (\bar{m}_1+\alpha_1)_{m_1}\cdots(\bar{m}_n+\alpha_n)_{m_n}\frac{u_1^{m_1}\cdots u_n^{m_n}}{m_1!\cdots m_n!}$$

$$=\frac{1}{\Delta(x_1,\cdots,x_n)}(1+x_1)^{\alpha_1}\cdots(1+x_n)^{\alpha_n}, \tag{4.9}$$

where

$$\Delta(x_1,\cdots,x_n)=\det(\delta_{ij}-x_ia_{ij})$$

and

$$u_i=x_i\prod_{j=1}^{n}(1+x_j)^{-a_{ij}-\delta_{ij}} \qquad (i=1,2,\cdots,n). \tag{4.10}$$

Clearly this result reduces to (4.3) when $n=2$.

5. Factorial analogs ($n=2$). We shall first obtain analogs of the results of § 2. To do this we apply the difference operators

$$\Delta_\alpha f(\alpha)=f(\alpha+1)-f(\alpha), \qquad \Delta_\beta f(\beta)=f(\beta+1)-f(\beta)$$

and generally $\Delta_\alpha^r\Delta_\beta^s=\Delta_\beta^s\Delta_\alpha^r$. Since

$$\Delta_\alpha(x+\alpha)_m=m(x+\alpha+1)_{m-1},$$

$$\Delta_\alpha^r(x+\alpha)_m=\frac{m!}{(m-r)!}(x+\alpha+r)_{m-r}$$

and

$$\Delta_\alpha(1+x)^\alpha=x(1+x)^\alpha, \qquad \Delta_\alpha^r(1+x)^\alpha=x^r(1+x)^\alpha,$$

it follows from (4.3) that

$$\frac{x^ry^s}{\Delta}(1+x)^\alpha(1+y)^\beta=\sum_{m=r}^{\infty}\sum_{n=s}^{\infty}\frac{u^mv^n}{(m-r)!(n-s)!}(am+cn+\alpha+r)_{m-r}$$

$$\cdot(bm+dn+\beta s)_{n-s}, \tag{5.1}$$

where

$$u=x(1+x)^{-a-1}(1+y)^{-b}$$

$$v=y(1+x)^{-c}(1+y)^{-d-1}. \tag{5.2}$$

Exactly as in the proof of (2.4), (5.1) yields

$$x^ry^s(1+x)^\alpha(1+y)^\beta\cdot$$

$$=\sum_{m=r}^{\infty}\sum_{n=s}^{\infty}\{(ar+cn+\alpha+r)(bm+ds+\beta+s)-bc(m-r)(n-s)\} \tag{5.3}$$

$$\cdot(am+cn+\alpha+r+1)_{m-r-1}(bm+dn+\beta+s+1)_{n-s-1}$$

$$\cdot\frac{u^mv^n}{(m-r)!(n-s)!},$$

where it is understood that

$$(x+1)_{-1}=\frac{1}{x}.$$

In particular, for $(r, s)=(0, 0), (1, 0), (0, 1)$, we get

$$(1+x)^{\alpha}(1+y)^{\beta} = \sum_{m,n=0}^{\infty} (bm\alpha+cn\beta+\alpha\beta) \cdot(am+cn+\alpha+1)_{m-1}(bm+dn+\beta+1)_{n-1}\frac{u^m v^n}{m!n!}, \tag{5.4}$$

$$x(1+x)^{\alpha}(1+y)^{\beta} = \sum_{m=1}^{\infty}\sum_{n=0}^{\infty} \{(a+\alpha+1)(bm+\beta)+(b+\beta)cn\} \cdot m(am+cn+\alpha+2)_{m-2}(bm+dn+\beta+1)_{n-1}\frac{u^m v^n}{m!n!}, \tag{5.5}$$

$$y(1+x)^{\alpha}(1+y)^{\beta} = \sum_{m=0}^{\infty}\sum_{n=1}^{\infty} \{(d+\beta+1)(cn+\alpha)+(c+\alpha)bm\} \cdot n(am+cn+\alpha+1)_{m-1}(bm+dn+\beta+2)_{n-2}\frac{u^m v^n}{m!n!}. \tag{5.6}$$

Specializing further, we take $\alpha=\beta=0$ in (5.5) and (5.6) and we get

$$\begin{aligned} x &= \sum_{m=1}^{\infty}\sum_{n=0}^{\infty} bm(am+cn+2)_{m-1}(bm+dn+1)_{n-1}\frac{u^m v^n}{m!n!}, \\ y &= \sum_{m=0}^{\infty}\sum_{n=1}^{\infty} cn(am+cn+1)_{m-1}(bm+dn+2)_{n-1}\frac{u^m v^n}{m!n!}. \end{aligned} \tag{5.7}$$

Changing the notation slightly we have

THEOREM 5. *Put*

$$\begin{aligned} u &= x(1+x)^{-a}(1+y)^{-b} \\ v &= y(1+x)^{-c}(1+y)^{-d}. \end{aligned} \tag{5.8}$$

Then

$$\begin{aligned} x &= \sum_{m=1}^{\infty}\sum_{n=0}^{\infty} \frac{b}{bm+dn}\binom{am+cn}{m-1}\binom{bm+dn}{n}u^m v^n, \\ y &= \sum_{m=0}^{\infty}\sum_{n=1}^{\infty} \frac{c}{am+cn}\binom{am+cn}{m}\binom{bm+dn}{n-1}u^m v^n. \end{aligned} \tag{5.9}$$

Note that when $b=c=0$, (5.8) becomes

$$u=x(1+x)^{-a}, \qquad v=y(1+y)^{-d},$$

while (5.9) reduces to

$$x=\sum_{m=1}^{\infty}\frac{1}{m}\binom{am}{m-1}u^m, \qquad y=\sum_{n=1}^{\infty}\frac{1}{n}\binom{dn}{n-1}v^n,$$

in agreement with [7, p. 125, no. 211].

Substituting from (5.8) in (5.9) we get the following pair of identities.

$$\sum_{r=1}^{m}\sum_{s=0}^{n}(-1)^{m+n-r-s}\frac{b}{br+ds}\binom{ar+cs}{r-1}\binom{br+ds}{s} \cdot\binom{ar+m-r-1}{m-r}\binom{bs+n-s-1}{n-s}=\delta_{m,1}\delta_{n,0}, \tag{5.10}$$

$$\sum_{r=0}^{m}\sum_{s=1}^{n}(-1)^{m+n-r-s}\frac{c}{ar+cs}\binom{ar+cs}{r}\binom{br+ds}{s-1} \cdot\binom{ar+m-r-1}{m-r}\binom{bs+n-s-1}{n-s}=\delta_{m,0}\delta_{n,1}. \tag{5.11}$$

In the next place rewrite (4.3) and (5.4) in the form

$$\sum_{m,n=0}^{\infty}\binom{am+cn+\alpha-1}{m}\binom{bm+dn+\beta-1}{n}u^m v^n=\frac{1}{\Delta}(1+x)^{\alpha}(1+y)^{\beta}, \tag{5.12}$$

$$\sum_{m,n=0}^{\infty}A_{m,n}(\alpha,\beta)u^m v^n=(1+x)^{\alpha}(1+y)^{\beta}, \tag{5.13}$$

where now

$$A_{m,n}(\alpha,\beta)=\frac{bm\alpha+cn\beta+\alpha\beta}{(am+cn+\alpha)(bm+dn+\beta)}\binom{am+cn+\alpha}{m}\binom{bm+dn+\beta}{n},$$

$$\Delta=(1-(a-1)x)(1-(d-1)y)-bcxy$$

and u, v are given by (5.8). The formulas (5.12), (5.13) evidently imply the following theorem.

THEOREM 6. *We have, for arbitrary* α, α', β, β',

$$\sum_{r=0}^{m}\sum_{s=0}^{n}A_{r,s}(\alpha,\beta)A_{m-r,n-s}(\alpha',\beta')=A_{m,n}(\alpha+\alpha',\beta+\beta'), \tag{5.14}$$

$$\sum_{r=0}^{m}\sum_{s=0}^{n}\binom{ar+cs+\alpha-1}{r}\binom{br+ds+\beta-1}{s}A_{m-r,n-s}(\alpha',\beta') =\binom{am-cn+\alpha+\alpha'-1}{m}\binom{bm+dn+\beta+\beta'-1}{n}. \tag{5.15}$$

6. Factorial analogs (n arbitrary). Applying the operator $\Delta_{\alpha_1}^{r_1}\cdots\Delta_{\beta_n}^{r_n}$ to (4.9) we get

$$\frac{1}{\Delta(x_1,\cdots,x_n)}x_1^{r_1}\cdots x_n^{r_n}(1+x_1)^{\alpha_1}\cdots(1+x_n)^{\alpha_n} \tag{6.1}$$
$$=\sum_{m_i=r_i}^{\infty}\prod_{j=1}^{n}\frac{u_j^{m_j}}{(m_j-r_j)!}(\bar{m}_j+r_j+\alpha_j)_{m_j-r_j},$$

where the notation is that of Theorem 4.

Define the operator

$$\Omega=\det(\delta_{ij}-a_{ij}\Delta_{\alpha_i})\qquad(i,j=1,2,\cdots,n), \tag{6.2}$$

obtained from $\Delta(x_1,\cdots,x_n)$ by replacing x_i by Δ_{α_i}. Then exactly as in the proof of (3.5),

$$(1+x_1)^{\alpha_1}\cdots(1+x_n)^{\alpha_n}=\Omega\sum_{m_i}\prod_{j=1}^{n}\frac{u_j^{m_j}}{2m_j!}(\bar{m}_j+\alpha_j)_{m_j}$$
$$=\sum_{m_i}\sum_{r_i}A(r_1,\cdots,r_n)\Delta_{\alpha_1}^{r_1}\cdots\Delta_{\alpha_n}^{r_n}\prod_{j=1}^{n}\frac{u_j^{m_j}}{m_j!}(\bar{m}_j+\alpha_j)_{m_j},$$

where each $r_i=0$ or 1. The inner sum is equal to

$$\sum_{r_i}A(r_1,\cdots,r_n)\prod_{j=1}^{n}\frac{u_j^{m_j}}{(m_j-r_j)!}(\bar{m}_j+r_j+\alpha_j)_{m_j-r_j}$$
$$=\sum_{r_i}A(r_1,\cdots,r_n)\prod_{i=1}^{n}\frac{m_i}{m_i+_i}\cdot\prod_{j=1}^{n}\binom{\bar{m}_j+m_j+\alpha_j-1}{m_j}u_i^{m_j}$$
$$=D_\alpha(m_1,\cdots,m_n)\prod_{i=1}^{n}(\bar{m}_i+\alpha_i)^{-1}\cdot\prod_{j=1}^{n}\binom{m_j+m_j+\alpha_j-1}{m_j}u_j^{m_j}$$
$$=D_\alpha(m_1,\cdots,m_n)\prod_{i=1}^{n}(m_i+m_i+\alpha_i)^{-1}\cdot\prod_{j=1}^{n}\binom{m_j+m_j+\alpha_j}{m_j}u_j^{m_j},$$

where

$$D_\alpha(m_1,\cdots,m_n)=\det((\bar{m}_i+\alpha_i)\delta_{ij}-m_ia_{ij}). \tag{6.3}$$

Thus we have

$$(1+x_1)^{\alpha_1}\cdots(1+x_n)^{\alpha_n} \tag{6.4}$$
$$=\sum_{m_i=0}^{\infty}D_\alpha(m_1,\cdots,m_n)\prod_{i=1}^{n}(m_i+m_i+\alpha_i)^{-1}\binom{\bar{m}_i+m_i+\alpha_i}{m_i}u_i^{m_i}.$$

This result can be simplified slightly by replacing the array $A=(a_{ij})$ by

$$A-I=(a_{ij}-\delta_{ij}).$$

Then (6.4) becomes

$$(1+x_1)^{\alpha_1}\cdots(1+x_n)^{\alpha_n} = \sum_{m_i=0}^{\infty} D_\alpha(m_1,\cdots,m_n)\prod_{j=1}^{n}(\bar{m}_j+\alpha_j)^{-1}\binom{\bar{m}_j+\alpha_j}{m_j}u_j^{m_j}. \tag{6.5}$$

Note that the coefficient $D_\alpha(m_1,\cdots,m_n)$ does not change.

Applying the operator Δ_{α_i} to (6.5) and then taking $\alpha_1=\cdots=\alpha_n=0$, we get

$$x_i = \sum_{m_j=0}^{\infty} C_i(m_1,\cdots,m_n)\prod_{j=1}^{n}(\bar{m}_j+\alpha_j)^{-1}\binom{\bar{m}_j+\alpha_j}{m_j}u_j^{m_j} \qquad (i=1,2,\cdots,n), \tag{6.6}$$

where, as in (3.6), $C_i(m_1,\cdots,m_n)$ denotes the cofactor of the element in the (i,i) position of

$$\det(\bar{m}_j\delta_{jk}-m_ja_{jk}) \qquad (j,k=1,2,\cdots,n).$$

We may state

THEOREM 7. *Given*

$$u_i = x_i\prod_{j=1}^{n}(1+x_j)^{-a_{ij}} \qquad (i=1,2,\cdots,n), \tag{6.7}$$

then the x_i are determined by (6.6).

Finally, to state convolution theorems, we make use of (4.9) and (6.5). It is convenient to rewrite (4.9) in the form

$$\sum_{m_i=0}^{\infty}\prod_{i=1}^{n}\binom{\bar{m}_i+\alpha_1-1}{m_i}u_i^{m_i} = \frac{1}{\Delta_0(x_1,\cdots,x_n)}(1+x_1)^{\alpha_1}\cdots(1+x_n)^{\alpha_n}, \tag{6.8}$$

where the u_i are now defined by (6.7) and

$$\Delta_0(x_1,\cdots,x_n)=\det((1+x_i)\delta_{ij}-x_ia_{ij}).$$

Also for brevity we put

$$A_{m_1,\cdots,m_n}(\alpha_1,\cdots,\alpha_n) = D_\alpha(m_1,\cdots,m_n)\prod_{j=1}^{m}(\bar{m}_j+\alpha_j)^{-1}\binom{\bar{m}_j+\alpha_j}{m_j},$$

$$B_{m_1,\cdots,m_n}(\alpha_1,\cdots,\alpha_n)=\prod_{j=1}^{n}\binom{\bar{m}_j+\alpha_j-1}{m_j}.$$

THEOREM 8. *We have, for arbitrary α_i, β_i,*

$$\sum_{r_i=0}^{m_i} A_{r_1,\cdots,r_n}(\alpha_1,\cdots,\alpha_n)A_{m_1-r_1,\cdots,m_n-r_n}(\beta_1,\cdots,\beta_n) = A_{m_1,\cdots mm_n}(\alpha_1+\beta_1,\cdots,\alpha_n+\beta_n), \tag{6.9}$$

$$\sum_{r_i=0}^{m_i} A_{r_1,\cdots,r_n}(\alpha_1,\cdots,\alpha_n)B_{m_1-r_1,\cdots,m_n-r_n}(\beta_1,\cdots,\beta_n)$$

$$= B_{m_1,\cdots,m_n}(\alpha_1+\beta_1,\cdots,\alpha_n+\beta_n). \tag{6.10}$$

It is easily verified that, for $n=2$, (6.9), (6.10) reduce to (5.14), (5.15), respectively.

REFERENCES

[1] L. CARLITZ, *An application of MacMahon's Master Theorem*, SIAM J. Appl. Math., 26 (1974), pp. 431–436.
[2] I. J. GOOD, *Generalizations to several variables of Lagrange's expansion with applications to stochastic processes*, Proc. Cambridge Philos. Soc., 56 (1960), pp. 367–379.
[3] ———, *A short proof of MacMahon's 'Master Theorem'*, Ibid., 58 (1962), p. 160.
[4] H. W. GOULD, *Some generalizations of Vandermonde's convolution*, Amer. Math. Monthly, 63 (1956), pp. 84–91.
[5] ———, *Final analysis of Vandermonde's convolution*, Ibid., 64 (1957), pp. 409–415.
[6] P. A. MACMAHON, *Combinatory Analysis I*, University Press, Cambridge, 1915.
[7] G. PÓLYA AND G. SZEGÖ, *Aufgaben und Lehrsätze aus der Analysis I*, Springer, Berlin, 1925.
[8] J. RIORDAN, *Combinatorial Identities*, John Wiley, New York, 1968.

SIAM J. MATH. ANAL.
Vol. 8, No. 2, April 1977

LINEAR CONTINUOUS OPERATORS IN L^p AND GENERALIZED RANDOM PROCESSES: A KERNEL REPRESENTATION*

REUVEN MEIDAN†

Abstract. A generalized impulse response representation is developed for linear and continuous operators from $\mathcal{D}$, the space of infinitely differentiable functions of compact support, into E, the space of complex-valued functions on a set Ω, equipped with the pointwise topology. This representation is employed in order to develop a similar representation for continuous linear operators from $\mathcal{D}$ into $L^p(\Omega)$, where Ω is now a measure space. These results are then applied in order to obtain a kernel representation for continuous linear operators from $L^p(R^n)$ into $L^q(R^m)$ and a representation for generalized random processes.

1. Introduction. The purpose of this work is to develop a kernel representation for linear and continuous operators in L^p, $1 \leqq p \leqq \infty$, the spaces of functions on R^n whose moduli to the power p are integrable. As is well known, these operators do not enjoy a kernel representation in an ordinary sense. The common counterexample usually employed to illustrate this fact is the identity operator. This operator will require the impulse functional $\delta(t-s)$ as a kernel, and since the latter is not a member of any of the L^p spaces, it cannot serve as an ordinary kernel. However, as will be shown in this work, if we suitably restrict the domain of definition of the operator to a subspace of L^p, a kernel representation does exist. Moreover, the representation admits a scalar product form. The same idea will enable us to provide a similar representation for generalized random processes.

The impulse response representation for linear continuous and time-invariant operators is well known (e.g. Zemanian [10]). Meidan [5] has introduced the concept of the generalized impulse response representation in order to be able to cope with operators which are time-varying. More precisely, it is shown in [5] that if u is a linear and continuous operator from $\mathcal{D}$ into $\mathcal{C}$, it admits an impulse response representation in a generalized sense. $\mathcal{D}$ denotes the space of infinitely differentiable testing functions on R^n with compact support equipped with the usual testing function topology. $\mathcal{C}$ is the space of continuous functions on R^m. By the generalized impulse response representation we mean that there exists a family F_y of distributions in $\mathcal{D}'$, the dual of $\mathcal{D}$, depending on the parameter $y \in R^m$, such that, the operator u is representable by

$$(u\phi)(y) = \langle K_y, \phi \rangle, \tag{1}$$

where ϕ denotes any testing function in the domain of u. Moreover, K_y can be expressed in terms of the response of the transpose operator u^t:

$$K_y = u^t \delta_y, \tag{2}$$

* Received by the editors August 21, 1975, and in revised form January 29, 1976.

† School of Engineering, Tel-Aviv University, Tel-Aviv, Israel. This work was carried out while the author was on visit at the National Research Institute for Mathematical Sciences of the CSIR, Pretoria, South Africa.

where δ_y denotes the shifted impulse functional in $\mathscr{C}'$, the dual of $\mathscr{C}$. Namely,

$$\delta_y(f) = \langle \delta_y, f \rangle = f(y), \qquad f \in \mathscr{C}, \quad y \in R^m. \tag{3}$$

However, in this work an extended version of the generalized impulse response representation is necessary. In § 2 we consider the space E as the range space. Let Ω be a set. E is the space of complex-valued functions on Ω equipped with the pointwise topology. It is shown that the generalized impulse response representation holds for linear and continuous operators from $\mathscr{D}$ into E. In a subsequent work [6] a similar result will be shown to hold even when the linearity of the operator is reliquished. However, this general result is not needed here, and we can dispense with the complications involved in this development.

The theory is valid for a fairly broad class of topological vector space serving as the domain space of the operator. However, the space $\mathscr{D}$ has been chosen for the following reasons.

(i) $\mathscr{D}$ is a rather restrictive space and is equipped with the rather strong testing function topology. On the other hand it is large enough to be dense in the L^p spaces, except for the case where $p = \infty$. The restrictiveness of the space and its strong topology assure a broad class of continuous operators. Its being dense in L^p provides its unique extension onto the whole space L^p, provided, of course, that this extensibility is possible from the viewpoint of continuity. On the other hand, the range space E is fairly broad and in view of its weak topology it does not place severe limitations on the class of permissible operators. In fact, the cases of interest will be obtained by limiting the range space to certain subspaces of E.

(ii) The use of the space $\mathscr{D}$ as the domain space is compatible with the concept of the generalized random process (e.g. Gel'fand [2]) which is involved in this work. By definition, a generalized random process is a linear and continuous operator from a space of testing functions into a space of random variables, where the latter space is defined over a probability space.

Further we introduce a measure on the set Ω. Namely $(\Omega, \mathscr{A}, \mu)$ is a measure space, where Ω is the set as introduced earlier, $\mathscr{A}$ is a σ-algebra of subsets of Ω and μ a σ-finite measure on the measurable space $(\Omega, \mathscr{A})$. We consider the spaces $L^p(\Omega)$ of (equivalence classes of) complex-valued functions on Ω. In accordance with the earlier notation, L^P denotes the case when $\Omega = R^n$ and the measure is the Lebesgue measure.

We consider linear and continuous operators from $\mathscr{D}$ into $L^P(\Omega)$ (§ 3) and show that this class of operators is in fact a subclass of the earlier class of operators, linear and continuous from $\mathscr{D}$ into E. Hence they too admit the generalized impulse response representation. This fact is used in order to provide representations for generalized random processes (§ 4) and continuous linear operators in L^P (§ 5).

In the development it is imperative that the operator is linear. It is still an open question whether the results can be extended for nonlinear operators. Of course, in such a case one would suspect a representation in terms of nonlinear functionals on $\mathscr{D}$ rather than the linear functions considered in this work.

Some basic results of functional analysis are quoted in this work. Although they may be found in any book on functional analysis, we preferred to use a single reference. The quotations in the work are taken from Treves's book [9].

2. Operators from $\mathcal{D}$ into E. Let $\mathcal{D}$ denote the space of infinitely differentiable functions on R^n with compact support equipped with Schwartz's testing function topology. E denotes the space of complex-values functions on a set Ω. We equip E with the pointwise topology, according to which a sequence $\{f_n(\omega)\}$ of functions in E converges if and only if $\{f_n(\omega)\}$ converges for each $\omega \in \Omega$.

The dual of $\mathcal{D}$ is denoted by $\mathcal{D}'$. It is the space of distribution on R^n. E' is the dual space of E. Let Δ denote the space spanned by the family of impulse functionals $\{\delta_\omega \mid \omega \in \Omega\}$, where

$$\delta_\omega(f) = \langle \delta_\omega, f \rangle = f(\omega), \qquad f \in E, \quad \omega \in \Omega. \tag{4}$$

Clearly, every δ_ω is a linear and continuous functional on E. Hence, $\Delta \subset E'$. In fact, it can be easily verified that $\Delta = E'$. Indeed, consider on E the weak topology generated by Δ. It is equal to the initial pointwise topology of E. By a well-known property of weak topologies (e.g. Dunford and Schwartz, [1, p. 421]), this establishes that $\Delta = E'$.

The topologies assigned to the dual spaces are the weak dual topologies. The operator u^t denotes the transpose operator of u. It is defined by

$$\langle u^t F, \phi \rangle = \langle F, u\phi \rangle \tag{5}$$

where $\phi \in \mathcal{D}$ and $F \in E'$. It is a linear and weakly continuous operator mapping E' into $\mathcal{D}'$.

At this point we are ready to introduce the generalized impulse response representation for operators which are linear and continuous from $\mathcal{D}$ into E.

THEOREM 1. *The following statements are equivalent.*

(i) *u is a linear and continuous operator from $\mathcal{D}$ into E.*

(ii) *u^t is a linear and weakly continuous operator from E' into $\mathcal{D}'$. u and u^t are transpose operators of each other. Let us denote*

$$F_\omega(x) = u^t \delta_\omega, \qquad \omega \in \Omega. \tag{6}$$

F_ω is a family of distribution in $\mathcal{D}'$ parametrized by $\omega \in \Omega$. It is the family of responses of u^t to the family of shifted impulses δ_ω.

(iii) *There exists a mapping $\omega \to F_\omega$ from Ω into $\mathcal{D}'$ such that the operator u can be represented by*

$$(u\phi)(\omega) = \langle F_\omega, \phi \rangle, \qquad \phi \in \mathcal{D}. \tag{7}$$

Proof. (i) $\Rightarrow$ (ii) follows from the standard theorem of functional analysis (e.g. Robertson and Robertson [7, p. 38, 39]).

(ii) $\Rightarrow$ (i) Let u^t be a linear and weakly continuous operator from E' into $\mathcal{D}'$. Consider its transpose. It is linear and weakly continuous from $\mathcal{D}$ into E. Since $\mathcal{D}$ is a Mackey space it is continuous for the initial topologies [7, p. 62]. This operator coincides with u.

(i) $\Rightarrow$ (iii) Assume u is continuous from $\mathcal{D}$ into E. Then for every $\omega \in \Omega$, $\phi \to (u\phi)(\omega)$ is a linear and continuous functional on $\mathcal{D}$. Hence there exists a functional, say F_ω, such that (7) holds.

(iii) $\Rightarrow$ (i) Conversely, if (iii) holds, the right-hand side of (7) defines a linear and continuous operator from $\mathcal{D}$ into E. It remains to show that the family F_ω

of statement (iii) is expressible via (6). Indeed, by the definition of the transpose,

$$\langle u^t \delta_\omega, \phi \rangle = \langle \delta_\omega, u\phi \rangle$$

and in view of the definition of δ_ω (equation (4))

$$\langle \delta_\omega, u\phi \rangle = (u\phi)(\omega),$$

which completes the proof.

Discussion. The theorem establishes a full equivalence among the three approaches. u is usually referred to as the physical operator, which represents the actual system. It operates on ordinary functions. u^t, its transpose, is a mathematical concept and operates between the spaces of linear functionals. Finally F_ω is a family of functionals which consists of the response of u^t to the family of shifted impulse functionals. Each one of these three approaches fully characterizes the situation.

3. Operators from $\mathscr{D}$ into $L^p(\Omega)$. In this section a generalized impulse response representation is pursued for linear and continuous operators from $\mathscr{D}$ into $L^p(\Omega)$. $L^p(\Omega)$ denotes the space of complex-valued measurable functions whose moduli to the power p are integrable in the measure space $(\Omega, \mathscr{A}, \mu)$. Ω is a set, $\mathscr{A}$ a σ-algebra of subsets in Ω and μ a σ-finite positive measure on the measurable space $(\Omega, \mathscr{A})$. p is a real number, $1 \leqq p \leqq \infty$. The topology on $L^p(\Omega)$ is the usual norm topology induced by the norm

$$\|f\|_p = \left[\int_\Omega |f|^p \, d\mu \right]^{1/p}, \qquad f \in L^p(\Omega). \tag{8}$$

It should be noted that no more generality is gained by considering the weak topology on $L^p(\Omega)$. More precisely, the family of linear operators which are continuous from $\mathscr{D}$ into $L^p(\Omega)$, when the latter is equipped with the weak topology, is identical to the family of linear operators continuous into $L^p(\Omega)$ for the norm topology. This is a direct consequence of the closed graph theorem.

The difficulty with these operators lies in the fact that on $L^p(\Omega)$ one cannot speak of the pointwise definition of a function. More precisely, in $L^p(\Omega)$ we are considering equivalence classes of functions. We identify functions which are μ-equivalent, i.e. functions which are equal except for a set of measure zero. However, it will be shown that in view of the particular structure of $\mathscr{D}$, the range of u is contained in the intersection of E and $L^p(\Omega)$. Moreover the continuity holds also relative to the pointwise topology on the range. Consequently this class of operators constitutes a subclass of the previous one. It follows that the generalized impulse response can be established for these operators as well.

Let K be a compact set in R^n and let $\mathscr{D}(K)$ denote the subset of $\mathscr{D}$ consisting of all test functions whose supports are contained in K. $\mathscr{D}(K)$ is thus a countably normed space. If u is the given operator which is linear and continuous from $\mathscr{D}$ into $L^p(\Omega)$; then u_K denotes its restriction to $\mathscr{D}(K)$. Now u_K is an operator defined on a countably normed space. Hence it is of finite order. Namely, if $\{\gamma_i | i = 0, 1, 2, \cdots\}$ denotes a sequence of nondecreasing seminorms in $\mathscr{D}(K)$, then there exists a finite i such that u_K is continuous on $\mathscr{D}(K)$ with respect to the seminorm γ_i. But $\mathscr{D}(K)$ is nuclear. Consequently, an integer $j > i$ exists such that the injection I_{ji}

from $\mathcal{D}_j(K)$, the space $\mathcal{D}(K)$ completed with respect to γ_j, into $\mathcal{D}_i(K)$ is nuclear. In other words the operator u_K from $\mathcal{D}(K)$ into $L^p(\Omega)$ is a nuclear operator.

We shall first show that u_K exhibits a generalized impulse response representation. Following the above discussion, u_K is a nuclear operator. Moreover, $L^p(\Omega)$ is separable. Hence (e.g. Treves [9, p. 482]) u admits the representation

$$u_K(\phi) = \sum_{i=1}^{\infty} \lambda_i \langle F_i, \phi \rangle f_i \tag{9}$$

where $\{\lambda_i | i = 1, 2, \cdots\}$ is a sequence of complex numbers which is absolutely convergent, i.e. $\sum_{i=1}^{\infty} |\lambda_i| < \infty$, $\{F_i\}$ is a bounded sequence in $\mathcal{D}(K)'$, the dual of $\mathcal{D}(K)$, and $\{f_i\}$ is a bounded sequence in $L^p(\Omega)$.

The series of the right-hand side of the representation (9) converges in the norm of $L^p(\Omega)$. However, it will now be shown that it also converges pointwise for almost every $\omega \in \Omega$.

LEMMA. *Let $\{f_i(\omega)\}$ be a bounded sequence in $L^p(\Omega)$, and $\{\lambda_i\}$ a sequence of complex numbers such that $\sum_{i=1}^{\infty} |\lambda_i| < \infty$. Then the sequence*

$$\sum_{i=1}^{\infty} |\lambda_i f_i(\omega)| \tag{10}$$

converges pointwise for almost every $\omega \in \Omega$.

Proof. We first rewrite series (10)

$$\sum_i |\lambda_i f_i| = \sum_i |\lambda_i^{1-1/p} (\lambda_i^{1/p} f_i)|. \tag{11}$$

We now apply Hölder's inequality to the right-hand side of (11) in order to obtain

$$\sum_i |\lambda_i f_i| \leqq \left(\sum_i |\lambda_i| \right)^{1-1/p} \left(\sum_i |\lambda_i f_i^p| \right)^{1/p} \tag{12}$$

Since, by hypothesis, $\sum_i |\lambda_i|$ converges, we can conclude that it would be sufficient to prove the pointwise convergence of

$$\sum_{i=1}^{\infty} |\lambda_i f_i(\omega)^p| \tag{13}$$

in order to obtain the pointwise convergence of (10). Consider ($p < \infty$),

$$\sum_{i=1}^{\infty} \int_{\Omega} |\lambda_i f_i(\omega)^P| \, d\mu = \sum_{i=1}^{\infty} |\lambda_i| \int_{\Omega} |f_i(\omega)|^P \, d\mu. \tag{14}$$

In view of the fact that the family $\{f_i\}$ is bounded in $L^P(\Omega)$ and $\sum |\lambda_i| < \infty$, we have that the series (14) converges. Consider the sequence $\{g_n\}$ of partial sums of (13)

$$g_n(\omega) = \sum_{i=1}^{n} |\lambda_i f_i(\omega)|^P. \tag{15}$$

$\{g_n\}$ is a monotone sequence of functions which, by the convergence of the series (14), converges in integral. Hence, there exists a subsequence of $\{g_n\}$ which converges pointwise almost everywhere. But since $\{g_n\}$ is a monotone sequence this subsequence can be taken to be $\{g_n\}$ itself. This completes the proof.

THEOREM 2. *Let u be a linear and continuous operator from $\mathcal{D}$ into $L^p(\Omega)$. Let K be a compact set in R^n and $\mathcal{D}(K)$ the subspace of $\mathcal{D}$ consisting of the test functions whose supports are contained in K. Let u_K denote the restriction of u to $\mathcal{D}(K)$. Then u_K admits the representation.*

$$(u_K\phi)(\omega) = \sum_{i=1}^{\infty} \lambda_i \langle F_i, \phi \rangle f_i(\omega) \tag{16}$$

where $\omega \in \Omega$, $\{\lambda_i\}$ is a sequence of complex numbers which converges absolutely, $\{F_i\}$ is a bounded family in $\mathcal{D}(K)'$ and $\{f_i(\omega)\}$ is a bounded family in $L^p(\Omega)$. The representation (16) *converges in the norm of $L^p(\Omega)$, as well as for the pointwise topology for almost every ω. Moreover, the operator u_K is a continuous operator from $\mathcal{D}$ into $L^p(\Omega) \cap E$.*

Proof. As was established above, the representation (16) and its convergence in the L^p norm follow from the nuclearity of u_K. Since $\{F_i\}$ is a bounded family in $\mathcal{D}(K)'$, the set of numbers $\{\langle F_i\phi \rangle\}$ is bounded for each $\phi \in \mathcal{D}(K)$. Hence the pointwise convergence of the series (16) follows from the lemma.

Now let $\{\phi_n\}$ be a sequence which converges to zero in $\mathcal{D}(K)$. We maintain that the sequence

$$\{\sup_i |\langle F_i, \phi_n \rangle|\} \tag{17}$$

converges to zero with n. Indeed, $\{F_i\}$ is a weakly bounded family in $\mathcal{D}(K)'$. Hence by the Banach–Steinhaus theorem (e.g. Treves [9, p. 349]), $\{F_i\}$ is equicontinuous. But the topology of the locally convex Hausdorff space $\mathcal{D}(K)$ is identical to the topology of uniform convergence on every equicontinuous subset of $\mathcal{D}(K)'$ (e.g. Treves [9, p. 369]). Consequently the sequence (17) converges to zero. This establishes the desired continuity of u_K and completes the proof.

The consequence of Theorem 2 is that u_K is in fact an operator continuous from $\mathcal{D}(K)$ into $E \cap L^p(\Omega)$ for both the pointwise and L^p topologies on the range space. We may therefore invoke Theorem 1 in order to establish a generalized impulse response representation for u_K. Consequently there exists a family $\{F_\omega^K\}$ of distributions in $\mathcal{D}(K)'$ such that for almost every ω

$$(u_K\phi)(\omega) = F_\omega^K(\phi), \qquad \phi \in \mathcal{D}(K). \tag{18}$$

We fix ω. For each compact set K of R^n a family of distributions $\{F_\omega^K\}$ exists. The families $\{F_\omega^K\}$ are compatible in the following sense: If K_1 and K_2 are two compact sets, then the restrictions of $\{F_\omega^{K_1}\}$ and $\{F_\omega^{K_2}\}$ to $K_1 \cap K_2$ coincide. Hence by a well-known theorem of distribution theory (e.g. Treves [9, p. 253]) there exists for almost every ω, a unique distribution F_ω on R^n such that its restriction to $\mathcal{D}(K)$ coincides with F_ω^K. This consideration provides the final result for the operator u, which is summarized in the following theorem.

THEOREM 3. *Let u be a linear and continuous operator from $\mathcal{D}$ into $L^p(\Omega)$. Then there exists a family $\{F_\omega\}$ of distributions in $\mathcal{D}'$, such that for almost every ω we have the following representation for the operator*:

$$(u\phi)(\omega) = \langle F_\omega, \phi \rangle, \qquad \phi \in \mathcal{D}. \tag{19}$$

Moreover,

$$F_\omega = u^t \delta_\omega \tag{20}$$

where u^t is the transpose operator of u. Its domain contains E'.

We shall now pursue a characterization of the family $\{F_\omega\}$. If $\{F_\omega\}$ is a family of distributions representing a linear operator u which is continuous from $\mathcal{D}$ into $L^p(\Omega)$, then the expression

$$I_p(\phi) = \int_\Omega |\langle F_\omega, \phi \rangle|^p \, d\mu, \tag{21}$$

exists for every $\phi \in \mathcal{D}$. It defines a functional on $\mathcal{D}$. It can be easily verified that the functional is continuous. However, I_p is not linear, hence it is not an element of $\mathcal{D}'$. In terms of I_p the following characterization of the operator can be established.

THEOREM 4. *Let u be a linear operator from $\mathcal{D}$ into $L^p(\Omega)$; u is continuous if and only if there exists a family $\{F_\omega | \omega \in \Omega\}$ of distributions in $\mathcal{D}'$ satisfying* (19) *and the integral on the right-hand side of* (21) *exists for every $\phi \in \mathcal{D}$, and defines a continuous functional I_p on $\mathcal{D}$.*

Proof. The proof of the direct statement is obvious. It remains to prove the converse. Assume $\{F_\omega\}$ has the above properties. Then it defines via (19) a linear operator u from $\mathcal{D}$ into E. In view of the existence of (21) for every $\phi \in \mathcal{D}$, u in fact maps $\mathcal{D}$ into $L^p(\Omega)$. By the assumed continuity of the functional I_p of (21) we have that u is continuous from $\mathcal{D}$ into $L^p(\Omega)$. Hence the proof is complete.

4. Generalized random processes. The concept of the generalized random process was introduced by Ito [4] and Gel'fand [2], [3]. In principle, it is defined to be a continuous linear operator from a topological vector space of testing functions into a topological vector space of random variables. The space of testing functions employed in this work is $\mathcal{D}$. However, as mentioned above, this is not mandatory and the theory can be established for other spaces of testing functions as well, provided they are nuclear. The space of random variables can be either the space E or the spaces $L^p(\Omega)$. Of course, in this context, one should identify the measure μ to be a probability measure. The following is the common definition of a generalized random process.

DEFINITION. Let $(\Omega, \mathcal{A}, \mu)$ be a probability space. A *generalized random process* is a linear and continuous operator u from the space of testing functions D into the space of random variables E (or $L^p(\Omega)$).

Our representation provides the interpretation of the generalized random process in term of another mapping, i.e. the mapping $\omega \to F_\omega$ from the probability space Ω into the space of distributions $\mathcal{D}'$. In other words, the generalized random process can be viewed as a random process whose sample "functions" F_ω are distributions. The mapping $\omega \to F_\omega$ is often called in the literature *random (Schwartz) distribution.* It is instructive to note that this mapping can be linearized by means of the concept of the transpose of the generalized random process. All these considerations are summarized in the following theorem.

THEOREM 5. *The following statements are equivalent*:

1. *There exists a generalized random process, namely a continuous linear operator u, from $\mathcal{D}$ into E.*

2. *There exists a mapping* $\omega \to F_\omega$ *from* Ω *into* $\mathcal{D}'$ *such that* u *is represented by* (19).

3. *There exists an operator* u^t, *linear and weakly continuous from* E' *into* $\mathcal{D}'$. u^t *is the transpose of* u *and* F_ω *can be expressed as the family of impulse responses of* u^t (*equation* (20)).

5. Continuous operators in L^p. Let u be a linear and continuous operator from $L^p(R^n)$ into $L^q(R^m)$. Consider the restriction of u to $\mathcal{D}(R^n)$ as a subspace of $L^p(R^n)$. Certainly the restricted operator is continuous from $\mathcal{D}(R^n)$ into $L^q(R^m)$. Since $\mathcal{D}$ is dense in L^p, $1 \leqq p < \infty$, we can uniquely reconstruct the given operator from its restricted version. But the restricted operator enjoys a generalized impulse response representation via Theorem 3. We can also view this representation as a kernel representation. Hence a generalized kernel representation has been obtained for continuous linear operators which do not, in general, possess a kernel representation in the ordinary sense.

In system theory it is fairly common for the operator u to be translation-invariant. We refer to u as translation-invariant if it commutes with the shift operator. A necessary condition for translation invariance is the requirement that $m = n$. In other words, u should operate between $L^p(R^n)$ and $L^q(R^n)$. We denote by W^q_{-m} the Sobolev space of order m based on $L^q(R^n)$. Namely, W^q_{-m} is the space of distributions which are expressed as a finite sum of derivatives (in a distributional sense) up to order m of functions which are members of $L^q(R^n)$.

If u is a translation-invariant operator from $L^p(R^n)$ into $L^q(R^n)$, then it is well known (Schwartz [8]) that the family $\{F_y | y \in R^n\}$ can be expressed by the shift of a single distribution F by means of

$$(u\phi)(y) = \langle F(y-x), \phi(x)\rangle = F * \phi. \tag{22}$$

Equation (22) represents the convolutional representation of the operator u and F is called the convolutional kernel. We shall establish the following characterization of the convolutional kernel of a translation-invariant operator.

THEOREM 6. *Let* u *be a linear continuous and translation-invariant operator from* $L^p(R^n)$ *into* $L^q(R^n)$. *Then the convolutional kernel* F *can be expressed as*

$$F = \frac{\partial^n f_1}{\partial x_1 \cdots \partial x_n} + f_2 \tag{23}$$

where f_1 *and* f_2 *are members of* $L^q(R^n)$ *and the derivatives are interpreted in the distributional sense. Namely,* $F \in W^q_{-m}$.

Proof. Let h be the unit step function in R^n and λ a test function in $\mathcal{D}(R^n)$ which is equal to unity in a neighborhood of the origin. λh is a function in $L^p(R^n)$; hence it is in the domain of u. Consequently

$$u(\lambda h) = F * (\lambda h) \tag{24}$$

is in $L^q(R^n)$. Consider

$$\frac{\partial^n (F * \lambda h)}{\partial x_1 \cdots \partial x_n} \tag{25}$$

By a well-known property of the convolutional operator, the operation commutes with differentiation. Hence,

$$\frac{\partial^n (F * \lambda h)}{\partial x_1 \cdots \partial x_n} = F * \frac{\partial^n \lambda h}{\partial x_1 \cdots \partial x_n}. \tag{26}$$

But

$$\frac{\partial^n \lambda h}{\partial x_1 \cdots \partial x_n} = \lambda\delta + \gamma \tag{27}$$

where γ is a function in $\mathcal{D}$ and δ is the impulse at the origin. The validity of (27) is based on the fact that λ was assumed to be equal to a constant in a neighborhood of the origin. Since $\lambda\delta = \delta$, we have that the right-hand side of (26) is equal to

$$F * (\lambda\delta + \gamma) = F + F * \gamma. \tag{28}$$

But the second term of (28) is an element of $L^q(R^n)$ by the hypotheses related to u. Combining equations (26), (27) and (28) completes the proof.

REFERENCES

[1] N. DUNFORD AND J. T. SCHWARTZ, *Linear Operators*, part I, John Wiley, New York, 1958.

[2] I. M. GEL'FAND, *Generalized random processes*, Dokl. Akad. Nauk SSSR, 100 (1956), pp. 853–856. (In Russian.)

[3] I. M. GEL'FAND AND N. YA. VILENKIN, *Generalized Functions*, vol. 4, Academic Press, New York, 1964.

[4] K. ITO, *Stationary random distributions*, Mem. Coll. Sci. Kyoto Univ. Ser. A, 28 (1954), pp. 209–223.

[5] R. MEIDAN, *Translation varying linear operators*, SIAM J., Appl. Math., 22 (1972), pp. 419–436.

[6] ———, *On nonlinear translation varying operators*, J. Math. Anal. Appl. to appear.

[7] A. P. ROBERTSON AND W. ROBERTSON, *Topological Vector Spaces*, Cambridge University Press, London, 1964.

[8] L. SCHWARTZ, *Theories des Distributions*, vol. II, Hermann, Paris, 1959.

[9] F. TREVES, *Topological Vector Spaces, Distributions and Kernels*, Academic Press, New York, 1967.

[10] A. H. ZEMANIAN, *An N-port realizability theory based on the theory of distribution*, IEEE Trans. Circuit Theory, CT-10 (1963), pp. 265–277.

SIAM J. MATH. ANAL.
Vol. 8, No. 2, April 1977

ON A NONLINEAR VOLTERRA INTEGRAL EQUATION ON A HILBERT SPACE*

VIOREL BARBU†

Abstract. Nonlinear Volterra integral equations with singular kernels are considered. The existence and the asymptotic behavior of solutions is studied in a special case.

1. Introduction. This paper concerns integral equations of the form

$$x(t)+\int_0^t E(t-s)g(x(s))\,ds \ni f(t), \qquad \text{a.e. } t>0, \tag{1.1}$$

on a Hilbert space H. $E(t)$: $]0, \infty[\to L(H, H)$ is a positive definite kernel (see [16]) while g is a maximal monotone (multivalued) operator on H.

The special case, $H = R^1$ of this equation has been studied by many authors and more recently by J. J. Levin [9] (see [6] and [15] for significant results and references on this subject).

For comparison with other literature on Volterra integral equation on Hilbert spaces, papers by R. C. MacCamy [13], S. O. Londen [10] and J. S. W. Wong [16] may be cited. However, for the most part, these authors focus on asymptotic behavior of solutions, which involve boundedness of g, thereby excluding many interesting g's which are partial differential operators.

In a recent paper [2] (on these lines see also [1]) the author has obtained some results concerning the existence and asymptotic behavior of solutions in the case in which $E(t)$ is a monotone and convex scalar continuous kernel and g is a cyclically maximal monotone (unbounded) operator in H. Recently S. O. Londen [11] has extended these results to the case in which the kernel is not necessarily decreasingly convex.

Our purpose here is to generalize these results to a class of operatorial singular kernels of the form $E(t) = a(t)S(t)$ where $a(t)$ is a real-valued positive definite function and $S(t)$ is a continuous semigroup of linear bounded operators on H. The original motivation for this case came from a class of partial differential equations arising in the study of heat conduction and is outlined in the last section of the paper.

2. The main results. Throughout this paper the symbol H will denote a real Hilbert space with norm $|\cdot|$ and inner product $(\cdot, \cdot)$. We first review some definitions and basic results concerning maximal monotone operators and convex functions in Hilbert spaces. For other results in this field relevant to the present paper, we refer the reader to the books [3] and [4].

Let g be a nonlinear multivalued (possibly) operator from H into itself. We shall use the following notations:

$$D(g) = \{u \in H;\ g(u) \neq \emptyset\}, \qquad R(g) = \cup\{g(u);\ u \in D(g)\}.$$

* Received by the editors August 21, 1975, and in revised form December 11, 1975.
† Faculty of Mathematics, University of Iasi, Iasi, Rumania.

The operator g is said to be monotone in $H\times H$ if

$$(y_1-y_2, x_1-x_2)\geqq 0 \quad \text{for all } x_i\in D(g),\quad y_i\in g(x_i),\quad i=1,2. \tag{2.1}$$

The monotone operator g is said to be strictly monotone if inequality (2.1) holds with equality if and only if $x_1=x_2$ and $y_1=y_2$. A monotone operator g is said to be maximal monotone if it admits no proper monotone extensions in $H\times H$.

Let g be a maximal monotone operator in $H\times H$. Then the range of $I+\lambda g$ is all of H and for every $\lambda>0$ the operator $(I+\lambda g)^{-1}$ is well-defined and nonexpansive on H.

An important class of maximal monotone operators in $H\times H$ is the subdifferentials of lower semicontinuous convex functions defined on H. Let φ: $H\to]-\infty, +\infty]$ be a lower semicontinuous convex function, nonidentically $+\infty$ on H. Let

$$D(\varphi)=\{u\in H;\ \varphi(u)<+\infty\}$$

$$\partial\varphi(u)=\{y\in H;\ \varphi(u)\leqq\varphi(v)+(y, u-v) \text{ for all } v\in H\}. \tag{2.2}$$

The multivalued operator $u\to\partial\varphi(u)$ is called the subdifferential of φ. If φ happens to be Gâteaux differentiable at u, then $\partial\varphi(u)$ is reduced to a single point which is just the Gâteaux differential of φ at u.

For every $\lambda>0$ let $\varphi_\lambda: H\to]-\infty, +\infty[$ be the convex function defined by

$$\varphi_\lambda(u)=\inf\left\{\frac{|u-v|^2}{2\lambda}+\varphi(v);\ v\in H\right\}. \qquad u\in H. \tag{2.3}$$

The function φ_λ is Fréchet differentiable on H and (see [4], [5])

$$\partial\varphi_\lambda=\lambda^{-1}(I-(I+\lambda\,\partial\varphi)^{-1}). \tag{2.4}$$

Furthermore, $\varphi_\lambda(u)\leqq\varphi(u)$ for all $\lambda>0$, $u\in H$ and

$$\lim_{\lambda\to\infty}\varphi_\lambda(u)=\varphi(u) \quad \text{for all } u\in H. \tag{2.5}$$

We shall denote by $L^2_{\text{loc}}(0,\infty; H)$ the space of all H-valued strongly measurable functions $u:]0,\infty[\to H$ such that $\int_0^T|u(t)|^2\,dt<+\infty$ for every $T>0$. Denote also by $H^1_{\text{loc}}(0,\infty; H)$ the space of all functions $u\in L^2_{\text{loc}}(0,\infty; H)$ which are absolutely continuous on every compact $[0, T]$ and with the first derivative (which exists almost everywhere) in $L^2(0, T; H)$.

A function $u:]0,\infty[\to H$ is called solution of (1,1) if $u\in L^2_{\text{loc}}(0,\infty; H)$ and there is a function $w:]0,\infty[\to H$ such that

$$w\in L^2_{\text{loc}}(0,\infty; H),\qquad w(t)\in g(u(t)) \quad \text{a.e. } t>0, \tag{2.6}$$

$$u(t)+\int_0^t E(t-s)w(s)\,ds=f(t) \quad \text{a.e. } t>0. \tag{2.7}$$

For the sake of simplicity we shall write $g(u)$ instead of w.

It is assumed throughout that $E(t)$ is of the form

$$E(t)=a(t)S(t) \quad \text{for } t>0, \tag{2.8}$$

where $S(t)\colon [0,\infty[\to L(H,H)$ is a continuous semigroup of linear bounded operators on H and $a(t)$ is a real-valued function defined on $R_+^1 =]0,\infty[$. In addition to this, the following conditions on a and S will be assumed:

(i) $a \in C^1(]0,\infty[) \cap L^1(0,1)$.

(ii) $(-1)^k a^{(k)}(t) \geqq 0$ *for all* $t>0$ *and* $k=0,1$.

(iii) $a'(t)$ *is nondecreasing on* $]0,\infty[$.

(iv) *The infinitesimal generator* A *of the semigroup* $S(t)$ *is self-adjoint and dissipative* (i.e., $(Ax,x) \leqq 0$ for all $x \in H$).

It is well known that condition (iv) implies that, for every $t>0$, operator $S(t) \in L(H,H)$ is self-adjoint and contractive, i.e., $\|S(t)\|_{L(H,H)} \leqq 1$. Furthermore, the semigroup $S(t)$ is everywhere differentiable on $]0,\infty[$ and

$$|d/dt\, S(t)x| \leqq |x|/t \quad \text{for all } t>0 \text{ and } x \in H. \tag{2.9}$$

We will impose the following two conditions on g:

(j) $g = \partial\varphi$ *where* $\varphi\colon H \to]-\infty, +\infty]$ *is convex, lower semicontinuous, nonidentically* $+\infty$ *and*

$$\inf\{\varphi(u);\, u \in H\} > -\infty. \tag{2.10}$$

(jj) $\varphi(S(t)x) \leqq \varphi(x)$ *for every* $x \in H$ *and* $t>0$.

Note that condition (jj) can be equivalently expressed as (see [4], [5])

$$(Ax, g_\lambda(x)) \leqq 0 \quad \text{for all } x \in D(A) \text{ and } \lambda > 0. \tag{2.11}$$

The main result is

THEOREM 1. *Suppose that conditions* (i)–(iv) *and* (j), (jj) *hold. Let* $f \in H^1_{\text{loc}}(0,\infty; H)$ *be given such that*

$$f(0) \in D(\varphi), \qquad Af \in L^2_{\text{loc}}(0,\infty; H). \tag{2.12}$$

Then there exists at least one solution $u \in L^2_{\text{loc}}(0,\infty; H)$ *of* (1.1). *If, in addition,* g *is strictly monotone, the solution is unique.*

Suppose further, that $a(t) \neq 0$ *for all* $t>0$, $f' - Af \in L^2(0,\infty; H)$ *and*

$$\lim_{|u|\to+\infty} \varphi(u) = +\infty. \tag{2.13}$$

Then $u \in L^\infty(0,\infty; H)$ *and*

$$a^2(t) \int_0^t |g(u(s))|^2\, ds \quad \textit{bounded on } [0,\infty[. \tag{2.14}$$

3. Proof of Theorem 1. Let $[0,T]$ be any finite interval and let $L\colon L^2(0,T;H) \to L^2(0,T;H)$ be the linear operator defined by

$$(Lu)(t) = \int_0^t E(t-s)u(s)\, ds \quad \text{a.e. } t \in]0,T[, \tag{3.1}$$

where $E(t) = a(t)S(t)$. Since $a \in L^1(0,1)$ and $\|S(t)\|_{L(H,H)} \leqq 1$, we may infer that L is continuous from $L^2(0,T;H)$ into itself.

We begin with the following simple lemma.

LEMMA 1. *Suppose that conditions* (i)–(iv) *hold. Then the operator L is positive, i.e.,*

$$\int_0^t (Lu(s), u(s))\, ds \geqq 0 \quad \textit{for all } u \in L^2(0, T; H) \textit{ and } 0 \leqq t \leqq T. \tag{3.2}$$

Proof. Let $(L_\varepsilon u)(t) = \int_0^t E(t+\varepsilon - s)u(s)\, ds$. Inasmuch as $\lim_{\varepsilon \to 0} L_\varepsilon u = L_u$ in the strong topology of $L^2(0, T; H)$ for every $u \in L^2(0, T; H)$, it suffices to prove that for every $\varepsilon > 0$ the operator L_ε is positive on $L^2(0, T; H)$. For the sake of simplicity we set

$$E_\varepsilon(t) = E(t+\varepsilon) \quad \text{for } t > 0 \text{ and } \varepsilon > 0.$$

As earlier mentioned (see (2.9)), condition (iv) implies that

$$\frac{d}{dt} S(t)x = AS(t)x \quad \text{for all } t > 0 \text{ and } x \in H \tag{3.3}$$

and therefore

$$\frac{d^k}{dt^k} S(t)x = (AS(t/k))^k x \quad \text{for all } t > 0,\ x \in H \text{ and } k = 1, 2, \cdots.$$

In particular, this implies that $S(t) \in C^k(]0, \infty[; L(H, H))$ for every k and

$$(S(t)x, x) \geqq 0, \qquad \left(\frac{d^2}{dt^2} S(t)x, x\right) \geqq 0 \quad \text{for all } t > 0 \text{ and } x \in H. \tag{3.4}$$

Since A is dissipative, it follows by (3.3) that

$$\frac{d}{dt} |S(t)x|^2 \leqq 0 \quad \text{for all } t > 0 \text{ and } x \in H.$$

Therefore

$$\left(\frac{d}{dt} S(t)x, x\right) = \frac{1}{2} \frac{d}{dt} |S(t/2)x|^2 \leqq 0 \quad \text{for all } t > 0 \text{ and } x \in H. \tag{3.5}$$

Combined with (i)–(iii), relations (3.4) and (3.5) imply that

$$(-1)^k E_\varepsilon^{(k)}(t) \geqq 0 \quad \text{for all } t > 0 \text{ and } k = 0, 1, 2. \tag{3.6}$$

(We have used the symbol $\geqq$ for the positiveness.) Corollary 4.1 in [12] can therefore be applied to the present situation to conclude that L_ε is positive on $L^2(0, T; H)$ as claimed.

We now turn to the proof of Theorem 1. Let $T > 0$ be such that $a(T) > 0$. First we shall prove that (1.1) has a solution $u \in L^2(0, T; H)$ on the interval $[0, T]$. This will be proved in several steps.

Step 1. Suppose first that $a(0) < +\infty$. Consider the approximating equations

$$u_\lambda(t) + \int_0^t a(t-s)S(t-s)g_\lambda(u_\lambda(s))\, ds = f(t) \quad \text{a.e. } t \in]0, T[, \tag{3.7}$$

where $g_\lambda = \lambda^{-1}(I - (I + \lambda g)^{-1}) = \partial \varphi_\lambda$ and $\lambda > 0$. Since g_λ is everywhere defined and Lipschitzian on H, it is obvious that (3.7) has a unique bounded solution

$u_\lambda : [0, T] \to H$. The next lemma collects some a priori estimates for u_λ which are needed in the proof.

LEMMA 2. *Let u_λ be the solution of* (3.7). *Then the following inequality holds*

$$2\varphi_\lambda(u_\lambda(t)) + a(t)\int_0^t |g_\lambda(u_\lambda(s))|^2\,ds \leqq 2\varphi_\lambda(f(0)) + \frac{1}{a(t)}\int_0^t |f'(s) - Af(s)|^2\,ds \tag{3.8}$$

for all $t \in [0, T]$ and $\lambda > 0$.

Proof. Since $S \in C^1(]0, \infty[; L(H, H))$, the equation

$$u_{\lambda,\varepsilon}(t) + \int_0^t a(t-s)S(t+\varepsilon-s)g_\lambda(u_{\lambda,\varepsilon}(s))\,ds = f(t) \tag{3.9}$$

has a unique solution $u_{\lambda,\varepsilon} \in C^1([0, T]; H)$ for every $\varepsilon > 0$. Furthermore, inasmuch as g_λ is Lipschitzian on H and $\lim_{\varepsilon \to 0} S(\varepsilon)x = x$ for every $x \in H$, we may infer that

$$\lim_{\varepsilon \to 0} u_{\lambda,\varepsilon}(t) = u_\lambda(t) \quad \text{for every } t \in [0, T] \tag{3.10}$$

and therefore

$$g_\lambda(u_{\lambda,\varepsilon}) \xrightarrow{\varepsilon \to 0} g_\lambda(u_\lambda) \quad \text{strongly in } L^2(0, T; H). \tag{3.11}$$

Next we differentiate (3.9) and use (3.3) to get

$$\begin{aligned} &u'_{\lambda,\varepsilon}(t) + a(0)S(\varepsilon)g_\lambda(u_{\lambda,\varepsilon}(t)) \\ &\quad + \int_0^t a'(t-s)S(t+\varepsilon-s)g_\lambda(u_{\lambda,\varepsilon}(s))\,ds \\ &\quad + A(f(t) - u_{\lambda,\varepsilon}(t)) = f'(t) \quad \text{a.e. } t \in]0, T[. \end{aligned} \tag{3.12}$$

Multiplying both sides of (3.12) by $g_\lambda(u_{\lambda,\varepsilon}(t))$ and integrating over $]0, t[$, we obtain, since

$$\frac{d}{dt}\varphi_\lambda(u_{\lambda,\varepsilon}(t)) = (\partial\varphi_\lambda(u_{\lambda,\varepsilon}(t)), u'_{\lambda,\varepsilon}(t)) \quad \text{a.e. } t \in]0, T[,$$

and by (2.11)

$$(Au_{\lambda,\varepsilon}(t), g_\lambda(u_{\lambda,\varepsilon}(t))) \leqq 0 \quad \text{for } t \in [0, T],$$

the inequality

$$\begin{aligned} &\varphi_\lambda(u_{\lambda,\varepsilon}(t)) + a(0)\int_0^t (S(\varepsilon)g_\lambda(u_{\lambda,\varepsilon}(s)), g_\lambda(u_{\lambda,\varepsilon}(s)))\,ds \\ &\quad + \int_0^t \Big(g_\lambda(u_{\lambda,\varepsilon}(s)), \int_0^s a'(s-\zeta)S(s+\varepsilon-\zeta)g_\lambda(u_{\lambda,\varepsilon}(\zeta))\,d\zeta\Big)\,ds \\ &\quad \leqq \varphi_\lambda(f(0)) + \int_0^t |f'(s) - Af(s)|\,|g_\lambda(u_{\lambda,\varepsilon}(s))|\,ds. \end{aligned} \tag{3.13}$$

This inequality combined with (3.10) and (3.11) yields

$$\varphi_\lambda(u_\lambda(t))+a(0)\int_0 |g_\lambda(u_\lambda(s))|^2\,ds \leqq \varphi_\lambda(f(0)) + \int_0^t |g_\lambda(u_\lambda(s))|\,ds \int_0^s |a'(s-\zeta)|\,|g_\lambda(u_\lambda(\zeta)|\,d\zeta + \int_0^t |f'(s)-Af(s)|\,|g_\lambda(u_\lambda(s))|\,ds \quad \text{for } 0 \leqq t \leqq T. \tag{3.14}$$

Then by applying Schwarz's inequality in (3.14) and making use of condition (ii) we obtain the inequality (3.8) as claimed. Since $\varphi_\lambda(f(0)) \leqq \varphi(f(0))$ and $\varphi_\lambda(u) \geqq \varphi((I+\lambda g)^{-1}u)$ for all $\lambda > 0$ and u in H, condition (j) and inequality (3.8) imply that

$$\{g_\lambda(u_\lambda)\} \quad \text{bounded in } L^2(0, T; H). \tag{3.15}$$

From (3.7) it follows that

$$\left(\int_0^T |u_\lambda(t)|^2\,dt\right)^{1/2} \leqq \left(\int_0^T |f(t)|^2\,dt\right)^{1/2} + \left(\int_0^T |g_\lambda(u_\lambda(t))|^2\,dt\right)^{1/2} \|a\|_{L^1(0,\,T)}. \tag{3.16}$$

Now we may take weakly convergent subsequences. More precisely, there exist $\{u_{\lambda_n}\} \subset \{u_\lambda\}$ such that $\lambda_n \to 0$ as $n \to \infty$ and

$$\begin{aligned} u_{\lambda_n} &\to u \quad \text{weakly in } L^2(0, T; H), \\ g_{\lambda_n}(u_{\lambda_n}) &\to g \quad \text{weakly in } L^2(0, T; H). \end{aligned} \tag{3.17}$$

Clearly,

$$u(t)+\int_0^t a(t-s)S(t-s)g(s)\,ds = f(t) \quad \text{a.e. } t \in]0, T[.$$

Moreover, the maximal monotonicity of the operator $u \to g(u)$ together the positivity of L imply that $u(t) \in D(g)$ and $g(t) \in g(u(t))$ a.e. $t \in]0, T[$. The argument is that used in author's paper [2, Thm. 1] and it will not be repeated here. Thus we may conclude as an intermediate step of the proof that under condition $a(0) < +\infty$, (1.1) has a solution $u(t)$ on the interval $]0, T[$. We note for later use that this solution also satisfies the inequality

$$2\varphi(u(t))+a(t)\int_0^t |g(u(s))|^2\,ds \leqq 2\varphi(f(0)) + \frac{1}{a(t)}\int_0^t |f'(s)-Af(s)|^2\,ds \quad \text{for } t \in [0, T]. \tag{3.18}$$

This follows from (3.8) and (3.17), since

$$\varphi_\kappa(u_\lambda) \geqq \varphi((I+\lambda g)^{-1}u_\lambda), \qquad \lambda g_\lambda(u_\lambda), = u_\lambda - (I+\lambda g)^{-1}u_\lambda$$

and the function φ is lower semicontinuous on H.

Step 2. Now we shall prove essentially the same result as before but in the general case $a \in L^1(0, 1)$. For every $\varepsilon > 0$, we denote by a_ε the function $a(t+\varepsilon)$. According to the first step of the proof, for all sufficiently small ε, the integral equation

$$u_\varepsilon(t) + \int_0^t a_\varepsilon(t-s)S(t-s)g(u_\varepsilon(s))\,ds \ni f(t) \quad \text{a.e. } t \in]0, T[\tag{3.19}$$

has at least one solution $u_\varepsilon \in L^\infty(0, T; H)$. Since $a \in L^1(0, 1)$ and $a(t+\varepsilon) > 0$ for $t \in [0, T]$, it follows from (3.18) and (3.19) that $\{u_\varepsilon\}$ and $\{g(u_\varepsilon)\}$ remain in a bounded subset of $L^2(0, T; H)$.

Therefore, without any loss of generality, we may assume that

$$\begin{aligned} u_\varepsilon &\to u \quad \text{weakly in } L^2(0, T; H), \\ g(u_\varepsilon) &\to g \quad \text{weakly in } L^2(0, T; H) \end{aligned} \tag{3.20}$$

as $\varepsilon \to 0$. We will prove that u satisfies (1.1) on the interval $]0, T[$. The argument is as before but with some simplifications. Define

$$(L_\varepsilon u)(t) = \int_0^t a_\varepsilon(t-s)u(s)\,ds \quad \text{for } u \in L^2(0, t; H).$$

Note that

$$\begin{aligned} &\int_0^T (u_\varepsilon(t) - u_{\varepsilon'}(t), g(u_\varepsilon(t)) - g(u_{\varepsilon'}(t)))\,dt \\ &\quad + \int_0^T (u_\varepsilon(t) - u_{\varepsilon'}(t), (L_\varepsilon g(u_\varepsilon))(t) - L_{\varepsilon'}(g(u_{\varepsilon'}))(t))\,dt = 0 \end{aligned} \tag{3.21}$$

while

$$\begin{aligned} \limsup_{\varepsilon, \varepsilon' \to 0} \int_0^T (g(u_\varepsilon(t)) - g(u_{\varepsilon'}(t)), \int_0^t (a(t+\varepsilon-s) - a(t+\varepsilon'-s)) \\ \cdot S(t-s)g(u_\varepsilon(s))\,ds)\,dt = 0 \end{aligned} \tag{3.22}$$

because $a \in L^1(0, T)$ and $\{g(u_\varepsilon)\}$ is bounded in $L^2(0, T; H)$. Since the operator L_ε is positive on $L^2(0, T; H)$, it follows from (3.21) and (3.22) that

$$\limsup_{\varepsilon, \varepsilon' \to 0} \int_0^T (u_\varepsilon(t) - u_{\varepsilon'}(t), g(u_\varepsilon(t)) - g(u_{\varepsilon'}(t)))\,dt \leqq 0. \tag{3.23}$$

Inasmuch as the operator $u(t) \to (g(u))(t)$ is maximal monotone in $L^2(0, T; H)$, it follows by (3.20) and (3.23) that $g(t) \in g(u(t))$ a.e. on $]0, T[$ (see, e.g., [3, p. 42]).

Step 3. We complete the proof of existence by showing that the function $u(t)$, found before, can be continued on $]T, +\infty[$. The argument is essentially the same as that used in the proof of Theorem 1 in [2]. We reproduce it here. Consider the

integral equation

$$v(t)+\int_0^t a(t-s)S(t-s)g(v(s))\,ds \ni f_0(t) \quad \text{a.e. } t\in]0, T[, \tag{3.24}$$

where $f_0(t)=f(T+t)-\int_0^t a(T+t-s)S(T+t-s)g(u(s))\,ds$. It should be noted that $f_0\in H^1(0, T; H)$ and $Af_0\in L^2(0, T; H)$. (The latter follows from the earlier mentioned fact that $S(t)$ is differentiable on $]0, \infty[$).

Moreover, since $u(t)$ satisfies (1.1) almost everywhere on $]0, T[$, no loss of generality results assuming that $f_0(t)\in D(\varphi)$. Thus according to Step 2, we may conclude that (3.24) has a solution $v\in L^2(0, T; H)$. A simple check with (3.24) yields that $\tilde{u}$ defined by

$$\tilde{u}(t)=\begin{cases} u(t) & \text{for } 0<t<T, \\ v(t-T) & \text{for } T<y<2T, \end{cases}$$

satisfies (1.1) on the interval $]0, 2T[$. Thus by repeatedly making use of this argument, one obtains a solution $u(t)\in L^2_{\text{loc}}(0, \infty; H)$ of (1.1) defined (almost everywhere) on the whole half-axis.

If g is strictly monotone, then from the positivity of L it is immediate that the solution is unique. The remaining part of Theorem 1 is a simple consequence of (3.18). This completes the proof.

We now pause briefly to discuss the integral equation (1.1) with more general kernels $E(t)$ than (2.8). More precisely, we make the following assumptions on $E(t)$:

(a) *For every $t>0$, $E(t)$ is a self-adjoint and bounded linear operator on H.*

(b) $E\in L^1(]0, 1[; L(H, H))\cap C^1(]0, \infty[; L(H, H))$.

(c) $(-1)^k E^{(k)}(t)\geqq 0$ *for all* $t>0$ *and* $k=0, 1$.

(d) $E'(t)-E'(s)\geqq 0$ *for* $t>s>0$.

(e) *There exists* $T>0$ *such that* $E^{-1}(T)\in L(H, H)$.

The arguments used in the proof of Theorem 1 can be applied here, with only minor modifications, provided one first use the results given in [11] to verify that for every $\varepsilon>0$, the operator $u(t)\to\int_0^t (E(t+\varepsilon-s)-\gamma I)u(s)\,ds$ is D-positive in the sense of [16] on the interval $[0, T]$ (γ is a positive number). Despite the generality of conditions (a)–(d), it should be emphasized that condition (e) is rather restrictive, thereby excluding many interesting cases. In the particular case, $E(t)=a(t)S(t)$, studied before, this condition fails unless $S(t)$ is a group of linear continuous operators on H. We expect to give details in a later paper.

4. An example. We shall illustrate the general problem studied before on the following nonlinear boundary value problem

$$\begin{aligned} &\frac{\partial u}{\partial t}(t, x)-\Delta u(t, x)=0 \quad \text{for } x\in\Omega, t>0, \\ &u(0, x)=u_0(x) \quad \text{for } x\in\Omega, \\ &\frac{\partial u}{\partial n}(t, x)\in -g(u(t, x)) \quad \text{for } x\in\Gamma, t>0, \end{aligned} \tag{4.1}$$

where Ω is an open subset of R^n with sufficiently smooth boundary Γ and g is a maximal monotone graph in $R^1 \times R^1$. Here $\Delta = \sum_{i=1}^n \partial^2/\partial^2 x_i$ and $\partial/\partial n$ is the outward normal derivative to Γ. This problem occurs in the generalized heat transfer between solids and gases and was studied by many authors (see [5], [8], [14]). There are other problems of physical interest which reduce to a problem of the form (4.1), where g is multivalued and not everywhere definite on R^1 (see [7]).

By using local charts we may restrict our attention to the case where $\Omega = \{(x', x_n) \in R^n; x_n > 0\}$. Thus in order to avoid a lengthy technical discussion, we will consider only the special case of problem (4.1) where

$$\frac{\partial u}{\partial t} - \Delta u = 0 \quad \text{for } x_n > \text{ and } t > 0, \tag{4.2}$$

$$u(0, x) = 0 \quad \text{for } x_n > 0, \tag{4.3}$$

$$\frac{\partial u}{\partial x_n} \in -g(u) \quad \text{for } x_n = 0 \text{ and } t > 0. \tag{4.4}$$

First, we observe that the linear problem consisting of (4.2), (4.3) and the boundary condition

$$\frac{\partial u}{\partial x_n}(t, x', 0) = h(t, x'), \qquad t > 0, \quad x' \in R^{n-1},$$

can be solved, for example, by Fourier transforms, and the result is

$$u(t, x) = (2\sqrt{\pi})^{-n} \int_0^t (t-s)^{-n/2} \exp\left(-(|x_n|^2 + |x' - \xi'|^2)/4(t-s)\right) \cdot h(s, \xi')\, d\xi'\, ds.$$

Thus we are led to the nonlinear integral equation

$$u(t) + (2\sqrt{\pi})^{-1} \int_0^t (t-s)^{-1/2} S(t-s) g(u(s))\, ds \ni 0, \qquad t > 0 \tag{4.5}$$

on the space $L^2(R^{n-1})$ where $u(t) = u(t, x')$, $(g(u(t)))(x') = g(u(t, x'))$ and $S(t)$: $[0, \infty[\to L(L^2(R^{n-1}), L^2(R^{n-1}))$ is defined by

$$(S(t)h)(x') = \begin{cases} (2\sqrt{\pi})^{-n+1} t^{-(n-1)/2} \displaystyle\int_{R^{n-1}} \exp\left(-|x' - \xi'|^2/4t\right) h(\xi')\, d\xi' & \text{for } t > 0 \\ h(x') & \text{for } t = 0. \end{cases}$$

The family of operators $S(t)$ is known to form an analytic semigroup of linear contractions on $L^2(R^{n-1})$ with infinitesimal generator Δ: $\Delta h = \sum_{i=1}^{n-1} \partial^2 h/\partial x_i^2$, and domain $D(\Delta) = \{h \in L^2(R^{n-1}); \Delta h$ exists in the sense of distributions and belongs to $L^2(R^{n-1})\}$.

Theorem 1 can therefore be applied to the present situation, where $a(t) = (2\sqrt{\pi})^{-1}t^{-1/2}$ and $\varphi : L^2(R^{n-1}) \to]-\infty, +\infty]$ is defined by

$$\varphi(u) = \int_{R^{n-1}} j(u(x'))\, dx'$$

and $j : R^{-1} \to]-\infty, +\infty]$ is such that $\partial j = g$.

REFERENCES

[1] S. AIZICOVICI, *Abstract integral equations of Volterra type*, Atti Acad. Naz. Lincei, 6 (1975), pp. 868–879.

[2] V. BARBU, *Nonlinear Volterra equations in Hilbert space*, this Journal, 6 (1975), pp. 728–741.

[3] ———, *Nonlinear Semigroups and Evolution Equations in Banach Space*, Publishing House Acad. Soc. Rep. Romania, Bucharest, and Noordhoff, Leyden, The Netherlands, 1975.

[4] H. BRÉZIS, *Opérateurs Maximaux Monotones et Semigroupes de Contractions dans les Espaces de Hilbert*, Math. Studies, No. 5, North-Holland, Amsterdam, 1973.

[5] ———, *Monotonicity Methods in Hilbert Spaces and some Applications to Nonlinear Partial Differential Equations*, Contributions to Nonlinear Functional Analysis, E. Zarantonello, ed., Academic Press, New York, 1972.

[6] C. CORDUNEANU, *Integral Equations and Stability of Feedback Systems*, Academic Press, New York, 1973.

[7] G. DUVAUT AND J. L. LIONS, *Sur les Inéquations en Méchanique et en Physique*, Dunod, Paris, 1972.

[8] A. FRIEDMAN, *Generalised heat transfer between solids and gases under nonlinear boundary conditions*, J. Math. Mech., 8 (1959), pp. 161–184.

[9] J. J. LEVIN, *A bound on the solutions of a Volterra equation*, Arch. Rational Mech. Anal., 52 (1973, pp. 339–349.

[10] S. O. LONDEN, *Volterra equations on a Hilbert space*, Trans. Amer. Math. Soc., to appear.

[11] ———, *On an integral equation in a Hilbert space*, Rep. 1527, Math. Research Center, Madison, Wisconsin, 1975.

[12] R. C. MACCAMY AND J. S. WONG, *Stability theorems for some functional equations*, Trans. Amer. Math. Soc., 164 (1972), pp. 1–37.

[13] R. C. MACCAMY, *Nonlinear Volterra equations on a Hilbert space*, J. Differential Equations, 17 (1975), pp. 232–254.

[14] W. R. MANN AND F. WOLF, *Heat transfer between solids and gases under nonlinear boundary conditions*, Quart. Appl. Math., 9 (1951), pp. 163–184.

[15] R. K. MILLER, *Nonlinear Volterra Integral Equations*, W. A. Benjamin, New York, 1971.

[16] J. S. W. WONG, *Positive definite functions and Volterra integral equations*, Bull. Amer. Math. Soc., 80 (1974), pp. 679–682.

SIAM J. MATH. ANAL.
Vol. 8, No. 2, April 1977

A REPRESENTATION FOR DISTRIBUTIONAL SOLUTIONS OF PARABOLIC PROBLEMS*

HAROLD D. MEYER†

Abstract. A boundary-integral representation is derived for distributional solutions of the parabolic problem

$$\sum_{|p|,\,|q|\leq m} (-1)^{|p|} D_x^p (a_{pq}(x,t) D_x^q u) + \frac{\partial u}{\partial t} = 0$$

in a finite cylinder, where the base of the cylinder is an analytic manifold. The method involves use of a duality principle at the boundary and extends results in [12]. A corollary provides a version of the result suitable for practical applications.

1. Introduction. The purpose of the present paper is to derive a boundary-integral type representation for distributional solutions of parabolic problems where the domain is a finite cylinder with bounded base whose boundary is an analytic manifold. A corollary provides a more restrictive version of the above which is suitable for practical applications. The problem considered is

$$\sum_{|p|,\,|q|\leq m} (-1)^{|p|} D_x^p (a_{pq}(x,t) D_x^q u) + \frac{\partial u}{\partial t} = 0, \tag{1.1}$$

where specifics about notation, domains and initial-boundary conditions will be discussed later.

Work here extends results in [12]. Other representations in the literature of related interest are found in [7, p. 32], [9], [13], [15], [16]. Our representations are presented for two reasons. First, they provide a characterization for distributional solutions; also and more important, the second of these can be used in numerical continuation procedures similar to those in [2], [4], [5], [13], [15].

2. Preliminaries. Much of the discussion in this and the next section is treated in great depth in Lions and Magenes [11]. The reader is referred there for additional details.

Let Ω be an open, bounded domain in the n-dimensional Euclidean space R_x^n and have boundary Γ, which is required to be an $(n-1)$-dimensional analytic variety with Ω locally only on one side. We take $R_x^n \times R_t$ to be the Cartesian product of R_x^n and the one-dimensional Euclidean space R_t, with typical point $(x, t) = (x_1, \cdots, x_n, t)$, and define $Q = \Omega \times (0, T) \subset R_x^n \times R_t$, $0 < T < \infty$, $\Sigma = \Gamma \times (0, T)$. Designate $\bar{Q}$ and $\bar{\Omega}$ to be the closures of Q and Ω, and ∂Q to be the entire boundary of Q. Relative to other set later on, overbars will also represent closures.

Let $D_{x_i} = \partial/\partial x_i$, $D_t = \partial/\partial t$, and $D_x^\alpha = D_{x1}^{\alpha_1} \cdots D_{x_n}^{\alpha_n}$, where $\alpha = (\alpha_1 \cdots, \alpha_n)$ and the α_i are nonnegative integers. For n-tuples such as α, we set $|\alpha| = \alpha_1 + \cdots + \alpha_n$. All functions here will be complex-valued. Take $\mathscr{D}(Q)$, $\mathscr{D}(\bar{Q})$, $\mathscr{E}(Q)$, $\mathscr{D}(\bar{\Omega})$, $\mathscr{D}(\bar{\Sigma})$ and $C^\infty(\Gamma)$ to be the standard spaces of infinitely differentiable functions provided with the usual topologies [11], [8]; $\mathscr{D}'(Q)$ is the strong dual of $\mathscr{D}(Q)$, i.e., the

* Received by the editors February 28, 1975, and in final revised form October 6, 1975.

† Department of Mathematics, Texas Tech University, Lubbock, Texas 79409.

standard Schwartz distribution space. Let $\mathscr{D}_{2m,1}(\bar{Q})$, $\mathscr{D}_{2m,1}(\bar{\Sigma})$ and $\mathscr{D}_{2m,1}(G)$, for G open, be the standard Gevrey spaces [11, p. 8] of order $2m$ in t and order 1 in x.

For functions defined on Q, we take

$$A = A(x, t, D_x) = \sum_{|p|,|q| \leqq m} (-1)^{|p|} D_x^p (a_{pq}(x, t) D_x^q), \tag{2.1}$$

$$P = A(x, t, D_x) + D_t, \tag{2.2}$$

where p and q are n-tuples treated as α before. The operators A^* and $P^* = A^*(x, t, D_x) - D_t$ are, respectively, the formal adjoints of A and P.

Concerning A, we require that

(I) the coefficients $a_{pq} \in \mathscr{D}_{2m,1}(\bar{Q})$ and that there be a $t_0 > 0$ such that each $a_{pq}(x, t) = a_{pq}(x)$ for $0 \leqq t \leqq t_0$;

(II) given any $t_1 \in [0, T]$ and any $\theta \in [-\pi/2, \pi/2]$, $A(x, t_1, D_x) + (-1)^m e^{i\theta} D_y^{2m}$ is properly elliptic (see [10, Chap. 2, § 1]) in $\bar{\Omega} \times R_y$ (R_y is the analogue with respect to y of R_t);

(III) $-A(0) = -A(x, 0, D_x)$ is an infinitesimal generator (see [1, p. 9, 15]) of an analytic semi-group.

In (III), $-A(0)$ is regarded as an unbounded operator having domain $D(A(0)) = \{u | D_x^p u \in L^2(\Omega),\ |p| \leqq 2m;\ B_j(x, 0, D_x)u = 0,\ j = 0, \cdots, m-1,\ B_j\text{'s}$ as defined below$\} \subset L^2(\Omega)$. Note that (III) implies a parallel statement for $-A^*(0) = -A^*(x, 0, D_x)$. By our definition of A, the operator P is parabolic in accordance with Petrowski [14].

Let $\{B_0, \cdots, B_{m-1}\}$ represent a system of boundary operators on $\bar{\Sigma}$, where

$$B_j = B_j(x, t, D_x) = \sum_{|h| \leqq m_j} b_{jh}(x, t) D_x^h, \qquad j = 0, \cdots, m-1, \tag{2.3}$$

and $0 \leqq m_j \leqq 2m - 1$.

It is required that

(a) the coefficients $b_{jh} \in \mathscr{D}_{2m,1}(\bar{\Sigma})$ and that there be a $t_0 > 0$ such that each $b_{jh}(x, t) = b_{jh}(x)$ for $0 \leqq t \leqq t_0$;

(b) given any $t_0 \in [0, T]$, $\{B_0(x, t_0, D_x), \cdots, B_{m-1}(x, t_0, D_x)\}$ is a normal system (see [10, p. 113]) on Γ;

(c) given any $t_1 \in [0, T]$ and any $\theta \in [-\pi/2, \pi/2]$, $\{B_0(x, t_1, D_x), \cdots, B_{m-1}(x, t_1, D_x)\}$ covers (see [10, p. 113]) the operator $A(x, t_1, D_x) + (-1)^m e^{i\theta} D_y^{2m}$ on $\Gamma \times R_y$.

In (c), $A(x, t_1, D_x) + (-1)^m e^{i\theta} D_y^{2m}$ is considered as an operator in $R_x^n \times R_y$.

Corresponding to $\{B_0, \cdots, B_{m-1}\}$, there exist [11, p. 209] systems of boundary operators $\{C_0, \cdots, C_{m-1}\}$, $\{S_0, \cdots, S_{m-1}\}$, $\{T_0, \cdots, T_{m-1}\}$ for which the standard Green's formula [11, p. 209] is valid. These operators have the same form as (2.3). The coefficients c_{jh}, s_{jh}, t_{jh}, which correspond to the b_{jh}, belong to $\mathscr{D}_{2m,1}(\bar{\Sigma})$ [11, p. 209]. Further, the s_{jh} may be chosen to be independent of t. The c_{jh} and t_{jh} will then be functions of x alone for $0 \leqq t \leqq t_0$ [11, p. 212].

3. The problem. After a few more preliminaries, we shall specify precisely the solutions for which representations will be given.

Let

$$X = \{u \in \mathscr{D}(\bar{Q}) | P^* u \in \mathscr{D}(Q),\ u(x, T) = 0,\ C_j u = 0, j = 0, \cdots, m-1\}, \tag{3.1}$$

$$Y = \{u \in \mathscr{D}'(Q) | Pu \in \Xi'(Q)\}. \tag{3.2}$$

The first of these spaces is provided with the inductive limit topology given by the union of the Fréchet spaces

$$X_K = \{u \in \mathscr{D}(\bar{Q}) | P^*u \in \mathscr{D}_K(Q),\ u(x, T) = 0,\ C_j u = 0,\ j = 0, \cdots, m-1\}, \tag{3.3}$$

where the K's are an increasing sequence of compact sets. In (3.3), the $\mathscr{D}_K(Q)$ are the standard Fréchet spaces made up of those functions in $\mathscr{D}(Q)$ having support in the respective K's. Each of the X_K's has the topology generated by the natural seminorms for this situation. The topology on X is the weakest locally convex one such that the linear maps $u \to u$ of X into $\mathscr{D}(\bar{Q})$ and $u \to P^*u$ of X into $\mathscr{D}(Q)$ are continuous. The space Y is equipped with the weakest locally convex topology for which the linear maps $u \to u$ of Y into $\mathscr{D}'(Q)$, and $u \to Pu$ of Y into $\Xi'(Q)$ are continuous. Denote Y' to be the strong dual of Y. In (3.2), $\Xi'(Q)$ is the standard distributional space described in [11, p. 210].

Let us define V to be the image of X under the map

$$\gamma u = (u(x, 0), T_0 u, \cdots, T_{m-1} u). \tag{3.4}$$

For any $L > 0$, take

$$\begin{aligned} V_L = \{&(u', u_0, \cdots, u_{m-1}) \in D^L((A^*)^\infty(0); k!) \times [\mathscr{D}([0, T]; \mathscr{H}_L(\Gamma))]^m | \\ &\text{(a) } \exists C > 0 \text{ for which } \|D_t^k u_j(\cdot, t)\|_{\mathscr{H}_L} \leqq CL^k k!, \text{ when } 0 \leqq t \leqq 1/L,\ j = 0, \cdots, m-1,\ k = 0, 1, \cdots; \\ &\text{(b) } D_t^k u_j(x, 0) = T_j(x, 0)((A^*)^k(0) u'(x)) \text{ for } x \in \Gamma,\ j = 0, \cdots, m-1,\ k = 0, 1, \cdots; \\ &\text{(c) } \exists C > 0 \text{ for which } \|D_t^k u_j(\cdot, t)\|_{\mathscr{H}_L} \leqq CL^k (k!)^{2m}, \text{ when } 1/L \leqq t \leqq T,\ j = 0, \cdots, m-1,\ k = 0, 1, \cdots; \\ &\text{(d) } u_j(t) = 0 \text{ for } T - 1/L \leqq t \leqq T,\ j = 0, \cdots, m-1\}. \end{aligned} \tag{3.5}$$

When we refer to conditions (a)–(d) later, we shall be referring to (a)–(d) in the brackets above.

In (3.5), as in [11, pp. 10 and 133],

$$\mathscr{H}_L(\Gamma) = \left\{u \in C^\infty(\Gamma) | \|u\|_{\mathscr{H}_L} \equiv \sup_{\substack{x \in \Gamma \\ k = 0, 1, \cdots}} \left| \frac{\Delta_\Gamma^k u(x)}{L^k (2k)!} \right| < \infty \right\}, \tag{3.6}$$

$$D^L(A^{*\infty}(0); k!) = \left\{u \in L^2(\Omega) | \|u\|_{D^L} \equiv \sup_{k = 0, 1, \cdots} \frac{\|A^{*k}(0) u\|_{L^2(\Omega)}}{L^k k!} < \infty \right\}, \tag{3.7}$$

for $L > 0$, where Δ_Γ is the Laplace–Beltrami operator on Γ [6]. Further, $\mathscr{D}([0, T]; \mathscr{H}_L(\Gamma))$ is the standard Schwartz space of infinitely differentiable functions on $[0, T]$ with values in $\mathscr{H}_L(\Gamma)$ ([11, p. 11], $\mathscr{D}([0, T]; \mathscr{H}_L(\Gamma)) = \mathscr{D}(\mathscr{I}, [0, T]; \mathscr{H}_L(\Gamma))$ there, with $[0, T] \subset \mathscr{I}$).

A topology for V_L is given by the norm

$$\|(u', u_0, \cdots, u_{m-1})\|_{W_L} = \|u'\|_{D^L} + \sum_{j=0}^{m-1} \sup_{k=0,1,\cdots} \left\{ \sup_{0\leqq t\leqq 1/L} \frac{\|D_t^k u_j(\cdot, t)\|_{\mathscr{H}_L}}{L^k k!} + \sup_{1/L\leqq t\leqq T} \frac{\|D_t^k u_j(\cdot, t)\|_{\mathscr{H}_L}}{L^k (k!)^{2m}} \right\}, \tag{3.8}$$

making it a Banach space according to the discussion in [11, p. 214].

It is shown in [11, pp. 213–214] that V, as defined previously, is equal to $\cup_{L=1}^{\infty} V_L$ with the inductive limit topology. We denote the strong dual by V'.

Define the boundary trace σ: $\mathscr{D}(\bar{Q}) \to \mathscr{D}(\bar{\Omega}) \times [\mathscr{D}(\bar{\Sigma})]^m$ to be given by

$$\sigma u = (u(x, 0), B_0 u, \cdots, B_{m-1} u) \tag{3.9}$$

for $u \in \mathscr{D}(\bar{Q})$. We have the following trace result proved in [11, p. 218].

LEMMA 3.1. *The map σ: $\mathscr{D}(\bar{Q}) \to \mathscr{D}(\bar{\Omega}) \times [\mathscr{D}(\bar{\Sigma})]^m$ can be extended by continuity to a continuous linear map of Y into V'. The topologies are the weak $\sigma(Y, Y')$ for Y and the weak $\sigma(V', V)$ for V'. This new trace will also be designated σ.*

Next, let $\mathscr{H}(\Gamma)$ be the space of analytic functions on Γ [11, p. 10]. It is the inductive limit of the spaces $\mathscr{H}_L(\Gamma)$ from before. By $\mathscr{H}'(\Gamma)$ is meant the strong dual of $\mathscr{H}(\Gamma)$. Take $\mathscr{D}_{-,2m}((0, T); \mathscr{H}(\Gamma))$ to be the standard space of infinitely differentiable functions on $(0, T)$ with values in $\mathscr{H}(\Gamma)$ (see [11, p. 13], with $M_k = (k!)^{2m}$). Their support is bounded on the left of $(0, T)$. Let $\mathscr{D}'_{+,2m}((0, T); \mathscr{H}'(\Gamma))$ represent the strong dual.

Further, define $D(A^{*\infty}(0); k!)$ [11, *p.* 132] to be the inductive limit of the spaces $D^L(A^{*\infty}(0); k!)$. (Each of the $D^L(A^{*\infty}(0); k!)$ has the Banach space topology provided by the norms $\|\cdot\|_{D^L}$.)

Remark. Note that the sequence $\{k!\}$ in $D(A^{*\infty}(0); k!)$ does not satisfy the assumption of nonquasi-analyticity [11, (1.23), p. 91]. Thus as discussed in [11], $D(A^{*\infty}(0); k!)$ could degenerate into a trivial space. For our case, however, by assumption (III) of § 2, $-A^*(0)$ is the infinitesimal generator of an *analytic* semigroup. Hence by footnote 1 of [11, p. 146], $D(A^{*\infty}(0); k!)$ is nontrivial.

We further remark that, because $D(A^{*\infty}(0); k!)$ is nontrivial, the representation [11, (7.74), p. 144], used in the proof of Lemma 4.1 below, is valid. The same proof as used there carries over to (7.74) for the present situation. In view of this, (4.1) will be nontrivial.

There exists no general description of elements in V'; however, we can specify a subclass V^* for which there is a description. We define V^* to consist of the functionals of form

$$L(u') + \sum_{j=0}^{m-1} L_j(u_j) \quad \forall (u', u_0, \cdots, u_{m-1}) \in V, \tag{3.10}$$

where $(L, L_0, \cdots, L_{m-1})$ denotes any element in $(D(A^{*\infty}(0); k!))' \times [\mathscr{D}'_{+,2m}((0, T); \mathscr{H}'(\Gamma))]^m$. By the definition of V (in terms of the V_L), clearly $V \subset D(A^{*\infty}(0); k!) \times [\mathscr{D}_{-,2m}((0, T); \mathscr{H}(\Gamma))]^m$. Thus $V^* \subset V'$. The description of elements in V^* will be given in the next section.

Now we can prescribe exactly the solutions for which representations will be given. These will be the distributions falling in the space

$$\mathbb{P} = \{u \in \mathscr{D}'(Q) | Pu = 0\}. \tag{3.11}$$

For the corollary, they will be the solutions

$$\mathbb{P}^* = \{u \in \mathscr{D}'(Q) | Pu = 0, \sigma u \in V^*\}. \tag{3.12}$$

Clearly, solutions in $\mathbb{P}$ have their traces in V'.

4. Structure of elements in V^*. Let var (μ) denote the total variation of a measure μ on its set of definition, which for our purposes will be either Ω or Σ. Later we shall need the following characterization of elements in V^*:

LEMMA 4.1. *Let $F \in V^*$. Then there exist measures μ_k on Ω and sets of measures $\mu_{kl}^{(j)}$, $j = 0, \cdots, m-1$, on Σ such that*

$$F(u) = \sum_{k=0}^{\infty} \int_{\Omega} A^{*k}(0) u' \, d\mu_k + \sum_{j=0}^{m-1} \sum_{k,l=0}^{\infty} \int_{\Sigma} \Delta_{\Gamma}^{l} D_i^l u_j \, d\mu_{kl}^{(j)} \tag{4.1}$$

$$\forall u = (u', u_0, \cdots, u_{m-1}) \in V.$$

The measures μ_k satisfy the condition

$$\sum_{k=0}^{\infty} L^k k! \operatorname{var}(\mu_k) < \infty \quad \forall L > 0. \tag{4.2}$$

The measures $\mu_{kl}^{(j)}$ can be selected such that for some $T^* = (T_0^*, \cdots, T_{m-1}^*)$,

$$\mu_{kl}^{(j)} = 0 \quad \text{for } t < T_j^*, \quad k, l = 0, 1, \cdots, \quad j = 0, \cdots, m-1, \tag{4.3}$$

$$\sum_{k=0}^{\infty} L^k (2k)! \operatorname{var}(\mu_{kl}^{(j)}(\cdot, t)) < \infty \quad \forall L > 0, \quad \text{any } l = 0, 1, \cdots, \tag{4.4}$$

$$j = 0, \cdots, m-1, \quad \text{and any } t \in (0, T).$$

Also they can be selected such that for any continuous function $\chi: (0, T) \to \mathbb{C}$ (one-dimensional complex space) having compact support,

$$\sum_{k,l=0}^{\infty} M^l (l!)^{2m} \int_{\Sigma} \Delta_{\Gamma}^{k} v \chi \, d\mu_{kl}^{(j)} < \infty \quad \forall M > 0,$$
$$\forall v \in \mathscr{D}^0((0, T); \mathscr{H}(\Gamma)) \quad \text{and } j = 0, \cdots, m-1. \tag{4.5}$$

In (4.5), $\mathscr{D}^0((0, T); \mathscr{H}(\Gamma))$ *consists of all continuous functions of* $(0, T) \to \mathscr{H}(\Gamma)$ *with compact support, provided with the usual Schwartz topology* [11, p. 19].

Proof. Let $(L, L_0, \cdots, L_{m-1})$ be as in (3.10). Taking $E = L^2(\Omega)$ and $M_k = k!$, we have from [11, (7.74)–(7.75), p. 144] (see also the Remark here following Lemma 3.1) that

$$L(u') = \sum_{k=0}^{\infty} \langle e_k, A^{*k}(0) u' \rangle_{L^2} \forall u' \in D(A^{*\infty}(0); k!) \tag{4.6}$$

for some sequence $\{e_k\} \subset L^2(\Omega)$, where

$$\sum_{k=0}^{\infty} L^k k! \|e_k\|_{L^2} < \infty \quad \forall L > 0. \tag{4.7}$$

In the above, $\langle \cdot, \cdot \rangle_{L^2}$ and $\| \cdot \|_{L^2}$ are the inner product and norm in $L^2(\Omega)$.

Taking $\mu_k(R) = \int_R e_k \, dx$, where R is any measurable subset of Ω, we have from (3.7) that

$$L(u') = \sum_{k=0}^{\infty} \int_{\Omega} A^{*k}(0) u' \, d\mu_k(x). \tag{4.8}$$

By (4.7), the μ_k satisfy (4.2).

Next, let $M_l = (l!)^{2m}$ in [11, Thm. 5.2, p. 20]. Then

$$L_j = \sum_{l=0}^{\infty} D_t^l \mu_l^{(j)}, \qquad j = 0, \cdots, m-1, \tag{4.9}$$

for some sequence of measures $\mu_l^{(j)}$ in $(\mathscr{D}^0((0, T); \mathscr{H}'(\Gamma)))'$, the strong dual of $\mathscr{D}^0((0, T); \mathscr{H}(\Gamma))$. These are measures on $(0, T)$ with their values in $\mathscr{H}'(\Gamma)$ [11, p. 19]. (Expression (4.9) should be taken in the sense that

$$L_j(u_j) = \sum_{l=0}^{\infty} (-1)^l \langle \mu_l^{(j)}, D_t^l u_j \rangle \quad \forall u_j \in \mathscr{D}'_{+,2m}((0, T); \mathscr{H}'(\Gamma)),$$

where $\langle \cdot, \cdot \rangle$ represents the duality between $(\mathscr{D}^0((0, T); \mathscr{H}'(\Gamma)))'$ and $\mathscr{D}^0((0, T); \mathscr{H}(\Gamma))$.) By [11, (5.21)–(5.22), p. 20], the measures $\mu_l^{(j)}$ satisfy

$$\sum_{l=0}^{\infty} M^l (l!)^{2m} \chi \operatorname{var}(\mu_l^{(j)}) < \infty \quad \forall M > 0, \tag{4.10}$$

and have their support with respect to t bounded on the left.

Since these measures have their values in $\mathscr{H}'(\Gamma)$, by [11, Thm. 3.1, p. 11], we can write that

$$\mu_l^{(j)}(x, t) = \sum_{k=0}^{\infty} \Delta_{\Gamma}^k \mu_{kl}^{(j)}(x, t), \qquad l = 0, 1, \cdots, \tag{4.11}$$

where the $\mu_{kl}^{(j)}$ are measures on Γ satisfying (see [11, (3.7), p. 11]),

$$\sum_{k=0}^{\infty} L^k (2k)! \operatorname{var}(\mu_{kl}^{(j)}) < \infty \quad \forall L > 0, \qquad l = 0, 1, \cdots. \tag{4.12}$$

Combining (4.9) and (4.11) and writing the result in integral form, we have that

$$L_j(u_j) = \sum_{k,l=0}^{\infty} \int_{\Sigma} \Delta_{\Gamma}^k D_t^l u_j \, d\mu_{kl}^{(j)} \quad \forall u_j \in \mathscr{D}'_{+,k}((0, T); \mathscr{H}'(\Gamma)). \tag{4.13}$$

Further, (4.10) and (4.12) give (4.4) and (4.5).

Combining these results with those for L, the lemma is proved.

5. The Green's function. It is a standard result, for our hypotheses, that a Green's function $G(x, t; \xi, \tau)$ satisfying

$$
\begin{aligned}
&P^*G(x, t; \xi, \tau) = \delta(x-\xi, t-\tau), \qquad (\xi, \tau) \in Q, \\
&C_jG(x, t; \xi, \tau) = 0, \qquad (\xi, \tau) \in \Sigma, \quad j = 0, \cdots, m-1, \\
&G(x, t; \xi, T) = 0, \qquad \xi \in \Omega,
\end{aligned} \tag{5.1}
$$

will exist. In the above, δ is the Dirac delta distribution, while the operations P^* and C_j are performed with respect to ξ and τ. From the above, for $(x, t) \in Q$ fixed and $G(\xi, \tau) \equiv G(x, t; \xi, \tau)$, G satisfies

$$
\begin{aligned}
&P^*G(\xi, \tau) = 0, \qquad (\xi, \tau) \in Q - \{(x, t)\}, \\
&C_jG(\xi, \tau) = 0, \qquad (\xi, \tau) \in \Sigma, \quad j = 0, \cdots, m-1, \\
&G(\xi, T) = 0, \qquad \xi \in \Omega.
\end{aligned} \tag{5.2}
$$

In terms of G, any solution $u \in \mathbb{P} \cap \mathscr{D}(\bar{Q})$ can be written as

$$
\begin{aligned}
u(x, t) = &\sum_{j=0}^{m-1} \int_{\Sigma} B_j u(\xi, \tau) T_j G(x, t; \xi, \tau)\, d\sigma \\
&+ \int_{\Omega} u(\xi, 0) G(x, t; \xi, 0)\, d\xi,
\end{aligned} \tag{5.3}
$$

where the B_j, T_j operate with respect to ξ and τ, $d\xi = d\xi_1 \cdots d\xi_n$, and the first integration on the right is with respect to ξ and τ.

Remark 1. One way to see that (5.1) has a solution is to apply [11, Thm. 3.5, p. 220]. To do this, take $T-t$ in place of t in (3.41)–(3.42) there. Also, take

$$\sigma u = (u(x, T), C_0 u, \cdots, C_{m-1} u),$$

and redefine the spaces X, Y, V, V' accordingly. Then one is studying a problem

$$
\begin{aligned}
&P^*u = f \quad \text{in the sense of } \mathscr{D}'(Q), \\
&\sigma u = g,
\end{aligned}
$$

in place of (3.41)–(3.42), with "initial" data now on $t = T$. Taking $f = \delta \in \Xi'(Q)$ and $g = 0$, the theorem provides a solution $G \in \mathscr{D}'(Q)$.

Further, recall the first equation of (5.2) and observe that P^* satisfies the hypotheses of [11, Thm. 1.1, p. 192] on $Q - \{(x, t)\}$. Thus $G \in C^\infty(Q - \{(x, y)\})$.

Expression (5.3) follows in the usual fashion using the Green's formula [11, (3.3), p. 209] along with our boundary information.

Remark 2. Note that $G(\xi, \tau)$ will be zero for $t < \tau \leqq T$. To see this, observe that

$$
\begin{aligned}
&P^*G(\xi, \tau) = 0, \qquad (\xi, \tau) \in \Omega \times (t, T], \\
&C_jG(\xi, \tau) = 0, \qquad (\xi, \tau) \in \Gamma \times (t, T], \quad j = 0, \cdots, m-1, \\
&G(\xi, T) = 0, \qquad \xi \in \Omega.
\end{aligned} \tag{5.4}
$$

Problem (5.4) has a unique solution in $\Omega \times (t, T]$ (by [11, Prop. 3.1, p. 209], taking $Q = \Omega \times (t, T)$ there). Since $u = 0$ is a solution, we thus have that G must be zero.

LEMMA 5.1. *As a function of* (ξ, τ), $\gamma G(x, t; \xi, \tau) = (G(x, t; \xi, 0), T_0 G(x, t; \xi, \tau), \cdots, T_{m-1} G(x, t; \xi, \tau))$ *is a vector in* V *for any* $(x, t) \in Q$.

Proof. The proof is suggested by work in [11]. As before write $G(\xi, \tau) \equiv G(x, t; \xi, \tau)$. It is sufficient to show that $\gamma G \in V_L$ for some $L > 0$, where V_L is given by (3.5).

First note that in a neighborhood N_0 of $\tau = 0$, by (5.2),

$$\begin{aligned} &P^* G(\xi, \tau) = 0, \\ &C_j G(\xi, \tau) = 0 \quad \text{on } \Sigma, \ j = 0, \cdots, m-1. \end{aligned} \tag{5.5}$$

By our assumptions on A, A^* is independent of τ here. Thus $A^*(\tau) = A^*(0)$. The hypotheses of [11, Remark 7.11, p. 146] are satisfied. Using this, G is analytic with respect to τ in N_0 and $G \in D(A^{*\infty}(0); k!)$. Thus for some $L_1 > 0$, $G(\xi, 0) \in D^{L_1}(A^{*\infty}(0); k!)$, and, by the analyticity of G and the fact that the coefficients of the T_j are in $\mathscr{D}_{2m,1}(\bar{\Sigma})$, $T_j G$, $j = 0, \cdots, m-1$, satisfies condition (a) of (3.5).

Next, observe that by the first equation in (5.5), since $A^*(\tau) = A^*(0)$,

$$D_\tau G(\xi, \tau) = A^*(0) G(\xi, \tau) \quad \text{in } N_0.$$

Thus

$$D_\tau^k G(\xi, \tau) = A^{*k}(0) G(\xi, \tau) \quad \text{in } N_0, \quad k = 0, 1, \cdots.$$

From this, we get

$$D_\tau^k (T_j G(\xi, 0)) = T_j(\xi, 0)(A^{*k}(0) G(\xi, 0))$$

for $\xi \in \Gamma$, $j = 0, \cdots, m-1$, $k = 0, 1, \cdots$, and condition (b) of (3.5) holds for some $L_2 > 0$.

In a neighborhood N_Σ of Σ, we again have (5.5). By [11, Thm. 2.3, p. 206], then $G(\xi, \tau) \in \mathscr{D}_{2m,1}(N_\Sigma)$. This means that G is an analytic function of ξ in N_Σ. Also the coefficients of T_j, as mentioned, are in $\mathscr{D}_{2m,1}(\bar{\Sigma})$, so that $T_j G \in \mathscr{H}(\Gamma)$ for each $\tau \in (0, T)$. Further, since both the coefficients of T_j and G are in the Gevrey class of order $2m$ in t, condition (c) of (3.5) now follows for $T_j G$, $j = 0, \cdots, m-1$, for some $L_3 > 0$.

From the above discussion, we also have that $T_j G$ is infinitely differentiable in τ, on Σ. Thus each $T_j G \in \mathscr{D}([0, T]; \mathscr{H}_{L_4}(\Gamma))$ for some $L_4 > 0$.

Lastly, since $G(\xi, \tau) = 0$ for $t < \tau \leqq T$, condition (d) holds for $T_j G$, $j = 0, \cdots, m-1$, with respect to some $L_5 > 0$.

Picking $L = \max\{L_1, L_2, L_3, L_4, L_5\}$, then conditions (a)–(d) of (3.5) are satisfied for this L. Further, $\gamma G \in D^L(A^{*\infty}(0); k!) \times [\mathscr{D}([0, T]; \mathscr{H}_L(\Gamma))]^m$, and we are done.

6. An approximation theorem. Before obtaining the representations, we shall need an approximation theorem. It is proved using methods suggested by work of Saylor [15] and Lions and Magenes [11].

THEOREM 6.1. *The space* $\mathbb{P} \cap \mathscr{D}(\bar{Q})$ *is dense in* $\mathbb{P}$ *relative to the* $\mathscr{E}(Q)$ *topology.*

To prove this we shall first need a lemma.

LEMMA 6.1. *The subspace* $P^* \mathscr{D}(Q)$ *is closed in* $\mathscr{D}(Q)$.

Proof. It is clear that $P^* \mathscr{D}(Q)$ is a subspace of $\mathscr{D}(Q)$. It remains to show that, given a sequence $\{v_k\} \subset \mathscr{D}(Q)$ such that $P^* v_k \to w$ with $w \in \mathscr{D}(Q)$, then w must be in $P^* \mathscr{D}(Q)$. Take $P^* v_k = w_k$.

Note that by [11, Prop. 3.1, p. 209], since $w_k \to w$ in $\mathcal{D}(Q)$, there will exist a $v \in X$ such that $v_k \to v$ in X with $P^*v = w$. We must show that $v \in \mathcal{D}(Q)$. Since $v \in X$, we already know that v is infinitely differentiable. We still need to prove that v has compact support. To do this, we shall first show that each v_k has its support in some compact set $K \subset Q$, where K is independent of k.

Recalling that $P^*v_k \to P^*v$ in $\mathcal{D}(Q)$, we have that there is a compact set K' (see [8, Thm. 1.3.1, p. 5]) such that

$$P^*v_k = 0 \quad \text{in } Q - K' \tag{6.1}$$

for all v_k. We proceed now as in [11, p. 211]. Note that, given any t, by (6.1) and [11, Thm. 1.2, p. 202], each v_k is analytic with respect to x in $\Omega - K'_t$, where $K'_t = \{(x, \tau) \in K' | \tau = t\}$. Further, since each $v_k \in \mathcal{D}(Q)$, each v_k is zero in a neighborhood of ∂Q. Thus using the analyticity result above, there are neighborhoods N_0, N_T and N_Σ, respectively, of $t = 0$, $t = T$, and of Σ, where $v_k = 0$. Since the region of analyticity in x depends only on K', the neighborhoods N_0, N_T and N_Σ also depend only on K' and not on k. Thus it follows that each v_k has its support contained in some compact subset K of Q, which is the same for each k.

Recall that $v_k \to v$ in X. By the nature of the topology on X, then $v_k \to v$ in $\mathcal{D}(\bar{Q})$. Thus since each $v_k = 0$ in $Q - K$, the limit, i.e., v, is also zero there. Hence v has compact support and the lemma is proved.

Proof of Theorem 6.1. Let us demonstrate first that $P^*\mathcal{D}(Q)$ is closed in $\mathcal{D}(Q)$ with respect to the $\mathcal{D}'(Q)$ topology.

Note that for some $\varepsilon > 0$, by [3, pp. 197–201], we can extend the operators A and A^* to a domain $Q_\varepsilon = \Omega_\varepsilon \times (-\varepsilon, T + \varepsilon)$, such that the operators retain the same properties as before. In the above, Ω_ε is such that $\bar{\Omega} \subset \Omega_\varepsilon$. Call the extensions A_ε and A^*. Take $P_\varepsilon = A_\varepsilon + D_t$ and $P_\varepsilon^* = A_\varepsilon^* - D_t$. These operators will have the same properties as P and P^*. In particular, coefficients will be in $\mathcal{D}_{2m,1}(\bar{Q}_\varepsilon)$.

We need to show that if L is any continuous antilinear form on $\mathcal{D}'(Q)$ and if $L(u) = 0$ $\forall u \in S \cap \mathcal{E}(Q)$, then $L(w) = 0$ $\forall w \in S$. Since $\mathcal{D}(Q)$ is reflexive, we can write that

$$L(u) = \langle v, \bar{u} \rangle = \int_Q f(x, t)\bar{u}(x, t)\, dx\, dt = 0 \quad \forall u \in S \cap \mathcal{D}(\bar{Q}) \tag{6.2}$$

for some $f \in \mathcal{D}(Q)$, where $\langle \cdot\,, \cdot \rangle$ represents the duality between $\mathcal{D}(Q)$ and $\mathcal{D}'(Q)$. Let u_ε be a solution of

$$P_\varepsilon u_\varepsilon = 0 \quad \text{on } Q_\varepsilon,$$

so that $u_\varepsilon|_Q \in S \cap \mathcal{D}(\bar{Q})$. Let f_ε be the extension of f to all of Q_ε by taking it to be zero outside of Q. Then

$$\int_{Q_\varepsilon} f_\varepsilon(x, t)\bar{u}_\varepsilon(x, t)\, dx\, dt = 0, \tag{6.3}$$

using (6.2).

Take $P_\varepsilon^{-1}(0) = \{u \in \mathcal{D}'(Q_\varepsilon) | P_\varepsilon u = 0\}$. Lemma 6.1 applies for P_ε^* and Q_ε. Thus $P_\varepsilon^*\mathcal{D}(Q_\varepsilon)$ is closed, and this implies that $P_\varepsilon^*\mathcal{D}(Q_\varepsilon)$ and the polar of $P_\varepsilon^{-1}(0)$ are equal. By (6.3), f_ε falls in this polar. Hence $f_\varepsilon \in P_\varepsilon^*\mathcal{D}(Q_\varepsilon)$, and there exists a

$v \in \mathcal{D}(Q_\varepsilon)$ such that

$$P_\varepsilon^* v = f_\varepsilon \quad \text{in } Q_\varepsilon. \tag{6.4}$$

Let ∂Q_ε be the entire boundary of Q_ε and if Γ_ε is the boundary of Ω_ε, take $\Sigma_\varepsilon = \Gamma_\varepsilon \times (-\varepsilon, T+\varepsilon)$. Since $v \in \mathcal{D}(Q_\varepsilon)$, there exists a neighborhood of ∂Q_ε in which $v = 0$. Also, by its definition, f_ε has compact support in Q. Thus the same argument that was used in Lemma 6.1 to show that each of the v_k had support in $Q - K$ can be repeated here to show that v has compact support in Q (see also [11, p. 211]). (Here, the neighborhoods N_0, N_T and N_ε are replaced by neighborhoods $N_{-\varepsilon}$, $N_{\varepsilon+T}$ and $N_{\varepsilon,\Sigma}$, respectively, of $t = -\varepsilon$, $t = T+\varepsilon$ and Σ_ε. These can be chosen so that their intersections with $\bar{Q}$ are, respectively, neighborhoods of $t = 0$, $t = T$ and Σ.) Hence $v \in \mathcal{D}(Q)$, and

$$P^* v = f \quad \text{in } Q. \tag{6.5}$$

Now let w be any solution in S. Applying (6.5), we have that

$$L(w) = \langle f, \bar{w} \rangle = \langle P^* v, \bar{w} \rangle = \langle f, \overline{Pw} \rangle = 0 \quad \forall w \in S,$$

since $Pw = 0$. Thus $S \cap C^\infty(\bar{Q})$ is dense in S with respect to the $\mathcal{D}'(Q)$ topology.

Next let us show that the topologies induced on S by $\mathcal{D}'(Q)$ and $C^\infty(Q)$ are the same. Let S_1 denote the space S with the $C^\infty(Q)$ topology, and S_2 the same with the $\mathcal{D}'(Q)$ topology. Let $I_{12}: S_1 \to S_2$ and $I_{21}: S_2 \to S_1$ be the respective injection mappings. Let G_{12} and G_{21} be the respective graphs of I_{12} and I_{21}.

Take $\{(u_k, I_{12}u_k)\} = \{(u_k, u_k)\}$ to be a convergent subsequence in G_{12}. Then $u_k \to u$ in S_1 and $u_k \to v$ in S_2 for some $(u, v) \in S_1 \times S_2$. But $u_k \to u$ in S_1 implies $u_k \to u$ in S_2. Hence $u = v$ and $(u, v) = (u, u) \in G_{12}$, and G_{12} is closed in $S_1 \times S_2$. Thus I_{12} is a closed linear operator [18, Def. 2, p. 77].

By the closed graph theorem for Fréchet spaces [18, Thm. 1, p. 79], then, I_{12} is continuous. A similar argument holds for I_{21}. Thus S_1 and S_2 are homeomorphic and the $\mathcal{D}'(Q)$ and $\mathcal{E}(Q)$ topologies are equivalent. Hence the theorem is proved.

7. The representations. We are now ready for the representations.

THEOREM 7.1. *Let $u \in \mathcal{D}'(Q)$ be any solution in $\mathbb{P}$. Then for $(x, t) \in Q$, u has the representation*

$$u(x, t) = \langle \sigma u, \gamma G(x, t; \cdot, \cdot) \rangle, \tag{7.1}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality between V' and V.

Proof. For any $u \in \mathbb{P}$, by Theorem 6.1, there exists a sequence $\{u_k\}$ in $\mathbb{P} \cap \mathcal{D}(\bar{Q})$ such that $u_k \to u$ in $\mathcal{E}(Q)$. Using (5.3) and the fact that $\gamma G \in V$ by Lemma 5.1, then

$$\begin{aligned} u_k(x, t) &= \sum_{j=0}^{m-1} \int_\Sigma B_j u_k(\xi, \tau) T_j G(x, t; \xi, \tau)\, d\sigma + \int_\Omega u_k(\xi, 0) G(x, t; \xi, 0)\, d\xi \\ &= \langle \sigma u_k, \gamma G(x, t; \cdot, \cdot) \rangle. \end{aligned} \tag{7.2}$$

Let $k \to \infty$. The left-hand side of (7.2) converges to $u(x, t)$ for all $(x, t) \in Q$. Since $u_k \to u$ in $\mathcal{E}(Q)$, the same is true in Y relative to the $\sigma(Y, Y')$ topology. By

the continuity of σ (Lemma 3.1), then

$$\langle \sigma u_k, \gamma G(x, t; \cdot, \cdot)\rangle \to \langle \sigma u, \gamma G(x, t; \cdot, \cdot)\rangle$$

for all $(x, t) \in Q$. Thus (7.1) results.

As a corollary, we have the following representation which is suitable for practical applications.

COROLLARY 7.1. *Let $u \in \mathscr{D}'(Q)$ be any solution in $\mathbb{P}^*$* (i.e., *a solution with trace in V^**). *Then there exist measures μ_k on Ω satisfying* (4.2) *and sets of measures $\mu_{kl}^{(j)}$ on Σ, $j = 0, 1, \cdots, m-1$, each with support bounded at the right of* $(0, T)$ *and satisfying* (4.3), (4.4) *and* (4.5), *such that for any $(x, t) \in Q$, u can be represented as*

$$\begin{aligned} u(x, t) &= \sum_{k=0}^{\infty} \int_{\Omega} A^{*k}(0) G(x, t; \xi, 0)\, d\mu_k(\xi) \\ &\quad + \sum_{j=0}^{m-1} \sum_{k,l=0}^{\infty} \int_{\Sigma} \Delta_{\Gamma}^{k} D_{\tau}^{l} T_j G(x, t; \xi, \tau)\, d\mu_{kl}^{(j)}(\xi, \tau). \end{aligned} \tag{7.3}$$

In the above, the notation is the same as that used previously and the operations $A^*(0)$, Δ_Γ, D_τ, T_j are performed with respect to (ξ, τ).

Proof. In (4.1), take $F = \sigma(u)$, $u = G(x, t; \cdot, \cdot)$. Then $\langle \sigma u, \gamma G(x, t; \cdot, \cdot)\rangle$ of (7.1) is given by the right-hand side of (4.1). Combining this representation with (7.1) then gives (7.3).

REFERENCES

[1] P. L. BUTZER AND H. BERENS, *Semi-Groups of Operators and Approximation*, Springer-Verlag, New York, 1967.

[2] J. R. CANNON AND C. K. MILLER, *Some problems in numerical analytic continuation*, SIAM J. Numer. Anal., 2 (1965), pp. 87–98.

[3] L. CARLESON, *On universal moment problems*, Math. Scand., 9 (1961), pp. 197–206.

[4] J. DOUGLAS, JR., *A numerical method for analytic continuation*, Boundary Problems in Differential Equations, R. E. Langer, ed., University of Wisconsin Press, Madison, 1960, pp. 179–189.

[5] ———, *Approximate continuation of harmonic and parabolic functions*, Numerical Solution of Partial Differential Equations, J. H. Bramble, ed., Academic Press, New York, 1966, pp. 353–364.

[6] S. HELGASON, *Differential Geometry and Symmetric Spaces*, Academic Press, New York, 1962.

[7] K. HOFFMAN, *Banach Spaces of Analytic Functions*, Prentice-Hall, Englewood Cliffs, N.J., 1962.

[8] L. HÖRMANDER, *Linear Partial Differential Operators*, Springer-Verlag, Berlin, 1964.

[9] G. JOHNSON, JR., *Harmonic functions on the unit disc. I*, Illinois J. Math., 12 (1968), pp. 366–385.

[10] J. L. LIONS AND E. MAGENES, *Non-Homogeneous Boundary Value Problems and Applications*, vol. I, Springer-Verlag, New York, 1973.

[11] ———, *Non-Homogeneous Boundary Value Problems and Applications*, vol. III, Springer-Verlag, New York, 1973.

[12] H. D. MEYER, *A representation for a distributional solution of the heat equation*, this Journal, 5 (1974), pp. 708–722.

[13] ———, *Half-plane representations and harmonic continuation*, this Journal, 7 (1976), pp. 713–721.

[14] I. G. PETROWSKI, *On the Cauchy problem for systems of linear partial differential equations in a domain of non-analytic functions*, Bjull. Moskov, Gos. Univ. Ser. A., 1 (1938), no. 7, pp. 1–74.

[15] R. SAYLOR, *A generalized boundary-integral representation for solutions of elliptic partial differential equations*, Doctoral thesis, Rice Univ., Houston, Texas, 1966.

[16] ———, *Boundary values of solutions of elliptic equations*, Indiana Univ. Math. J., 24 (1975), pp. 907–913.

[17] F. TRÈVES, *Topological Vector Spaces, Distributions and Kernels*, Academic Press, New York, 1967.

[18] K. YOSIDA, *Functional Analysis*, Academic Press, New York, 1965.

SIAM J. MATH. ANAL.
Vol. 8, No. 2, April 1977

SINGULAR PERTURBATIONS IN THE FIRST BOUNDARY VALUE PROBLEM FOR PARABOLIC EQUATIONS*

CHARLES J. HOLLAND†

Abstract. The first boundary value problem for singularly perturbed semilinear parabolic equations is considered. The regular and ordinary boundary layer expansions are derived using probabilistic methods.

1. Introduction. In this paper using estimates from the theory of stochastic differential equations we derive the regular and ordinary boundary layer expansions for a class of singularly perturbed semilinear parabolic partial differential equations in two independent variables. The approach taken here was also used in [4] to establish the regular and ordinary boundary layer expansions for semilinear elliptic equations. Previous treatments of singularly perturbed parabolic equations include the work of Aronson [1] in which the zeroth order expansion for linear equations was established.

2. Development. Let $x=(x_1, x_2)\in\mathbb{R}^2$. For $\varepsilon>0$ consider the equation

$$\mathscr{L}^\varepsilon\phi=\varepsilon\phi_{x_1x_1}+a(x)\phi_{x_1}-b(x)\phi_{x_2}+F(x,\phi)=0 \tag{1}$$

in the open rectangle $R=(0,1)\times(0,1)$ with boundary data $\phi=\Lambda$ along the bottom $S_1=\{(x_1,0):0\leqq x_1\leqq 1\}$, and lateral sides $S_2=\{(0,x_2):0\leqq x_2\leqq 1\}$, $S_3=\{(1,x_2):0\leqq x_2\leqq 1\}$. Denote the top of the rectangle by $S_4=\{(x_1,1):0\leqq x_1\leqq 1\}$. The expansions will be established on certain subsets of $\bar{R}$ ($\bar{}$ denotes closure) in which the solutions to (1) are uniformly bounded and the method of characteristics yield a C^∞ solution to (1) when $\varepsilon=0$. The characteristics of (1) are solutions of the differential equations

$$x_1'=a(x), \qquad x_2'=-b(x). \tag{2}$$

The definitions are made to allow for a precise statement of the theorem. Definition 1 is a modification of the definition of regular multilateral given in [1].

DEFINITION 1. Let I be a closed subarc of $S_2\cup S_3\cup S_4$ such that no characteristic of (2) is tangent to S_2, S_3, or S_4 and such that every characteristic which starts on I enters R, for increasing t, and first leaves R via a closed subarc $J\subset S_1\cup S_2\cup S_3$. In addition let no characteristic starting on I be tangent to S_1, S_2, S_3 at J. The closed region bounded by I, J and the characteristics joining their endpoints is a *regular multilateral D.*

DEFINITION 2. For r positive, let

$$V_r=\{x:x\in D \text{ and dist } (x, I\cap S_2)\leqq r\},$$

$$W_r=\{x:x\in D \text{ and dist } (x, I\cap S_3)\leqq r\},$$

and let $M_r=D-(W_r\cup V_r)$.

* Received by the editors, January 17, 1975.

† Department of Mathematics, Purdue University, West Lafayette, Indiana 47907.

DEFINITION 3. For each $\varepsilon > 0$, let $\mathscr{S}_\varepsilon \subset \mathbb{R}^2$ and let Z^ε be defined on $\mathscr{S}_\varepsilon$. Then Z^ε *converges uniformly to* Z *on* $\mathscr{S}_\varepsilon$ if, for any $\mu > 0$, there exists $\gamma > 0$ such that if $x \in \mathscr{S}_\varepsilon$ and $\varepsilon < \gamma$, then

$$|Z^\varepsilon(x) - Z(x)| < \mu.$$

We shall use three regular multilaterals D, D', D''. Let us use the notation that if w is a quantity defined with respect to D, then $w'(w'')$ is the corresponding quantity defined with respect to $D'(D'')$.

THEOREM 1. *Let there exist positive constants* ε_0, K^*, q, *and* m *with* $0 < m < \frac{1}{2}$ *such that the following hold*:

(A1) D, D', D'' *are regular multilaterals with* $I'' \subset$ *interior* I', $I' \subset$ *interior* I.

(A2) a, b *are* C^∞ *functions on* D, F *is a* C^∞ *function on* $D \times (-\infty, \infty)$.

(A3) $b > q > 0$ *on* D.

(A4) *The method of characteristics defines a* C^∞ *solution* ϕ^0 *to* (1) *with* $\varepsilon = 0$ *on* D *taking boundary data* $\phi^0 = \Lambda$ *on* J.

(A5) *For* $0 < \varepsilon < \varepsilon_0$ *there exists a* C^2 *solution* ϕ^ε *to* (1) *on* R *with* $\phi^\varepsilon = \Lambda$ *on* $S_1 \cup S_2 \cup S_3$ *satisfying* $|\phi^\varepsilon| \leqq K^*$ *on* D.

(A6) $\delta(\varepsilon) = \varepsilon^m$.

Then there exists functions $\theta_1, \theta_2, \cdots$ *bounded on* D *and functions* $\psi^0, \chi_1, \chi_2, \cdots$ *bounded on* $[0, \infty) \times I' \cap S_2$ *and* $\tilde{\psi}^0, \tilde{\chi}, \tilde{\chi}_2, \cdots$ *bounded on* $[0, \infty) \times I' \cap S_3$ *such that for any positive integer* n,

$$\text{(i)}\quad \phi^\varepsilon(x) = \phi^0(x) + \psi^0\left(\frac{x_1}{\varepsilon}, x_2\right) + \tilde{\psi}^0\left(\frac{1-x_1}{\varepsilon}, x_2\right) + \sum_{j=1}^{n} \left[\theta_j(x) + \chi_j\left(\frac{x_1}{\varepsilon}, x_2\right) + \tilde{\chi}_j\left(\frac{1-x_1}{\varepsilon}, x_2\right)\right]\varepsilon^j + o(\varepsilon^n) \tag{3}$$

uniformly for x *on* D''.

$$\text{(ii)}\quad \phi^\varepsilon(x) = \phi^0(x) + \sum_{j=1}^{n} \theta_j(x)\varepsilon^j + o(\varepsilon^n) \tag{4}$$

uniformly for x *on* $M''_{\delta(\varepsilon)}$.

$$\text{(iii)}\quad \phi^\varepsilon(x) = \phi^0(x) + \psi^0\left(\frac{x_1}{\varepsilon}, x_2\right) + \sum_{j=1}^{n} \left[\theta_j(x) + \chi_j\left(\frac{x_1}{\varepsilon}, x_2\right)\right]\varepsilon^j + o(\varepsilon^n) \tag{5}$$

uniformly for x *on* $M''_{\delta(\varepsilon)} \cup V''_{\delta(\varepsilon)}$.

$$\text{(iv)}\quad \phi^\varepsilon(x) = \phi^0(x) + \tilde{\psi}^0\left(\frac{1-x_1}{\varepsilon}, x_2\right) + \sum_{j=1}^{n} \left[\theta_j(x) + \tilde{\chi}_j\left(\frac{1-x_1}{\varepsilon}, x_2\right)\right]\varepsilon^j + o(\varepsilon^n) \tag{6}$$

uniformly for x *on* $M''_{\delta(\varepsilon)} \cup W''_{\delta(\varepsilon)}$.

If $D \cap S_2$ is the empty set, then the functions $\psi^0, \chi_1, \chi_2, \cdots$ are identically zero. If $D \cap S_3$ is the empty set, then the functions $\tilde{\psi}^0, \tilde{\chi}_1, \tilde{\chi}_2, \cdots$ are identically zero.

Remarks. Equation (4) is the regular expansion in the theory of singular perturbations. The functions $\psi^0, \chi_1, \chi_2, \cdots$ are a result of the ordinary boundary layer along $D \cap S_2$; the functions $\tilde{\psi}^0, \tilde{\chi}_1, \tilde{\chi}_2, \cdots$ are a result of the ordinary boundary layer along $D \cap S_3$. Equations for the functions appearing in (3) are described below.

The coefficient θ_k in (4) satisfies the equation to make the coefficient of ε^k identically zero in the formal expansion of (1) in powers of ε. By direct calculation θ_k satisfies

$$\mathscr{L}^0(\theta_k) + F_\phi(x, \phi^0(x))\theta_k + \Gamma_k + (\theta_{k-1})_{x_1 x_1} = 0, \tag{7}$$

$\theta_0 = \phi^0$, with boundary data $\theta_k = 0$ on J. $\Gamma_1 = 0$ and in general Γ_k is a polynomial in $\theta_1, \cdots, \theta_{k-1}$ of degree k, with coefficients $F_{\phi\phi}, F_{\phi\phi\phi}, \cdots$ evaluated at $(x, \phi^0(x))$. If $F(x, \phi)$ is linear in ϕ, as is the case in [1], then $\Gamma_k = 0$ for any k.

Equations for the functions $\psi^0, \chi_1, \chi_2, \cdots$ are found by formally substituting the expansion (5) into (1). This technique for elliptic equations was demonstrated in Appendix A of [4]. One finds that $\psi^0 = \psi^0(x)$ satisfies on $[0, \infty) \times I \cap S_2$

$$\psi^0_{x_1 x_1} + a(0, x_2)\psi^0_{x_1} = 0 \tag{8}$$

with boundary conditions

$$\psi^0(0, x_2) = \Lambda(0, x_2) - \phi^0(0, x_2),$$
$$\psi^0(\infty, x_2) = 0.$$

Note that for fixed x_2, (8) is an ordinary differential equation in the variable x_1 with solution

$$\psi^0(x_1, x_2) = [\Lambda(0, x_2) - \phi^0(0, x_2)] \cdot \exp[-a(0, x_2)x_1]. \tag{9}$$

Henceforth write θ for θ_1 and χ for χ_1. Then $\chi = \chi(x)$ satisfies

$$\chi_{x_1 x_1} + a(0, x_2)\chi_{x_1} + \left[\int_0^1 F_\phi(0, x_2, \phi^0(0, x_2) + \lambda\psi^0(x))\, d\lambda\right]\psi^0 + a_{x_1}(0, x_2)x_1\psi^0_{x_1} - b(0, x_2)\psi^0_{x_2} = 0 \tag{10}$$

with boundary conditions

$$\chi(0, x_2) = -\theta(0, x_2),$$
$$\chi(\infty, x_2) = 0.$$

It is easy to show that the functions ψ^0, χ_k satisfy an exponential decay in x_1 as $x_1 \to +\infty$.

The equations for $\tilde{\psi}^0, \tilde{\chi}_k$ are found similarly.

Proof of the theorem. Let the positive integer n in (3)–(6) be fixed. First, the expansion (4) is derived on $M'_{\delta(\varepsilon)}$. Let $\tilde{Z}_{-1}, \tilde{Z}_0, \tilde{Z}_1, \cdots, \tilde{Z}_n$ be regular multilaterals with the properties that $D = \tilde{Z}_{-1} \supset \tilde{Z}_0 \supset \cdots \tilde{Z}_n = D'$ and there exists some positive constant $\tilde{p}$ such that distance $(\tilde{Z}_k, D - \tilde{Z}_{k-1}) > \tilde{p}$, $k = 0, 1, \cdots n$. Define $Z^\varepsilon_k = \tilde{Z}_k - V_{[(k+1)\delta(\varepsilon)/(n+1)]}$ and $\theta^\varepsilon_1 = \varepsilon^{-1}(\phi^\varepsilon - \phi^0)$, $\theta^\varepsilon_k = \varepsilon^{-1}(\theta^\varepsilon_{k-1} - \theta_{k-1})$,

$k = 2, \cdots, n$. Then to prove (4) we show that $\phi^\varepsilon \to \phi^0$ uniformly on Z_0^ε and $\theta_k^\varepsilon \to \theta_k$ uniformly on Z_k^ε, $k = 1, 2, \cdots, n$. We note that θ_k^ε satisfies

$$\mathscr{L}^\varepsilon(\theta_k^\varepsilon) + A^\varepsilon \theta_k^\varepsilon + \Gamma_k^\varepsilon + (\theta_{k-1})_{x_1 x_1} = 0. \tag{11}$$

Here

$$A_k^\varepsilon = \int_0^1 F_\phi(x, \phi^0 + \lambda(\phi^\varepsilon - \phi^0))\, d\lambda, \qquad \Gamma_1^\varepsilon = 0,$$

$$\Gamma_2^\varepsilon = \theta_1 \theta_1^\varepsilon \int_0^1 \int_0^1 F_{\phi\phi}(x, \phi^0 + \lambda\mu(\phi^\varepsilon - \phi^0))\, d\mu\, d\lambda, \cdots.$$

We now make a probabilistic connection between θ_k^ε and θ_k. Redefine the functions a, b outside D so that there exists a constant M which is both a bound for $|a|$, $|b|$ and a Lipschitz constant for a, b on $\mathbb{R}^2$. Let ξ_x^ε be the solution to the Ito stochastic differential equation

$$\begin{aligned} d\xi_1 &= a(\xi)\, dt, \\ d\xi_2 &= -b(\xi)\, dt + (2\varepsilon)^{1/2}\, dw \end{aligned} \tag{12}$$

with initial condition $\xi_x^\varepsilon(0) = x$. An application of Gronwall's inequality and standard estimates on Brownian motion yields

$$\begin{aligned} &\Pr\{\|\xi_x^\varepsilon - \xi_x^0\|_t \geqq \delta(\varepsilon)/2(n+1)\} \\ &\qquad \leqq \frac{4(n+1)e^{Mt}}{\delta(\varepsilon)} \left(\frac{\varepsilon t}{\pi}\right)^{1/2} \cdot \exp\left(\frac{-\delta(\varepsilon)^2\, e^{-2Mt}}{4\varepsilon t(n+1)^2}\right) \end{aligned} \tag{13}$$

where $\|\xi_x^\varepsilon - \xi_x^0\|_t = \sup_{0 \leqq t' \leqq t} |\xi_x^\varepsilon(t') - \xi_x^0(t')|$.

For $x \in Z_0^\varepsilon$ let τ_x^ε denote the exit time of ξ_x^ε from the interior of Z_{-1}^ε and let $\gamma_x^\varepsilon = \min(\tau_x^\varepsilon, (1/q) + 1)$. From the Ito stochastic differential rule one obtains

$$\begin{aligned} \theta_1^\varepsilon(x) = E\Big\{ &\int_0^{\gamma_x^\varepsilon} D_x^\varepsilon(t)(\phi^0)_{x_1 x_1}(\xi_x^\varepsilon(t))\, dt \\ &+ D_x^\varepsilon(\gamma_x^\varepsilon)\theta_1^\varepsilon(\xi_x^\varepsilon(\gamma_x^\varepsilon)) \Big\}. \end{aligned} \tag{14}$$

Let $t_1 \geqq 1/q + 1$ be fixed. For all sufficiently small ε, if $x \in Z_0^\varepsilon$ and $\|\xi_x^\varepsilon - \xi_x^0\|_{t_1} < \delta(\varepsilon)/2(n+1)$, then $\xi_x^\varepsilon(\tau_x^\varepsilon) \in J$. Hence, from (13) since $\delta(\varepsilon) = \varepsilon^m$ with $0 < m < \frac{1}{2}$,

$$|\tau_x^\varepsilon - \tau_x^0| \to 0, \quad \|\xi_x^\varepsilon - \xi_x^0\|_{t_1} \to 0, \quad |\xi_x^\varepsilon(\tau_x^\varepsilon) - \xi_x^0(\tau_x^0)| \to 0 \tag{15}$$

in probability uniformly on Z_0^ε. Since if $\xi_x^\varepsilon(\gamma_x^\varepsilon) \in J$, $\theta_1^\varepsilon(\xi_x^\varepsilon(\gamma_x^\varepsilon)) = 0$ and $|\theta_1^\varepsilon(\xi_x^\varepsilon(\gamma_x^\varepsilon))| \leqq C\varepsilon^{-1}$ otherwise, then using the estimates (13) and (15), one obtains that $\lim_{\varepsilon \to o} |\phi^\varepsilon - \phi^0| = \lim_{\varepsilon \to 0} |\varepsilon\theta_1^\varepsilon| = 0$ uniformly on Z_0^ε.

Now for each $k = 1, 2, \cdots, n$ and $x \in Z_k^\varepsilon$ let τ_x^ε be the exit time of ξ_x^ε from Z_{k-1}^ε. Then applying the Ito stochastic differential rule with $\gamma_x^\varepsilon = \min(\tau_x^\varepsilon, (1/q) + 1)$, we obtain

$$\begin{aligned} \theta_k^\varepsilon(x) = E\Big\{ &\int_0^{\gamma_x^\varepsilon} D_x^\varepsilon(t)[\Gamma_k^\varepsilon(\xi_x^\varepsilon(t)) + (\theta_{k-1})_{x_1 x_1}(\xi_x^\varepsilon(t))]\, dt \\ &+ D_x^\varepsilon(\gamma_x^\varepsilon)\theta_k^\varepsilon(\xi_x^\varepsilon(\gamma_x^\varepsilon)) \Big\} \end{aligned} \tag{16}$$

with $D_x^\varepsilon(t) = \exp \int_0^t A^\varepsilon(\xi_x^\varepsilon(s))\, ds$. Similarly the method of characteristics yields

$$\theta_k(x) = \int_0^{\tau_x^\varepsilon} D_x(t)[\Gamma_k^0(\xi_x^0(t)) + (\theta_{k-1})_{x_1x_1}(\xi_x^0(t))]\, dt \tag{17}$$

since $\theta_k = 0$ when $x \in J$. When $\xi_x^\varepsilon(\gamma_x^\varepsilon) \in J$, $\theta_k^\varepsilon(\xi_x^\varepsilon(\gamma_x^\varepsilon)) = 0$ and $|\theta_k^\varepsilon(\xi_x^\varepsilon(\gamma_x^\varepsilon))| \leqq C\varepsilon^{-k}$ otherwise since ϕ^ε is bounded on D by assumption. Using the estimate (13) we obtain $E_x\{D_x^\varepsilon(\gamma_x^\varepsilon)\theta_k^\varepsilon(\xi_k^\varepsilon(\gamma_x^\varepsilon))\} \to 0$ uniformly on Z_k^ε. By induction one shows that $\Gamma_k^\varepsilon \to \Gamma_k$ uniformly on Z_{k-1}^ε. Using (13) and the definition of τ_x^ε one obtains that (15) is valid uniformly for $x \in Z_k^\varepsilon$. Therefore $\theta_k^\varepsilon \to \theta_k$ uniformly on Z_k^ε. This proves (4).

The verification of the boundary layer expansions (5) and (6) is similar to the proof of the corresponding boundary layer expansions for elliptic equations in [4]. We only prove (5) with $n = 0$ to illustrate the necessary changes from the proof in [4].

Choose ε^* sufficiently small so that for $\varepsilon < \varepsilon^*$

$$\left[\frac{k_1' + k_1''}{2}, \frac{k_2' + k_2''}{2}\right] \times [0, \delta(\varepsilon)] \supset V''_{\delta(\varepsilon)},$$

and

$$\left[\frac{2k_1' + k_1''}{3}, \frac{k_2'' + 2k_2'}{3}\right] \times [0, \delta(\varepsilon)] \subset V'_{\delta(\varepsilon)}.$$

Define

$$Q^\varepsilon = [(k_1' + k_1'')/2, (k_2' + k_2'')/2] \times [0, \delta(\varepsilon)\varepsilon^{-1}],$$
$$\tilde{Q}^\varepsilon = [(2k_1' + k_1'')/3, (k_2'' + 2k_2')/3] \times [0, \delta(\varepsilon)\varepsilon^{-1}],$$

and $\psi^\varepsilon(x) = \phi^\varepsilon(\varepsilon x_1, x_2) - \phi^0(\varepsilon x_1, x_2)$. Then to establish (5) with $n = 0$ it suffices to show that $\psi^\varepsilon \to \psi^0$ uniformly on Q^ε.

Let $\Phi = \Phi(x_1, x_2)$. Define the operator $\mathcal{M}^\varepsilon$ by

$$\mathcal{M}^\varepsilon \Phi = \Phi_{x_1x_1} + a(\varepsilon x_1, x_2)\Phi_{x_1} - \varepsilon b(\varepsilon x_1, x_2)\Phi_{x_2}.$$

Then ψ^ε satisfies the equation

$$\mathcal{M}^\varepsilon \psi^\varepsilon + \varepsilon V(x, \varepsilon) = 0 \tag{18}$$

with

$$V(x, \varepsilon) = [-F(\varepsilon x_1, x_2, \phi^\varepsilon(\varepsilon x_1, x_2)) + F(\varepsilon x_1, x_2, \phi^0(\varepsilon x_1, x_2)) + \varepsilon \phi^0_{x_1x_1}(\varepsilon x_1, x_2)].$$

We may also rewrite the equation (8) for ψ^0 in the form

$$\mathcal{M}^\varepsilon \psi^0 + \varepsilon H(x, \varepsilon) = 0 \tag{19}$$

with

$$H(x, \varepsilon) = b(\varepsilon x_1, x_2)\psi^0_{x_2}(x) - \left(\int_0^1 a_{x_1}(\lambda \varepsilon x_1, x_2)\, d\lambda\right) \cdot x_1 \psi^0_{x_1}(x).$$

We now make the probabilistic connection between ψ^ε and ψ^0. Consider for $x \in Q^\varepsilon$ the stochastic differential equation

$$d\eta^\varepsilon = \begin{pmatrix} a(\varepsilon\eta_1, \eta_2) \\ -\varepsilon b(\varepsilon\eta_1, \eta_2) \end{pmatrix} dt + \begin{pmatrix} 1 \\ 0 \end{pmatrix} dw \tag{20}$$

with initial condition $\eta^\varepsilon(0) = x$. By assumption the characteristics intersect I' nontangentially; hence there exists $\varepsilon^{**} < \varepsilon^*$, $\varepsilon^{**} > 0$ and a constant P so that $a(\varepsilon x_1, x_2) > P$ for $x \in Q^\varepsilon$. Let $d < (k_1'' - k_1')/6$. Then if $t < d/\varepsilon M$, η_x^ε cannot have exited through the lower boundary of $\tilde{Q}^\varepsilon$. Consider the problem for $\varepsilon < \bar{\varepsilon}$ where $\bar{\varepsilon} < \varepsilon^{**}$ and $(2\delta(\bar{\varepsilon}))/(\bar{\varepsilon}P) < d/(\bar{\varepsilon}M)$. Define τ_x^ε to be the first time $t \geqq 0$ that $\eta_1^\varepsilon(t) = 0$, or $\delta(\varepsilon)/\varepsilon$, and $\gamma_x^\varepsilon = \min[\tau_x^\varepsilon, 2\delta(\varepsilon)(\varepsilon P)^{-1}]$. Then

$$\psi^\varepsilon(x) = E_x\left[\int_0^{\gamma_x^\varepsilon} \varepsilon V(\eta^\varepsilon(t), \varepsilon)\, dt + \psi^\varepsilon(\eta^\varepsilon(\gamma_*^\varepsilon))\right] \tag{21}$$

and

$$\psi^0(x) = E_x\left[\int_0^{\gamma_x^\varepsilon} \varepsilon H(\eta^\varepsilon(t), \varepsilon)\, dt + \psi^0(\eta^\varepsilon(\gamma_x^\varepsilon))\right]. \tag{22}$$

We show that $\psi^\varepsilon \to \psi^0$ uniformly on Q^ε through use of the representations (21) and (22). Recall the form of $V(x, \varepsilon)$, $H(x, \varepsilon)$ and ψ^0. Since $\gamma_x^\varepsilon \leqq 2\delta(\varepsilon)/(\varepsilon P)$ and the functions $H(x, \varepsilon)$, $V(x, \varepsilon)$ are uniformly bounded on $Q^\varepsilon \times [0, \bar{\varepsilon}]$, then the expectations of the corresponding integrals in (21) and (22) converge uniformly to zero on Q^ε. We now need only show that

$$E|\psi^\varepsilon(\eta^\varepsilon(\gamma_x^\varepsilon)) - \psi^0(\eta^\varepsilon(\gamma_x^\varepsilon))| \to 0$$

uniformly for $x \in Q^\varepsilon$. Since ψ^ε, ψ^0 are uniformly bounded on $\tilde{Q}^\varepsilon$, $\psi^\varepsilon = \psi^0$ along I',

$$\left|\psi^\varepsilon\left(\frac{\delta(\varepsilon)}{\varepsilon}, x_2\right) - \psi^0\left(\frac{\delta(\varepsilon)}{\varepsilon}, x_2\right)\right| \to 0$$

uniformly for $x_2 \in [k_1', k_2']$, then we need only show that

$$\Pr\{\eta_1^\varepsilon(\gamma_x^\varepsilon) \neq 0 \text{ or } \delta(\varepsilon)/\varepsilon\} \to 0 \tag{23}$$

uniformly for $x \in Q^\varepsilon$. The estimate on $\psi^\varepsilon(\delta(\varepsilon)\varepsilon^{-1}, x_2) - \psi^0(\delta(\varepsilon)\varepsilon^{-1}, x_2)$ follows from the exponential decay in x_1 of ψ^0 and the result from the regular expansion (3) that $\psi^\varepsilon(\delta(\varepsilon)\varepsilon^{-1}, x_2) = \varepsilon\theta(\delta(\varepsilon), x_2) + o(\varepsilon)$ as $\varepsilon \to 0$.

We now establish (23). Now

$$\eta_1^\varepsilon(2\delta(\varepsilon)(\varepsilon P)^{-1}) > x_1 + 2\delta(\varepsilon)/\varepsilon + w(2\delta(\varepsilon)(\varepsilon P)^{-1}).$$

If $\eta_1^\varepsilon(\gamma_x^\varepsilon) \neq 0$ or $\delta(\varepsilon)/\varepsilon$, then $w(2\delta(\varepsilon)(\varepsilon P)^{-1}) \leqq -\delta(\varepsilon)\varepsilon^{-1}$. From standard estimates we have that

$$\Pr\{w(2\delta(\varepsilon)(\varepsilon P)^{-1}) \leqq -\delta(\varepsilon)\varepsilon^{-1}\} < 4\left(\frac{\varepsilon}{\pi P \delta(\varepsilon)}\right) \exp(-\delta(\varepsilon)P/(2\varepsilon)).$$

The last inequality implies (23) and therefore the result is proved for $\psi^\varepsilon \to \psi^0$. The higher order expansions follow by induction using the estimates derived above in a manner similar to the corresponding expansion in [4] for the elliptic case.

The boundary layer expansion (6) is derived in a similar manner. The details are omitted. Equation (3) now holds on D'' due to the exponential decay of the boundary layer functions. This completes the proof of the theorem. □

Remark. The assumption of the domain being a rectangle is not as severe as it first seems. Suppose that the domain R has boundary $S_1 \cup S_2 \cup S_3 \cup S_4$ with

$S_1 = \{(x_1, c) : a_1 \leqq x_1 \leqq b_1\}$,

$S_4 = \{(x_1, d) : a_2 \leqq x_1 \leqq b_2\}$, where $c < d$, and

$S_2 = \{(g(x_2), x_2) : c \leqq x_2 \leqq d\}$,

$S_3 = \{(h(x_2), x_2) : c \leqq x_2 \leqq d\}$ where g and h are smooth functions with $g < h$. Then the change of variables

$$(x_1, x_2) \to \left(\frac{x_1 - g(x_2)}{h(x_2) - g(x_2)}, \frac{x_2 - c}{d - c} \right)$$

transforms the region R into the rectangle previously treated. Equation (1) is transformed into a similar equation for which Theorem 1 is valid.

REFERENCES

[1] D. ARONSON, *Linear parabolic differential equations containing a small parameter*, J. Rational Mech. Anal., 5 (1956), pp. 1003–1014.

[2] W. ECKHAUS AND E. M. DE JAGER, *Asymptotic solutions of singular perturbations problems for linear differential equations of elliptic type*, Arch. Rational Mech. Anal., 23 (1966), pp. 26–86.

[3] W. H. FLEMING, *Stochastically perturbed dynamical systems*, Rocky Mountain J. Math., 4 (1974), pp. 407–433.

[4] C. HOLLAND, *Singular perturbation problems in elliptic partial differential equations*, J. Differential Equations, 20 (1976), pp. 248–265.

SIAM J. MATH. ANAL.
Vol. 8, No. 3, May 1977

ON A CONCEPT OF A DERIVATIVE AMONG FUNCTIONS DEFINED ON THE DYADIC FIELD*

JENÖ PAL†

Abstract. In papers [10], [11], P. L. Butzer, H. J. Wagner and F. Pichler examined the concept of a dyadic derivative for functions defined on $[0, \infty)$. In this paper, a corresponding inverse operator is defined for such functions, namely, the dyadic integral operator, and a dyadic calculus is developed. It is shown that the analogue of the fundamental theorems of the calculus holds for dyadic differentiation and integration.

1. Introduction. P. L. Butzer and H. J. Wagner ([1], [2], [3], [10]) introduced a concept of a derivative among real-valued functions defined on the dyadic group and the dyadic field, respectively. They proved among other things that the characters of the group (the Walsh–Paley functions and Walsh functions of continuous index, respectively) are arbitrarily often differentiable. Furthermore, they showed that the inverse operation of the derivation introduced on the dyadic group is a convolution with a certain function $W \in L(0, 1)$. In this paper, we are going to give the inverse operation of the derivation defined on the dyadic field. On the basis of Theorem 4.2 to be established, an explicit procedure of determining the Walsh–Fourier transform of functions in $L^2[0, \infty)$ is to be found in [18].

This paper is connected with a number of results in dyadic analysis based upon the Walsh system achieved so far. In this respect, it must be pointed out that Walsh functions play a dominant role in a series of applications. For the (more theoretical) applications of dyadic analysis to Walsh–Fourier analysis and approximation theory see [1], [2], [3], to dyadic partial differential equations see [10]; in the latter paper a dyadic analogue of the wave equation is solved, the solution of which is interpreted by H. F. Harmuth [17] in a speculative way. For the more practical applications, such as to system theory see [11], [13], [14], to information theory see [15], to hardware see [19], to digital signal processing see [20], to sequency multiplexing of digital signals, two-dimensional sequency filters for TV image processing and to radar see Harmuth [21].

The survey type papers by Gibbs and Ireland [16], Gibbs [23], Harmuth [21], H. Hübner [22] as well as the annual conferences on Walsh functions held in Washington, D.C., Hatfield, England and elsewhere give comprehensive accounts of most of these applications.

2. The essential notions and theorems employed. Let $\mathbb{R}_+$ denote the set of nonnegative real numbers. Arbitrary $x \in \mathbb{R}_+$ is available in dyadic form:

$$x = \sum_{j=-K}^{\infty} x_j / 2^{j+1}, \qquad (x_j \in \{0, 1\}, \quad K \in \mathbb{N} = \{0, 1, 2, \cdots\}). \tag{2.1}$$

If x is not a dyadic rational number, then expression (2.1) is unambiguous; if x is dyadically rational, then we consider the expression in which, from a certain index

* Received by the editors March 25, 1975, and in revised form December 1, 1975.
† Department II of Analysis, Eötvös Loránd University, Budapest, Hungary.

on, all 0's stand. If the expansion of $y \in \mathbb{R}_+$ is $y = \sum_{j=-M}^{\infty} y_j/2^{j+1}$, then let

$$x \dot{+} y := \sum_{j=-L}^{\infty} (|x_j - y_j|)/2^{j+1}, \qquad L = \max(K, M). \tag{2.2}$$

The generalized Walsh–Paley functions introduced by N. J. Fine [5] are defined by the following equality:

$$\psi_y(x) = (-1)^{\sum_{j=K}^{\infty} x_j y_{-(1+j)}}, \qquad (x, y \in \mathbb{R}_+).$$

The following assertions are easily provable (cf. [5]):

1.

$$\psi_y(x) = \psi_x(y), \qquad (x, y \in \mathbb{R}_+). \tag{2.3}$$

2.

$$\psi_y(x) = \bar{\psi}_{[y]}(x)\bar{\psi}_{[x]}(y), \qquad (x, y \in \mathbb{R}_+), \tag{2.4}$$

where $\{\bar{\psi}_n : n \in \mathbb{N}\}$ denotes the Walsh–Paley system, and $[x]$ the integer part of $x \in \mathbb{R}_+$.

3. Mappings $x \mapsto \psi_y(x)$ are the characters of the additive group of the dyadic field, i.e.,

$$\psi_y(x \dot{+} z) = \psi_y(x)\psi_y(z), \qquad (x, y, z \in \mathbb{R}_+). \tag{2.5}$$

Let $L^1(\mathbb{R}_+)$ denote the class of functions absolutely integrable, with the usual norm $\|f\|_1 = \int_0^{\infty} |f(x)|\, dx$. We define the Walsh–Fourier transform $\hat{f}$ of function f as follows (cf. [5]):

$$\hat{f}(y) = \int_0^{\infty} f(x)\psi_y(x)\, dx, \qquad (y \in \mathbb{R}_+). \tag{2.6}$$

The fundamental properties of the Walsh–Fourier transforms are the following:

1. The Walsh–Fourier transform is a bounded linear operator mapping from $L^1(\mathbb{R}_+)$ into $L^{\infty}(\mathbb{R}_+)$, whose norm is 1; namely,

$$\lim_{n\to\infty} \|f - f_n\|_1 = 0 \Rightarrow \lim_{n\to\infty} \hat{f}_n(y) = \hat{f}(y)), \qquad (y \in \mathbb{R}_+). \tag{2.7}$$

2. Let $(\tau_t f)(x) := f(x \dot{+} t)$. Then

$$(\tau_t f)\hat{} = \psi_t \cdot \hat{f} \qquad (f \in L^1(\mathbb{R}_+), \quad t \in \mathbb{R}_+). \tag{2.8}$$

3. The dual of the statement (2.8) also holds true:

$$(f \cdot \psi_t)\hat{} = \tau_t \hat{f}, \qquad (f \in L^1(\mathbb{R}_+), \quad t \in \mathbb{R}_+). \tag{2.9}$$

4. If $f, g \in L^1(\mathbb{R}_+)$, then

$$(f * y)\hat{}(y) = \hat{f}(y)\hat{g}(y) \qquad (y \in \mathbb{R}_+), \tag{2.10}$$

where $(f * g)(x) := \int_0^{\infty} f(u)g(x \dot{+} u)\, du$.

5. Let

$$S_w(f; x) := \int_0^{w} \hat{f}(t)\psi_x(t)\, dt, \qquad (x \in \mathbb{R}_+, \quad f \in L^1(\mathbb{R}_+)).$$

Then

$$\lim_{n\to\infty} \|S_{2^n}(f)-f\|_1 = 0 \qquad \text{(cf. [10])}. \tag{2.11}$$

We remark that we may also interpret the Walsh–Fourier transform of the functions in $L^2(\mathbb{R}_+)$ in the usual way (cf., e.g., [9]). This transformation is a linear isometry mapping from $L^2(\mathbb{R}_+)$ into $L^2(\mathbb{R}_+)$. Moreover, the convolution theorem, the inversion formula and the analogue of the Parseval formula hold true.

3. Definition of the concept of a derivative. Let $X(\mathbb{R}_+)$ comprehensively denote the class of w-continuous functions $f:\mathbb{R}_+\to\mathbb{R}$ (cf. [7]) and the class of functions absolutely integrable on pth power $(1\leqq p\leqq\infty)$, consecutively, with the usual norms. Let $\|\cdot\|_X$ denote these jointly.

DEFINITION 3.1 (cf. [10]). We say function $f\in X(\mathbb{R}_+)$ is *D-X differentiable*, if there exists a function $g\in X(\mathbb{R}_+)$ for which

$$\lim_{m\to\infty} \left\| \frac{1}{2} \sum_{j=-m}^{m} 2^j [f-\tau_{2^{-(j+1)}} f]-g \right\|_X = 0. \tag{3.1}$$

Function g we call the *D-X derivative* of f and denote it by $D^{[1]}f$. The *r-th* $(r\in\mathbb{P})$[1] *derivative of function* $f\in X(\mathbb{R}_+)$ is defined by the equality

$$D^{[r]}f := D^{[1]}(D^{[r-1]}f). \tag{3.2}$$

It is clear that the operation of derivation is a linear operator.

THEOREM 3.1. *The generalized Walsh–Paley functions are D-C differentiable arbitrary often, and*

$$D^{[r]}\psi_y = y^r\psi_y, \qquad (y\in\mathbb{R}_+). \tag{3.3}$$

For the proof the reader is referred to F. Pichler [11]. The proof of the theorem is based upon the identity

$$\frac{1}{2}\sum_{j=-\infty}^{\infty} 2^j[1-\psi_y(2^{-(j+1)})] = y, \qquad (y\in\mathbb{R}_+), \tag{3.4}$$

and on the property (2.5) of ψ_y.

The relationship of the derivation and of the Walsh–Fourier transform is given by the following

THEOREM 3.2.

(a) *If f, $D^{[r]}f\in L^1(\mathbb{R}_+)$, then*

$$(D^{[r]}f)\hat{}\,(y) = y^r\hat{f}(y), \qquad (y\in\mathbb{R}_+). \tag{3.5}$$

For proof see P. L. Butzer, H. J. Wagner in [10].

(b) *If $f\in L^2(\mathbb{R}_+)$ is r times D-L^2 differentiable, then*

$$(D^{[r]}f)\hat{}\,(y) = y^r\hat{f}(y), \qquad (y\in\mathbb{R}_+). \tag{3.6}$$

[1] $\mathbb{P}=\{1,2,3,\cdots\}$.

Proof of part (b) is analogous to that of part (a). One has only to make use of the isometry property of the Walsh–Fourier transform concerning functions in $L^2(\mathbb{R}_+)$, instead of (2.7).

4. The inverse operation of derivation. In this section we shall prove an analogue of Butzer and Wagner's theorem: the inverse operation of derivation is a convolution with a certain function W_a.

Before coming to the actual subject of this section, we are going to state some notions and lemmas, interesting in themselves. Let us introduce the kernel function D_ω, analogous to the Walsh–Dirichlet kernel function, by the following definition:

$$D_\omega(t) := \int_0^\omega \psi_y(t)\,dy, \qquad (t \in [0,\infty)). \tag{4.1}$$

Let us apply for this the transformation as follows, employing (2.4):

$$D_\omega(t) = \sum_{i=0}^{[\omega]-1} \bar{\psi}_i(t) \int_i^{i+1} \bar{\psi}_{[t]}(y)\,dy + \bar{\psi}_{[\omega]}(t) \int_{[\omega]}^{\omega} \bar{\psi}_{[t]}(y)\,dy.$$

With the same method, $D_\omega(t)$ may be written in the form

$$D_\omega(t) = \begin{cases} \bar{\psi}_{[\omega]}(t)\mathcal{F}_{[t]}(\omega - [\omega]), & (t \geqq 1), \\ \bar{D}_{[\omega]}(t) + \bar{\psi}_{[\omega]}(t)\mathcal{F}_{[t]}(\omega - [\omega]), & (0 \leqq t < 1), \end{cases} \tag{4.2}$$

where $\bar{D}_{[\omega]}$ is the Walsh–Dirichlet kernel function of order $[\omega]$. Moreover $\mathcal{F}_l(u) := \int_0^u \bar{\psi}_l(v)\,dv$ is the integral function of the lth Walsh–Paley function. We remark that $\mathcal{F}_{[t]}(\omega) = \mathcal{F}_{[t]}(\omega - [\omega])$ if $t \geqq 1$ and $\mathcal{F}_{[t]}(\omega) = \omega$ if $0 \leqq t < 1$. With the help of this remark, (4.2) may be written as

$$D_\omega(t) = \begin{cases} \bar{\psi}_{[\omega]}(t)\mathcal{F}_{[t]}(\omega), & (t \geqq 1), \\ \bar{D}_{[\omega]}(t) + (\omega - [\omega])\bar{\psi}_{[\omega]}(t), & (0 \leqq t < 1). \end{cases} \tag{4.3}$$

From this expression it is obvious that if $\omega \in \mathbb{N}$, then

$$D_\omega(t) = \begin{cases} 0, & (t \geqq 1), \\ \bar{D}_\omega(t), & (0 \leqq t < 1). \end{cases}$$

Define the kernel function K_ω analogous to the Fejér kernel function by

$$K_\omega(t) := \frac{1}{\omega} \int_0^\omega D_u(t)\,du, \qquad (t \in [0,\infty)). \tag{4.4}$$

This is affected by the following

LEMMA 4.1. $\|K_\omega\|_1 = O(1)$ $(\omega \to \infty)$.

Proof. (i) Let us investigate first $K_\omega(t)$ in the interval $[1, \infty)$. Then using the expression of $D_u(t)$ in (4.3), we arrive at the result that

$$\begin{aligned} K_\omega(t) &= \frac{1}{\omega}\int_0^\omega \bar{\psi}_{[u]}(t)\mathscr{F}_{[t]}(u)\,du \\ &= \frac{1}{\omega}\sum_{i=0}^{[\omega]-1}\bar{\psi}_i(t)\int_i^{i+1}\mathscr{F}_{[t]}(u)\,du + \frac{1}{\omega}\bar{\psi}_{[\omega]}(t)\int_{[\omega]}^{\omega}\mathscr{F}_{[t]}(u)\,du \\ &\equiv K_\omega^{(1)}(t) + K_\omega^{(2)}(t). \end{aligned}$$

Taking into account that function $u \mapsto \mathscr{F}_{[t]}(u)$ is 1-periodic, let us express $K_\omega^{(1)}(t)$ as

$$\begin{aligned} K_\omega^{(1)}(t) &= \frac{1}{\omega}\left(\sum_{i=0}^{[\omega]-1}\bar{\psi}_i(t)\right)\int_0^1 \mathscr{F}_{[t]}(u)\,du \\ &= \frac{1}{\omega}\bar{D}_{[\omega]}(t)\int_0^1 \mathscr{F}_{[t]}(u)\,du. \end{aligned}$$

Since

$$\int_0^1 \mathscr{F}_{[t]}(u)\,du = \begin{cases} 0, & ([t] \neq 2^k), \\ \dfrac{1}{2^{k+2}}, & ([t] = 2^k), \quad (k = 0, 1, 2, \cdots), \end{cases} \tag{4.5}$$

again,

$$\int_{2^k}^{2^{k+1}} |\bar{D}_{[\omega]}(t)|\,dt = O(\ln \omega), \qquad (k = 0, 1, 2, \cdots)$$

(cf. [6]) for this reason:

$$\begin{aligned} \int_1^\infty |K_\omega^{(1)}(t)|\,dt &= \frac{1}{\omega}\sum_{k=0}^{\infty}\frac{1}{2^{k+2}}\int_{2^k}^{2^{k+1}}|\bar{D}_{[\omega]}(t)|\,dt \\ &= O\left(\frac{\ln \omega}{\omega}\right)\sum_{k=0}^{\infty}\frac{1}{2^{k+2}} = O\left(\frac{\ln \omega}{\omega}\right). \end{aligned} \tag{4.6}$$

Investigate now the order of magnitude of the integral $\int_{[\omega]}^{\omega}\mathscr{F}_{[t]}(u)\,du = \int_0^{\omega-[\omega]}\mathscr{F}_{[t]}(u)\,du$ occurring in $K_\omega^{(2)}(t)$. If $[t] = 2^k + 2^{k_1} + \cdots + 2^{k_s}$, where $k > k_1 > \cdots > k_s \geqq 0$, then

$$\left|\int_0^{\omega-[\omega]}\mathscr{F}_{[t]}(u)\,du\right| \leqq \frac{1}{2^{k+k_1}}, \tag{4.7}$$

which may be simply verified. Employing this we find that

$$|K_\omega^{(2)}(t)| \leqq \frac{1}{\omega}\frac{1}{2^{k+k_1}},$$

where the dyadic expansion of $[t]$ is of the above form. Making use of this, we can easily give an upper estimation for the integral $\int_1^\infty |K_\omega^{(2)}(t)|\,dt$ as follows:

$$\int_1^\infty |K_\omega^{(2)}(t)|\,dt \leqq \frac{1}{\omega}\sum_{k=0}^{\infty}\sum_{i=2^k}^{2^{k+1}-1}\frac{1}{2^{k+k_1^{(i)}}}, \tag{4.8}$$

where $k_1^{(i)}$ is such that $i=2^k+2^{k_1^{(i)}}+\cdots+2^{k_s^{(i)}}$, $k>k_1^{(i)}>\cdots>k_s^{(i)}\geqq 0$. Forming the sum providing the upper estimation further, we obtain that

$$\frac{1}{\omega}\sum_{k=0}^{\infty}\sum_{i=2^k}^{2^{k+1}-1}\frac{1}{2^{k+k_1^{(i)}}}=\frac{1}{\omega}\sum_{k=0}^{\infty}\frac{1}{2^k}\sum_{i=2^k}^{2^{k+1}-1}\frac{1}{2^{k_1^{(i)}}}; \tag{4.9}$$

again

$$\sum_{i=2^k}^{2^{k+1}-1}\frac{1}{2^{k_1^{(i)}}}=\sum_{j=0}^{k-1}\frac{1}{2^j}\sum_{\substack{i\\ i=2^k+2^j+\cdots\\ k>j}}=\sum_{j=0}^{k-1}\frac{1}{2^j}2^j=k; \tag{4.10}$$

hence

$$\int_1^\infty |K_\omega^{(2)}(t)|\,dt\leqq\frac{1}{\omega}\sum_{k=0}^{\infty}\frac{k}{2^k}=O\left(\frac{1}{\omega}\right).$$

Comparing this with (4.6), we find

$$\int_1^\infty |K_\omega(t)|\,dt=O\left(\frac{\ln\omega}{\omega}\right). \tag{4.11}$$

(ii) Further examinations will concern behavior of $K_\omega(t)$ in $[0,1)$. For development (4.3) of $D_u(t)$ we may set

$$\begin{aligned}K_\omega(t)&=\frac{1}{\omega}\int_0^\omega(\bar{D}_{[u]}(t)+(\omega-[\omega])\bar{\psi}_{[u]}(t))\,du\\&=\frac{1}{\omega}\sum_{i=0}^{[\omega]-1}\int_i^{i+1}(\bar{D}_{[u]}(t)+(\omega-[\omega])\bar{\psi}_{[u]}(t))\,du\\&\quad+\frac{1}{\omega}\int_{[\omega]}^{\omega}(\bar{D}_{[u]}(t)+(\omega-[\omega])\bar{\psi}_{[u]}(t))\,du\\&\equiv\tilde{K}_\omega^{(1)}(t)+\tilde{K}_\omega^{(2)}(t).\end{aligned}$$

We may express $\tilde{K}_\omega^{(1)}(t)$ as

$$\begin{aligned}\tilde{K}_\omega^{(1)}(t)&=\frac{1}{\omega}\sum_{i=0}^{[\omega]-1}(\bar{D}_i(t)+(\omega-[\omega])\bar{\psi}_i(t))\\&=\frac{1}{\omega}[\omega]\bar{K}_{[\omega]}(t)+\frac{1}{\omega}(\omega-[\omega])\bar{D}_{[\omega]}(t),\end{aligned}$$

where $\bar{K}_{[\omega]}$ denotes the Fejér kernel function concerning Walsh–Fourier series. Using $\int_0^1|\bar{K}_{[\omega]}(t)|\,dt=O(1)$ (cf. [8]) and $\int_0^1|\bar{D}_{[\omega]}(t)|\,dt=O(\ln\omega)$ (cf. [6]) we obtain

$$\int_0^1|\tilde{K}_\omega^{(1)}(t)|\,dt=O(1)+O\left(\frac{\ln\omega}{\omega}\right). \tag{4.12}$$

Writing $\tilde{K}_\omega^{(2)}(t)$ in another form, we get that

$$\tilde{K}_\omega^{(2)}(t)=\frac{1}{\omega}(\omega-[\omega])\bar{D}_{[\omega]}(t)+\frac{1}{\omega}(\omega-[\omega])^2\bar{\psi}_{[\omega]}(t);$$

for this reason

$$\int_0^1 |\tilde{K}_\omega^{(2)}(t)|\,dt=O\left(\frac{\ln\omega}{\omega}\right)+1. \tag{4.13}$$

Taking (4.11), (4.12) and (4.13), the statement to be proved is evident.

Let W_a denote the function whose Walsh–Fourier transform is

$$\hat{W}_a(y):=\begin{cases}0, & (0\leqq y<a),\\ 1/y, & (a\leqq y<\infty),\end{cases} \tag{4.14}$$

where $a>0$ is an arbitrary number to be fixed later. (See similar construction in [12].) Since $\hat{W}_a\in L^2(0,\infty)$, by the inversion formula there actually exists a function $W_a\in L^2(0,\infty)$ for which (4.14) is fulfilled. Now we show that there holds the following

LEMMA 4.2. $W_a\in L(0,\infty)$.

Proof. Let

$$W_{a,m}(x):=\int_a^{2^m}\frac{1}{y}\psi_x(y)\,dy, \qquad (x\in[0,\infty)). \tag{4.15}$$

Let us investigate—in case of $n>m$—difference

$$W_{a,n}(x)-W_{a,m}(x)=\int_{2^m}^{2^n}\frac{1}{y}\psi_x(y)\,dy.$$

Partially integrating twice, we obtain

$$\begin{aligned}W_{a,n}(x)-W_{a,m}(x)=&\frac{D_{2^n}(x)}{2^n}-\frac{D_{2^m}(x)}{2^m}\\ &+\frac{K_{2^n}(x)}{2^n}-\frac{K_{2^m}(x)}{2^m}+2\int_{2^m}^{2^n}\frac{1}{y^2}K_y(x)\,dy.\end{aligned} \tag{4.16}$$

Since $\|D_{2^j}\|_1=1$ $(j\in\mathbb{N}_+)$ and $\|K_\omega\|_1=O(1)$ $(\omega\to\infty)$, it follows that

$$\|W_{a,n}-W_{a,m}\|_1=O\left(\frac{1}{2^m}\right), \qquad (m<n).$$

For this reason the sequence of functions $W_{a,m}(m\in\mathbb{N}_+)$ is convergent in norm $\|\cdot\|_1$; let us denote the limit function by W. We shall show that $W=W_a$. By (1.8) the Walsh–Fourier transform is continuous, so

$$\lim_{m\to\infty}\hat{W}_{a,m}(t)=\hat{W}(t), \qquad (t\in[0,\infty)).$$

Let us introduce function

$$f_{a,m}(u):=\frac{\chi_{[a,2^m)}(u)}{u}, \qquad (u>0,\quad m\in\mathbb{N}_+).$$

Now we may set $W_{a,m} = \hat{f}_{a,m}$. As $f_{a,m} \in L^2(0, \infty)$. according to the inversion formula concerning the Walsh–Fourier transform of functions in $L^2(0, \infty)$,

$$\hat{W}_{a,m}(t) = (\hat{f}_{a,m})\hat{}\,(t) = f_{a,m}(t)$$
$$= \begin{cases} 0, & (0 \leqq t < a), \\ \dfrac{1}{t}, & (a \leqq t < 2^m), \\ 0, & (2^m \leqq t < \infty). \end{cases}$$

Employing this, we may write

$$\hat{W}(t) = \lim_{m \to \infty} \hat{W}_{a,m}(t) = \begin{cases} 0, & (0 \leqq t < a), \\ \dfrac{1}{t}, & (a \leqq t < \infty), \end{cases}$$

that is, $W = W_a$, which was stated.

LEMMA 4.3. *Let $d_n W_a$ denote the n-th "difference quotient" function of function W_a; that is, let*

$$d_n W_a := \frac{1}{2} \sum_{j=-n}^{n} 2^j [W_a - \tau_{2^{-(j+1)}} W_a].$$

Then $\|d_n W_a\|_1 = O_a(1)$.

Proof. We carry over the proof of the assertion for the case $a = 1$; if $a \neq 1$, then the lemma may also be proved by similar argument. Value of the constant in $O(1)$ depends on a.

Let $n \in \mathbb{N}$ be an arbitrary fixed number; moreover, let $m > n$. Let us decompose the function $W_{1,m}$ defined under (4.15) as follows:

$$W_{1,m}(x) = \int_1^{2^n} \frac{1}{y} \psi_x(y)\, dy + \int_{2^n}^{2^m} \frac{1}{y} \psi_x(y)\, dy$$
$$\equiv \Phi_1(x) + \Phi_2(x).$$

Since $\|d_n W_{1,m}\|_1 \to \|d_n W_1\|_1$ $(m \to \infty)$ and by (4.16) $\|\Phi_2\|_1 = O(1/2^n)$ if $m \to \infty$, thus

$$\|d_n \Phi_2\|_1 = \left\| \frac{1}{2} \sum_{j=-n}^{n} 2^j [\Phi_2 - \tau_{2^{-(j+1)}} \Phi_2] \right\|_1$$
$$\leqq \sum_{j=-n}^{n} 2^j \|\Phi_2\|_1 = \sum_{j=-n}^{-1} 2^j \|\Phi_2\|_1 + \sum_{j=0}^{n} 2^j \|\Phi_2\|_1$$
$$\leqq \|\Phi_2\|_1 + \|\Phi_2\|_1 (2^{n+1} - 1) = O\left(\frac{1}{2^n}\right) 2^{n+1} = O(1).$$

Now we are going to show that $\|d_n\Phi_1\|_1 = O(1)$ also holds. By the definition,

$$d_n\Phi_1(x) = \frac{1}{2}\sum_{j=-n}^{n} 2^j\left[\int_1^{2^n}\frac{1}{y}\psi_x(y)\,dy - \int_1^{2^n}\frac{1}{y}\psi_{x\dotplus 2^{-(j+1)}}(y)\,dy\right] \tag{4.17}$$

$$= \int_1^{2^n}\frac{1}{y}\left(\frac{1}{2}\sum_{j=-n}^{n} 2^j[1-\psi_{2^{-(j+1)}}(y)]\right)\psi_x(y)\,dy.$$

Introducing notation

$$\sigma_n(y) := \frac{1}{2}\sum_{j=-n}^{n} 2^j[1-\psi_{2^{-(j+1)}}(y)] \quad \left(= \sum_{k=-n-1}^{n-1} y_k/2^{k+1}\right), \tag{4.18}$$

(4.17) may be written in the form

$$d_n\Phi_1(x) = \int_1^{2^n}\frac{1}{y}\sigma_n(y)\psi_x(y)\,dy.$$

Let us decompose this integral into two members as follows:

$$d_n\Phi_n(x) = -\int_1^{2^n}\frac{1}{y}(y-\sigma_n(y))\psi_x(y)\,dy + \int_1^{2^n}\psi_x(y)\,dy \equiv -\Phi_1^{(1)}(x)+\Phi_1^{(2)}(x).$$

As

$$\Phi_1^{(2)}(x) = \sum_{k=1}^{2^n-1}\bar{\psi}_k(x)\int_k^{k+1}\bar{\psi}_{[x]}(y)\,dy = \begin{cases}\bar{D}_{2^n}(x)-1, & (0\leqq x<1),\\ 0, & (1\leqq x<\infty),\end{cases}$$

so

$$\|\Phi_1^{(2)}\|_1 = \frac{1}{2^n}(2^n-1)+\left(1-\frac{1}{2^n}\right) = O(1).$$

Convert $\Phi_1^{(1)}(x)$ in the following way:

$$\Phi_1^{(1)}(x) = \sum_{k=1}^{2^n-1}\bar{\psi}_k(x)\int_k^{k+1}\frac{1}{y}(y-\sigma_n(y))\bar{\psi}_{[x]}(y)\,dy.$$

By this we may write

$$\|\Phi_1^{(1)}\|_1 = \sum_{l=0}^{\infty} \int_l^{l+1} \Big| \sum_{k=1}^{2^n-1} \bar{\psi}_k(x) \cdot \int_k^{k+1} \frac{1}{y}(y-\sigma_n(y))\bar{\psi}_l(y)\,dy \Big|\,dx \tag{4.19}$$

$$\leqq \sum_{l=0}^{\infty} \sum_{k=1}^{2^n-1} \Big| \int_k^{k+1} \frac{1}{y}(y-\sigma_n(y))\bar{\psi}_l(y)\,dy \Big|.$$

If we prove that

$$\sum_{l=0}^{\infty} \Big| \int_k^{k+1} \frac{1}{y}(y-\sigma_n(y))\bar{\psi}_l(y)\,dy \Big| = O\Big(\frac{1}{2^n}\Big), \tag{4.20}$$

$$(k = 1, 2, \cdots, 2^n - 1),$$

then from this by (4.19), $\|\Phi_1^{(1)}\|_1 = O(1)$ follows. Since $y \in [1, 2^n)$, thus by (4.18),

$$y - \sigma_n(y) = \sum_{j=n}^{\infty} y_j/2^{j+1} = \sum_{j=n}^{\infty} \frac{1-r_j(y)}{2^{j+2}}.$$

Replacing this into (4.20), we have

$$\sum_{l=0}^{\infty} \Big| \int_k^{k+1} \frac{1}{y}\Big(\sum_{j=n}^{\infty} \frac{1-r_j(y)}{2^{j+2}} \Big)\bar{\psi}_l(y)\,dy \Big|$$

$$\leqq \sum_{l=0}^{\infty} \Big(\sum_{j=n}^{\infty} \frac{1}{2^{j+2}} \Big| \int_k^{k+1} \frac{1}{y}(1-r_j(y))\bar{\psi}_l(y)\,dy \Big| \Big)$$

$$\leqq \sum_{j=n}^{\infty} \frac{1}{2^{j+2}} \sum_{l=0}^{\infty} \Big| \int_k^{k+1} \frac{1}{y}\bar{\psi}_l(y)\,dy \Big| \tag{4.21}$$

$$+ \sum_{j=n}^{\infty} \frac{1}{2^{j+2}} \sum_{l=0}^{\infty} \Big| \int_k^{k+1} \frac{1}{y}\bar{\psi}_{2^j+l}(y)\,dy \Big|$$

$$= 2 \sum_{j=n}^{\infty} \frac{1}{2^{j+2}} \sum_{l=0}^{\infty} \Big| \int_k^{k+1} \frac{1}{y}\bar{\psi}_l(y)\,dy \Big|.$$

Let us form integral $\int_k^{k+1} (1/y)\bar{\psi}_l(y)\,dy$ by partial integration done twice; we get that

$$\int_k^{k+1} \frac{1}{y}\bar{\psi}_l(y)\,dy = \int_k^{k+1} \frac{1}{y^2}\mathscr{F}_l(y)\,dy$$

$$= \frac{1}{y^2}L_l(y)\Big|_k^{k+1} + 2\int_k^{k+1} \frac{1}{y^3}L_l(y)\,dy,$$

where $\mathscr{F}_l$ is an integral function of $\bar{\psi}_l$ and L_l is that of $\mathscr{F}_l$, that is,

$$\mathscr{F}_l(y) = \int_0^y \bar{\psi}_l(t)\,dt, \qquad L_l(y) = \int_0^y \mathscr{F}_l(t)\,dt.$$

Using (4.5) and the fact that the function $t \mapsto \mathscr{F}_1(t)$ is 1-periodic in case of $l \geqq 1$, for the absolute value of the integrated part,

$$\left| \frac{1}{y^2} L_l(y) \Big|_k^{k+1} \right| = 0$$

is obtained if $l = 0$, and in case of $l \geqq 1$,

$$\left| \frac{1}{y^2} L_l(y) \Big|_k^{k+1} \right| = \left| \frac{1}{(k+1)^2}(k+1) \int_0^1 \mathscr{F}_l(t)\, dt - \frac{1}{k^2} \cdot k \int_0^1 \mathscr{F}_l(t)\, dt \right|$$

$$= \begin{cases} 0, & (l \neq 2^j) \\ \dfrac{1}{k(k+1)} \dfrac{1}{2^{j+2}}, & (l = 2^j), (j = 0, 1, 2, \cdots); \end{cases}$$

thus

$$\sum_{l=0}^{\infty} \left| \frac{1}{y^2} L_l(y) \Big|_k^{k+1} \right| = \frac{1}{k(k+1)} \sum_{j=0}^{\infty} \frac{1}{2^{j+2}}$$

$$= \frac{1}{2} \frac{1}{k(k+1)} < \infty, \qquad (k = 1, 2, \cdots, 2^n - 1). \tag{4.22}$$

As in case of $l = 0$, $L_l(y) = \int_0^y t\, dt = y^2/2$; thus

$$2 \left| \int_k^{k+1} \frac{1}{y^3} L_l(y)\, dy \right| \leqq \frac{1}{k} \tag{4.23}$$

if $l = 0$. In case of $l \geqq 1$, let us consider the following decomposition of $L_l(y)$:

$$L_l(y) = \sum_{i=0}^{[y]-1} \int_i^{i+1} \mathscr{F}_l(t)\, dt + \int_{[y]}^{y} \mathscr{F}_l(t)\, dt$$

$$= [y] \int_0^1 \mathscr{F}_l(t)\, dt + \int_0^{y-[y]} \mathscr{F}_l(t)\, dt.$$

Employing this and (4.7), we get for $l \geqq 1$,

$$2 \left| \int_k^{k+1} \frac{1}{y^3} L_l(y)\, dy \right| = 2k \int_0^1 \mathscr{F}_l(t)\, dt \cdot \int_k^{k+1} \frac{1}{y^3}\, dy$$

$$+ 2 \left| \int_k^{k+1} \frac{1}{y^3} \left(\int_0^{y-[y]} \mathscr{F}_l(t)\, dt \right) dy \right| \tag{4.24}$$

$$\leqq \frac{2}{k^2} \int_0^1 \mathscr{F}_l(t)\, dt + \frac{2}{k^3} \frac{1}{2^{k(l)+k(l)}},$$

where $l=2^{k_0^{(l)}}+2^{k_1^{(l)}}+\cdots+2^{k_s^{(l)}}$, $k_0^{(l)}>k_1^{(l)}>\cdots>k_s^{(l)}\geqq 0$. On the basis of (4.9), (4.10), (4.23) and (4.24),

$$\sum_{l=0}^{\infty}\left|2\int_k^{k+1}\frac{1}{y^3}L_l(y)\,dy\right|\leqq\frac{1}{k}+\frac{2}{k^2}\sum_{j=0}^{\infty}\frac{1}{2^{j+2}}+\frac{2}{k^3}\sum_{j=0}^{\infty}\frac{j}{2^j}<\infty, \qquad (k=1,2,\cdots,2^k-1) \tag{4.25}$$

follows. Taking (4.22) and (4.25) into consideration, we obtain

$$\sum_{l=0}^{\infty}\left|\int_k^{k+1}\frac{1}{y}\bar{\psi}_l(y)\,dy\right|<\infty, \qquad (k=1,2,\cdots,2^n-1), \tag{4.26}$$

which by (4.21) proves (4.20). We have therefore proved the lemma.

By the convolution theorem it is easy to prove

LEMMA 4.4. (a) *Let $f\in L(0,\infty)$ be an arbitrary function for which $\hat{f}(y)=0$ $(0\leqq y<a)$, and $g\in L(0,\infty)$ be a function, for which*

$$\hat{g}(y)=y\hat{f}(y), \qquad (y\in[0,\infty)). \tag{4.27}$$

Then $f=W_a*g$.

(b) *Let $f\in L^2(0,\infty)$ be an arbitrary function for which $\hat{f}(y)=0$ $(0\leqq y<a)$; moreover, $g\in L^2(0,\infty)$ be a function for which $\hat{g}(y)=y\cdot\hat{f}(y)$ $(y\in[0,\infty))$. Then $f=W_a*g$.*

The following theorem shows that inverse operation of the derivation defined in (3.1) is the convolution with a function W_a.

THEOREM 4.1. (a) *Let $f\in L(0,\infty)$ be D-L differentiable, and let $\hat{f}(y)=0$ $(0\leqq y<a)$. Then*

$$W_a*D^{[1]}f=f. \tag{4.28}$$

(b) *Let $f\in L^2(0,\infty)$ be D-L^2 differentiable, and again, let $\hat{f}(y)=0$ $(0\leqq y<a)$. Then*

$$W_a*D^{[1]}f=f. \tag{4.29}$$

Proof. Applying Lemma (4.4)—taking Theorem 3.2 into consideration—to function $g=D^{[1]}f$, we obtain the required assertion.

In the following we shall show that an equivalent of the well-known theorem concerning differentiability of an integral function also holds.

This is stated in

THEOREM 4.2. (a) *Let $f\in L(0,\infty)$ be an arbitrary function. Now W_a*f is D-L differentiable, and*

$$D^{[1]}(W_a*f)=f. \tag{4.30}$$

(b) *Let $f\in L^2(0,\infty)$ be an arbitrary function, for which $\hat{f}(y)=0$ $(0\leqq y<a)$ is satisfied. Then W_a*f is D-L^2 differentiable, and*

$$D^{[1]}(W_a*f)=f. \tag{4.31}$$

Proof. (a) By the definition

$$d_n(W_a * f)(x) = \frac{1}{2}\sum_{j=-n}^{n} 2^j\Big[\int_0^\infty W_a(x \mathring{+} y)f(y)\,dy - \int_0^\infty W_a(x \mathring{+} y \mathring{+} 2^{-(j+1)})f(y)\,dy\Big]$$

$$= \int_0^\infty \Big\{\frac{1}{2}\sum_{j=-n}^{n} 2^j[W_a(x \mathring{+} y) - W_a(x \mathring{+} y \mathring{+} 2^{-(j+1)})]\Big\}f(y)\,dy$$

$$= (d_n W_a * f)(x),$$

that is, $d_n(W_a * f) = d_n W_a * f = W_a * d_n f$. Let us define sequence of operators $T_n : L(0, \infty) \to L(0, \infty)$ $(n \in \mathbb{N})$ as follows:

$$T_n f := d_n W_a * f, \qquad (f \in L(0, \infty), \quad n \in \mathbb{N}).$$

By this definition validity of (4.30) is equivalent to the equality

$$\lim_{n\to\infty} \|T_n f - f\|_1 = 0. \tag{4.32}$$

Since

$$\|T_n f\|_1 = \|d_n W_a * f\|_1 \leqq \|d_n W_a\|_1 \|f\|_1,$$

that is, $\|T_n\| \leqq \|d_n W_a\|_1$—considering the assertion of Lemma (4.3)—it is sufficient to prove (4.32) for the elements of a class of functions everywhere dense in $L(0, \infty)$, as by the Banach–Steinhaus theorem fulfilment of (4.32) follows from this in case of arbitrary $f \in L(0, \infty)$. Let

$$\tilde{\psi}_m(x) := \begin{cases} \bar{\psi}_m(x), & (x \in [0, 1)), \\ 0, & (\text{otherwise}), \end{cases} \qquad (m \in \mathbb{N}).$$

We shall prove that

$$\lim_{n\to\infty} \|d_n W_a * \tilde{\psi}_m - \tilde{\psi}_m\|_1 = \lim_{n\to\infty} \|W_a * d_n \tilde{\psi}_m - \tilde{\psi}_m\|_1 = 0. \tag{4.33}$$

The "difference quotient" function $d_n \bar{\psi}_m$ may be expressed in the following manner:

$$d_n \tilde{\psi}_m(x) = \frac{1}{2}\sum_{j=-n}^{n} 2^j[\tilde{\psi}_m(x) - \tilde{\psi}_m(x \mathring{+} 2^{-(j+1)})]$$

$$= \frac{1}{2}\sum_{j=0}^{n} 2^j[\tilde{\psi}_m(x) - \tilde{\psi}_m(x \mathring{+} 2^{-(j+1)})] + \frac{1}{2}\sum_{j=-n}^{-1} 2^j[\tilde{\psi}_m(x) - \tilde{\psi}_m(x \mathring{+} 2^{-(j+1)})]$$

$$\equiv Q_1(x) + Q_2(x).$$

Considering the definition of $\tilde{\psi}_m$, $Q_n(x)$ may be obtained in the form

$$Q_1(x)=\begin{cases}\bar{\psi}_m(x)\dfrac{1}{2}\sum\limits_{j=0}^{n}2^j[1-\bar{\psi}_m(2^{-(j+1)})], & (x\in[0,1)),\\ 0, & \text{(otherwise)}.\end{cases}$$

Since

$$m_j=\frac{1-\bar{\psi}_m(2^{-(j+1)})}{2}$$

(where $m=\sum_{k=0}^{\infty}m_k\cdot 2^k$ is the dyadic expansion of m), thus

$$Q_1(x)=\begin{cases}m\cdot\bar{\psi}_m(x), & (x\in[0,1)),\\ 0, & \text{(otherwise)},\end{cases}\tag{4.34}$$

if $n>[\log_2 m]$.

By a simple change of indices $Q_2(x)$ may be changed to

$$\begin{aligned}Q_2(x)&=\frac{1}{2}\sum_{j=-n}^{-1}2^j[\tilde{\psi}_m(x)-\tilde{\psi}_m(x\mathring{+}2^{-(j+1)})]\\&=\frac{1}{2}\sum_{k=0}^{n-1}\frac{1}{2^{k+1}}[\tilde{\psi}_m(x)-\tilde{\psi}_m(x\mathring{+}2^k)].\end{aligned}$$

According to this, we may write

$$Q_2(x)=\begin{cases}\bar{\psi}_m(x)\dfrac{1}{2}\sum\limits_{k=0}^{n-1}\dfrac{1}{2^{k+1}}, & (x\in[0,1)),\\ \dfrac{1}{2^{j+2}}\tilde{\psi}_m(x\mathring{+}2^j), & (x\in[2^j,2^j+1)), \qquad j=0,1,\cdots,n-1,\end{cases}$$

or,

$$Q_2(x)=\begin{cases}\dfrac{1}{2}\left(1-\dfrac{1}{2^n}\right)\bar{\psi}_m(x), & (x\in[0,1)),\\ \dfrac{1}{2^{j+2}}\bar{\psi}_m(x-2^j), & (x\in[2^j,2^{j+1})), \quad j=0,1,\cdots,n-1.\end{cases}\tag{4.35}$$

Comparing this with (4.34), we obtain that in case of $n>[\log_2 m]$,

$$d_n\tilde{\psi}_m(x)=\begin{cases}\left(\dfrac{1}{2}-\dfrac{1}{2^{n+1}}+m\right)\bar{\psi}_m(x), & (x\in[0,1)),\\ \dfrac{1}{2^{j+2}}\bar{\psi}_m(x-2^j), & (x\in[2^j,2^j+1)),\end{cases}\tag{4.36}$$

$$j=0,1,\cdots,n-1.$$

Let us observe the function $W_a*d_n\tilde{\psi}_m-\tilde{\psi}_m=d_nW_a*\tilde{\psi}_m-\tilde{\psi}_m\in L(0,\infty)\cap L^2(0,\infty)$. Since

$$(\tilde{\psi}_m)\hat{}=\chi_{[m,m+1)}, \qquad (d_nW_a)\hat{}=\sigma_n\cdot\hat{W}_a,$$

(where σ_n is the function defined in (4.18))—which are easily provable—the Walsh–Fourier transform of this function may be put in the form:

$$(W_a * d_n\tilde{\psi}_m - \tilde{\psi}_m)\hat{}(y) = (d_n W_a * \tilde{\psi}_m - \tilde{\psi}_m)\hat{}(y)$$
$$= (\hat{W}_a(y)\sigma_n(y) - 1)\chi_{[m,m+1)}(y), \qquad (y \in [0, \infty)).$$

Let us now apply the inversion formula (1.11) concerning the Walsh–Fourier transform. It follows that

$$(d_n W_a * \tilde{\psi}_m)(t) - \tilde{\psi}_m(t)$$
$$= \int_m^{m+1} \left(\frac{1}{y}\sigma_n(y) - 1\right)\psi_t(y)\, dy, \qquad (t \in [0, \infty)).$$

Employing $\lim_{n\to\infty} \sigma_n(y) = y$ $(y \in [0, \infty))$, by Lebesgue's theorem, we arrive at the result

$$\lim_{n\to\infty} (d_n W_a * \tilde{\psi}_m)(t) = \tilde{\psi}_m(t), \qquad (t \in [0, \infty)). \tag{4.37}$$

By the presentation of $d_n\tilde{\psi}_m$ in (4.36),

$$|d_n\tilde{\psi}_m(x)| \leqq \begin{cases} m+1, & (x \in [0, 1)), \\ \dfrac{1}{2^{j+2}}, & (x \in [2^j, 2^j + 1)), \quad j \in \mathbb{N}_+, \end{cases} \tag{4.38}$$

in case of arbitrary $n > [\log_2 m]$. Denote by M_m the majorizing function on the right-hand side of (4.38). Since $M_m \in L(0, \infty)$, moreover,

$$|(d_n W_a * \tilde{\psi}_m)(x)| = |(W_a * d_n\tilde{\psi}_m)(x)|$$
$$= \left|\int_0^{\infty} W_a(x \mathring{+} y) d_n\tilde{\psi}_m(y)\, dy\right| \leqq (|W_a| * M_m)(x),$$
$$(x \in [0, \infty), \quad n > [\log_2 m]),$$

and $|W_a| * M_m \in L(0, \infty)$, we see that functions $d_n W_a * \tilde{\psi}_m$ have a common integrable majorant. From this and (4.37), equality (4.33)—which was to be proved—follows. As the class of functions arising from functions $\tilde{\psi}_m (m \in \mathbb{N})$ by translation is everywhere dense in $L(0, \infty)$ and $\|d_n W_a * \tau_h\tilde{\psi}_m - \tau_h\tilde{\psi}_m\|_1 = \|d_n W_a * \tilde{\psi}_m - \tilde{\psi}_m\|_1$ in case of arbitrary $h \in [0, \infty)$, the proof of part (a) of the theorem is complete.

(b) To verify equality $\lim_{n\to\infty} \|d_n W_a * f - f\|_2 = 0$, the isometry property of the Walsh–Fourier transform suffices to show that $\lim_{n\to\infty} \|(d_n W_a)\hat{} \cdot \hat{f} - \hat{f}\|_2 = 0$. But

$$(d_n W_a)\hat{}(y) = \frac{1}{2}\sum_{j=-n}^{n} 2^j[\hat{W}_a(y) - \hat{W}_a(y)\psi_{2^{-(j+1)}}(y)]$$
$$= \sigma_n(y)\hat{W}_a(y),$$

so that

$$\|(d_n W_a)\hat{} \cdot \hat{f} - \hat{f}\|_2^2 = \int_a^\infty |\hat{f}(y)|^2 \left| \frac{1}{y} \sigma_n(y) - 1 \right|^2 dy.$$

From this, (3.4) and Lebesgue's theorem, it follows that

$$\lim_{n \to \infty} \|(d_n W_a)\hat{} \cdot \hat{f} - \hat{f}\|_2 = 0,$$

which was to be proved.

Acknowledgment. I wish to express my thanks to Professor Ferenc Schipp for calling my attention to this range of problems and for his many helpful suggestions; I am also indebted to Professor P. L. Butzer who sent me the latest literature on the subject and helped to put the paper in its final form.

REFERENCES

[1] P. L. BUTZER AND H. J. WAGNER, *Walsh–Fourier series and the concept of a derivative*, Applicable Anal., 3 (1973), pp. 29–46.

[2] ———, *Approximation by Walsh polynomials and the concept of a derivative*, Applications of Walsh Functions, Proc. Symp. Applications of Walsh-Functions, Washington D.C., March 27–29, 1972, pp. 388–392.

[3] ———, *On a Gibbs-type derivative in Walsh–Fourier analysis with applications*, Proc. 1972 National Electronics Conf., Chicago, 1972, Oak Brook, Ill., 1972, pp. 393–398.

[4] R. B. CRITTENDEN, *Walsh–Fourier transforms*, Applications of Walsh Functions, Proc. Symp. and Workshop, Naval Research Lab., Washington, D.C., 1970, pp. 170–174.

[5] N. J. FINE, *The generalized Walsh functions*, Trans. Amer. Math. Soc., 69 (1950), pp. 66–77.

[6] ———, *On the Walsh functions*, Ibid., 65 (1949), pp. 372–414.

[7] G. MORGENTHALER, *Walsh–Fourier series*, Ibid., 84 (1957), pp. 472–507.

[8] S. YANO, *On approximation by Walsh functions*, Proc. Amer. Math. Soc., 2 (1951), pp. 962–967.

[9] A. ZYGMUND, *Trigonometric Series*. vol. II, Cambridge University Press, London, 1959.

[10] P. L. BUTZER AND H. J. WAGNER, *A calculus for Walsh functions defined on R_+*, Applications of Walsh Functions, Proc. Sympos., Naval Research Lab., Washington, D.C., April 18–20, 1973, pp. 75–81.

[11] F. PICHLER, *Walsh functions and linear system theory*, Applications of Walsh Functions, Proc. Sympos. Naval Research Lab., Washington, D.C., March 31–April 3, 1970, C. A. Bass, ed., pp. 175–182.

[12] C. WATARI, *Best approximation by Walsh polynomials*, Tôhoku Math. J. 15 (1963), pp. 1–5.

[13] F. PICHLER, *Dyadische Faltungsoperatoren zur Beschreibung linearer Systeme*, Sitzungsberichte der Oesterreichischen Akademie der Wissenschaften, Mathematisch-naturwissenschaftliche Klasse, Abteilung II, Band 180, Heft 1–3, 1971, pp. 69–87.

[14] F. PICHLER, J. E. MARSHALL AND J. E. GIBBS, *A System-Theory Approach to Electrical Measurements*, Hatfield, England, 1973.

[15] S. WANATABE, *Knowing and guessing: A Quantitative Study of Inference and Information*, John Wiley, New York, 1969.

[16] J. E. GIBBS AND B. IRELAND, *Walsh functions and differentiation*, Applications of Walsh Functions, Proc. Sympos. Naval Research Lab., Washington, D.C., March 18–20, 1974, 1974, pp. 147–176.

[17] H. F. HARMUTH, *Real numbers versus dyadic group as basis for models of time and space*, preprint.

[18] J. PAL, *On the connection between the concept of a derivative defined on the dyadic field and the Walsh–Fourier transform*, Ann. Univ. Sci. Budapest. Eötvös Sect. Math., to appear.

[19] T. S. DURRANI AND E. M. STAFFORD, *Hardware applications of Walsh functions*, Internat. J. Electronics, 33 (1972), pp. 53–65.

[20] N. AHMED AND K. R. RAO, *Orthogonal Transforms for Digital Signal Processing*, Springer-Verlag, Berlin-Heidelberg-New York, 1975.

[21] H. F. HARMUTH, *Research and development in the field of Walsh functions and sequency theory*, Advances in Electronics and Electron Physics, vol. 36, Academic Press, New York, 1974, pp. 195–264.

[22] H. HÜBNER, *Signalmultiplexbildung im Korrelations- und Matrixvielfach*, Symp. on Theory and Appl. of Walsh and Other Non-Sinusoidal Functions, Hatfield, England, 1975.

[23] J. E. GIBBS, *Differentiation and frequency on the dyadic group*, Rep. 18, National Physical Laboratory, Middlesex, England, 1975.

SIAM J. MATH. ANAL.
Vol. 8, No. 3, May 1977

EXTENT OF THE LEFT BRANCHING SOLUTION TO CERTAIN BIFURCATION PROBLEMS*

W. E. OLMSTEAD†

Abstract. A special class of nonlinear eigenvalue problems which exhibit branching to the left of (below) the smallest eigenvalue is considered. This particular class serves to illustrate a procedure for bounding the leftward extent of a nonnegative solution branch. The procedure relies upon a well-known result about upper and lower solutions associated with a monotone operator. In the situation of left-branching bifurcation, the more difficult determination of a suitable lower solution is achieved by using an explicit solution to a simpler nonlinear problem. A physical example relative to the buckling of a nonlinearly elastic rod is worked out in detail.

1. Introduction. The purpose of this paper is to demonstrate a procedure that can provide some global information about certain birfurcation problems which exhibit branching to the left of the lowest eigenvalue. While the method suggests some general applicability, we will only be concerned here with a limited class of problems which illustrate the idea.

Consider the nonlinear boundary value problem

$$u''(x)+\lambda f(u(x))=0, \qquad 0<x<l, \quad \lambda>0, \tag{1.1}$$

$$u'(0)=u(l)=0, \qquad u(0)\geqq 0. \tag{1.2}$$

We take $f(z)$ to be thrice continuously differentiable with respect to z and,

$$f'(z)\geqq 0, \quad z\geqq 0; \quad f(0)=0, \quad f'(0)>0. \tag{1.3}$$

Additional conditions on $f(z)$ will be imposed as needed.

This problem has the equivalent and sometimes more convenient formulation as an integral equation,

$$u(x)=\lambda\int_0^l g(x|\xi)f(u(\xi))\,d\xi\equiv Au(x), \tag{1.4}$$

where $g(x|\xi)=l-\xi+(\xi-x)H(x-\xi)\geqq 0$.

For appropriate choices of $f(u)$, it is well known that this problem can have nontrivial solutions which branch from $u\equiv 0$ at the eigenvalues λ_n, $n=1,2,\cdots$, of the linearized problem. The local behavior near the branch points is explicitly known (cf. Keller [3]). The existence and gross characterization of the solutions away from the branch points has also been investigated (cf. Rabinowitz [6], Wolkowisky [10]).

We will be concerned with the situation in which a bounded, nonnegative solution exists on some interval, $0<\lambda'\leqq\lambda<\lambda_1$, where λ' is not particularly close to zero. A typical case is depicted in the bifurcation diagram of Fig. 1. The branch corresponding to the nonnegative solution bifurcates from the trivial solution at the smallest eigenvalue λ_1 and extends to the left as far as some λ' before winding back to the right.

* Received by the editors October 6, 1975, and in revised form November 10, 1975.

† The Technological Institute, Northwestern University, Evanston, Illinois 60201. This work was supported by the National Science Foundation under Grant GP-44027.

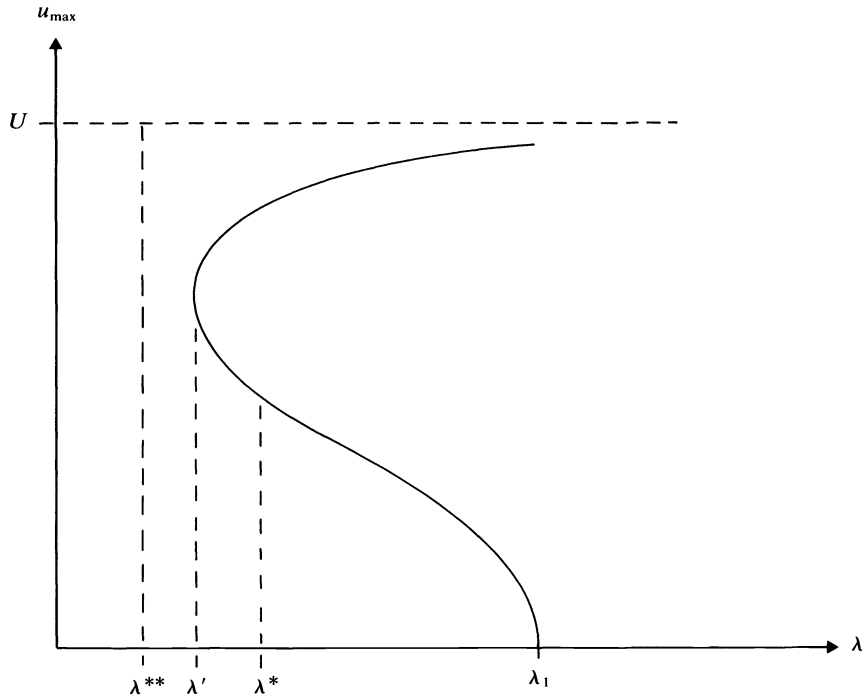


FIG. 1. *Bifurcation diagram*

Generally it is difficult to obtain the kind of detailed information about the entire bifurcation diagram which is shown in Fig. 1. Fortunately some limiting features of the diagram can frequently be determined. Results on the local behavior near λ_1 indicate how the branch starts. A boundedness result, $0 \leqq u \leqq U$, can often be obtained to limit the upward extent of the branch. The leftward extent of the branch is limited by a uniqueness result which establishes $u \equiv 0$ as the only solution for $0 < \lambda \leqq \lambda^{**}$. Also, a stability analysis would reveal the solution on the upper part of the branch as stable and the lower part as unstable.

Our main interest here is to determine more information about the leftward extent of the branch. While uniqueness provides some bound $\lambda^{**} \leqq \lambda'$, we rarely know how accurate this estimate is. We will show that, for certain problems like (1.1)–(1.2) some λ^* can be determined such that a nonnegative solution exists for $\lambda^* \leqq \lambda < \lambda_1$. This result together with the uniqueness results yields the estimate

$$0 < \lambda^{**} \leqq \lambda' \leqq \lambda^* < \lambda_1. \tag{1.5}$$

This inequality provides some useful information about the location of λ'.

A result like (1.5) can be important in many problems of physical interest. As one example, we will treat the buckling of a rod with a nonlinear compressibility property. The first buckled mode of the rod can be described with a bifurcation diagram like that of Fig. 1. The parameter λ is inversely proportional to the bending stiffness of the rod. For λ sufficiently small only the unbuckled state ($u \equiv 0$) of rod is possible. However, for values of λ which are significantly below

the classical critical value λ_1, it is possible for the rod to buckle. An estimate, like (1.5), gives some indication of the true critical value λ' where the buckling can first occur.

The procedure described here for determining an appropriate λ^* relies in part on certain monotone methods which have been widely used in the analysis of nonlinear problems (cf. Keller and Cohen [2], Krasnosel'skii [4], Sattinger [7], [8]).

Under the conditions indicated, the operator A of (1.4) is monotone with respect to the cone of nonnegative, continuous functions. Moreover, for such monotone operators it is known that if some $V \geqq 0$ and $v \geqq 0$ can be found such that $AV \leqq V$ and $Av \geqq v$ with $v \leqq V$ for all λ, $\lambda^* \leqq \lambda \leqq \lambda_1$, then (1.4) has at least one solution u, $v \leqq u \leqq V$.

When the branching is like that of Fig. 1, it is relatively easy to find a suitable V. Often, some positive constant will suffice. It is less easy to find a suitable v. A principal feature of the work here is the choice of v as the solution of a simpler nonlinear problem, closely related to (1.1), (1.2). For the class of problems considered, v is explicitly given in terms of elliptic functions. Thus the requirements that $v \leqq Av$ and $v \leqq V$ can lead to an explicit value of λ^*.

2. Basic results. We will use several basic results relevant to that branch of the nonnegative solution of (1.4) which emanates from the smallest eigenvalue λ_1 of the linearized problem. Collectively, these results provide a qualitative description of the bifurcation diagram of Fig. 1.

The bifurcation behavior at λ_1 requires some knowledge of the linearized problem

$$\psi''(x) + \lambda f'(0)\psi(x) = 0, \qquad 0 < x < l, \quad \lambda > 0, \tag{2.1}$$

$$\psi'(0) = \psi(l) = 0, \qquad \psi(0) = 1. \tag{2.2}$$

We only use the eigenfunction corresponding to the lowest eigenvalue, namely

$$\psi_1(x) = \cos\frac{\pi x}{2l}, \qquad \lambda_1 = \frac{\pi^2}{4l^2 f'(0)}. \tag{2.3}$$

Now the local branching at λ_1 is summarized in the following result.

THEOREM 1. *There exists a left-branching, nonnegative solution of* (1.4) *for* λ *sufficiently near* λ_1 *if*

(i) $f''(0) = 0$, $f'''(0) > 0$; *whereupon,*

$$u(x) = \left[\frac{8f'(0)}{f'''(0)}\left(1 - \frac{\lambda}{\lambda_1}\right)\right]^{1/2} \cos\frac{\pi x}{2l} + o\left[\left(1 - \frac{\lambda}{\lambda_1}\right)^{1/2}\right]; \tag{2.4}$$

or if

(ii) $f''(0) > 0$; *whereupon,*

$$u(x) = \left[\frac{3\pi f'(0)}{4f''(0)}\left(1 - \frac{\lambda}{\lambda_1}\right)\right] \cos\frac{\pi x}{2l} + o\left[\left(1 - \frac{\lambda}{\lambda_1}\right)\right]. \tag{2.5}$$

This theorem follows as a special case of that due to Keller [3].

To limit the leftward extent of this branch, we must impose some additional condition on $f(z)$ which will allow us to prove a uniqueness result for the trivial solution. It is clear from (1.4) that the operator A maps the space of nonnegative, square integrable functions into itself. Therefore for any continuous function $u(x)$ which is a solution of (1.4), we have

$$\|u\|_{L_2} = \|Au\|_{L_2} \leqq \frac{\lambda}{f'(0)\lambda_1}\|f(u)\|_{L_2}. \tag{2.6}$$

Suppose there exists a constant $\gamma > 0$ such that $f(u) \leqq \gamma u$, $u \geqq 0$; then clearly

$$\|u\|_{L_2} \leqq \frac{\lambda\gamma}{f'(0)\lambda_1}\|u\|_{L_2}. \tag{2.7}$$

If $\lambda\gamma < f'(0)\lambda_1$, then (2.7) implies that $\|u\|_{L_2} = 0$. Thus we have established the following result.

THEOREM 2. *Let there exist a constant $\gamma > 0$ such that $f(z) \leqq \gamma z$, $z \geqq 0$. If $\lambda < \lambda^{**} = f'(0)\lambda_1\gamma^{-1}$, then the only continuous nonnegative solution of* (1.4) *is $u \equiv 0$.*

This theorem provides some sufficient value $\lambda^{**} > 0$, which limits the leftward extent of the branch.

The upward extent of the branch is often limited by a boundedness property on $f(z)$. For example, it is easy to establish a result like

THEOREM 3. *Let there exist a constant $\beta > 0$ such that $f(z) \leqq \beta$, $z > 0$. Then any nonnegative, continuous solution of* (1.4) *satisfies*

$$u(x) \leqq U = \frac{\beta\pi^2}{8f'(0)}, \qquad 0 \leqq x \leqq l. \tag{2.8}$$

This result follows immediately from (1.4) upon replacing λ by λ_1 and noting that $\int_0^l g(x|\xi)\, d\xi \leqq \frac{1}{2}l^2$.

We now turn our attention toward the determination of a suitable λ^* which will help delineate the extent of the leftward branch. We will rely upon the following result.

THEOREM 4. *Let V and v be nonnegative, continuous functions which satisfy*

$$AV \leqq V, \quad Av \geqq v, \quad v \leqq V. \tag{2.9}$$

Then there exists a solution of $u = Au$ such that

$$v \leqq u \leqq V. \tag{2.10}$$

Here V and v are called, respectively, an *upper* and *lower solution* of (1.4).

Theorem 4 is a special case of Theorem 4.1 of Krasnosel'skii [4]. Essentially all that needs to be verified is that A is a monotone operator on the cone of nonnegative, continuous functions. That property does hold because if u_1 and u_2 are each continuous and nonnegative functions with $u_1 \geqq u_2$, $0 \leqq x \leqq l$, then

$$Au_1 - Au_2 = \lambda \int_0^l g(x|\xi) f'(\theta(\xi))[u_1(\xi) - u_2(\xi)]\, d\xi, \qquad 0 \leqq u_2 \leqq \theta \leqq u_1. \tag{2.11}$$

Since $g \geqq 0$, $f' \geqq 0$ and $\lambda > 0$; this representation clearly implies that $Au_1 \geqq Au_2$.

It should be pointed out that while $f'(z) \geqq 0$ is convenient for the approach presented here, this condition is not a necessary one for dealing with problems like (1.1)–(1.2). Indeed, Sattinger [7] has developed results with monotone methods under much weaker conditions on $f(z)$.

The proof of Theorem 4 involves the construction of certain monotone sequences which converge to a solution of $Au = u$. The results of Sattinger [7] indicate that this convergence is always to a stable solution. Thus the existence statement actually refers to the upper portion of the solution branch in Fig. 1.

Our task remains to find suitable functions V and v such that (2.9) holds for $0 < \lambda^* \leqq \lambda \leqq \lambda_1$. It is relatively easy to find an upper solution V. The determination of an appropriate lower solution v is more difficult.

Under the assumed condition on $f(z)$ in Theorem 3, it is straightforward to show that the constant U, defined by (2.8), provides an acceptable upper solution. For $\lambda \leqq \lambda_1$, we find that

$$AU = \lambda f(U) \int_0^l g(x|\xi)\, d\xi \leqq \frac{\lambda_1 f(U) l^2}{2} \leqq \frac{\beta \pi^2}{8 f'(0)} = U. \tag{2.12}$$

Therefore $V = U$ can be used, although a better choice may follow in certain individual examples.

To obtain candidates for a lower solution, we will introduce two nonlinear boundary value problems. In essence, these problems represent the two possible limiting forms of (1.1)–(1.2) where there is branching to the left of the smallest eigenvalue. First we consider

$$\hat{w}''(x) + \Lambda[\hat{w}(x) + a\hat{w}^3(x)] = 0, \qquad 0 < x < l, \quad \Lambda > 0, \tag{2.13}$$

$$\hat{w}'(0) = \hat{w}(l) = 0, \qquad \hat{w}(0) > 0. \tag{2.14}$$

Here $a > 0$ is a constant to be specified later. As an equivalent integral equation, this problem takes the form

$$\hat{w}(x) = \Lambda \int_0^l g(x|\xi)[\hat{w}(\xi) + a\hat{w}^3(\xi)]\, d\xi, \tag{2.15}$$

where $g(x|\xi)$ is the same as in (1.4).

The nonnegative solution of (2.13)–(2.14) or (2.15) which exhibits left-branching at the smallest eigenvalue $\Lambda_1 = \pi^2/4l^2$ of the linearized problem ($a = 0$) is given by

$$\hat{w}(x) = \left[\frac{2m}{a(1-2m)}\right]^{1/2} \operatorname{cn}\left(\left[\frac{\Lambda}{1-2m}\right]^{1/2} x|m\right), \qquad 0 \leqq x \leqq l, \tag{2.16}$$

where Λ depends on m through the relation

$$\frac{\Lambda}{\Lambda_1} = \left[\frac{2}{\pi} K(m)\right]^2 (1-2m), \qquad 0 \leqq m < \tfrac{1}{2}. \tag{2.17}$$

Here $\operatorname{cn}(z|m)$ is the Jacobian elliptic cosine function with parameter m, and $K(m)$ is the complete elliptic integral of the first kind. In (2.17) it can be shown from the properties of $K(m)$ that Λ is a monotonically decreasing function of m, $0 \leqq m < \frac{1}{2}$, with $\Lambda = \Lambda_1$ when $m = 0$ and $\Lambda \to 0$ as $m \to \frac{1}{2}$.

As the other candidate for a lower solution, we consider

$$\tilde{w}''(x)+\Lambda[\tilde{w}(x)+b\tilde{w}^2(x)]=0, \qquad 0<x<l, \quad \Lambda>0, \tag{2.18}$$

$$\tilde{w}'(0)=\tilde{w}(l)=0, \qquad \tilde{w}(0)>0. \tag{2.19}$$

Here $b>0$ is a constant to be specified later. Again we have the equivalent integral equation

$$\tilde{w}(x)=\Lambda\int_0^l g(x|\xi)[\tilde{w}(\xi)+b\tilde{w}^2(\xi)]\,d\xi. \tag{2.20}$$

The nonnegative solution of (2.18)–(2.19) or (2.20) which exhibits left-branching at the smallest eigenvalue $\Lambda_1=\pi^2/4l^2$ of the linearized problem ($b=0$) is given by

$$\tilde{w}(x)=\frac{3m}{2b\tau}\left\{\mathrm{sn}^2\left(\left[\frac{\Lambda}{4\tau}\right]^{1/2}l|m\right)-\mathrm{sn}^2\left(\left[\frac{\Lambda}{4\tau}\right]^{1/2}x|m\right)\right\}, \qquad 0\leqq x\leqq l, \tag{2.21}$$

where $\tau=(1-m+m^2)^{1/2}$, and Λ depends upon m through the relation

$$\frac{\Lambda}{\Lambda_1}=\frac{16\tau}{\pi^2}\{\mathrm{sn}^{-1}([1+m+\tau]^{-1/2})\}^2, \qquad 0\leqq m\leqq\tfrac{1}{2}. \tag{2.22}$$

Here $\mathrm{sn}\,(x|m)$ is the Jacobian elliptic sine function with parameter m, and $\mathrm{sn}^{-1}(y)$ is its principal inverse. In (2.22) it can be shown that Λ is a monotonically decreasing function of m, $0\leqq m\leqq\frac{1}{2}$, with $\Lambda=\Lambda_1$ when $m=0$, and $\Lambda=(0.77\ldots)\Lambda_1$ when $m=\frac{1}{2}$. The limitation here on the range of Λ is apparently due to the choice of functions used to express the solution.

In order that $\hat{w}$ and $\tilde{w}$ can be used as lower solutions of (1.4), we must make an appropriate selection of Λ, a and b. An obvious choice which corresponds to the two cases of Theorem 1 is

$$\begin{aligned} &\text{(i)} && \Lambda=\lambda f'(0); \quad a=f'''(0)/6f'(0), && f''(0)=0, \quad f'''(0)>0. \\ &\text{(ii)} && \Lambda=\lambda f'(0); \quad b=f''(0)/2f'(0), && f''(0)>0. \end{aligned} \tag{2.23}$$

Under (2.23) we find that as $m\to 0$, $\lambda\to\lambda_1$ while $\hat{w}(x)$ and $\tilde{w}(x)$ have precisely the bifurcation behavior indicated in (2.4) and (2.5), respectively. That is, for λ near λ_1, $u(x)\sim\hat{w}(x)$ in case (i) and $u(x)\sim\tilde{w}(x)$ in case (ii).

In case (i), we have from (1.4), (2.15) and (2.23) that

$$A\hat{w}-\hat{w}=\lambda\int_0^l g(x|\xi)\left[f(\hat{w}(\xi))-f'(0)\hat{w}(\xi)-\frac{f'''(0)}{6}\hat{w}^3(\xi)\right]d\xi. \tag{2.24}$$

Analogously in case (ii), we have from (1.4), (2.20) and (2.23) that

$$A\tilde{w}-\tilde{w}=\lambda\int_0^l g(x|\xi)\left[f(\tilde{w}(\xi))-f'(0)\tilde{w}(\xi)-\frac{f''(0)}{2}\tilde{w}^2(\xi)\right]d\xi. \tag{2.25}$$

In either case if the integral is nonnegative, we clearly have $A\hat{w}\geqq\hat{w}$ or $A\tilde{w}\geqq\tilde{w}$ and hence a lower solution. This provides the basis for the following results.

THEOREM 5. *Let* $f(z) \leqq \beta$, $z \geqq 0$ *with* $f''(0) = 0$ *and* $f'''(0) > 0$. *Suppose some* $\lambda^* > 0$ *can be found such that*

$$f(\hat{w}) \geqq f'(0)\hat{w} + \frac{f'''(0)}{6}\hat{w}^3 \quad \text{and} \quad \hat{w} \leqq \frac{\beta\pi^2}{8f'(0)}, \qquad 0 \leqq x \leqq l, \tag{2.26}$$

for all λ, $\lambda^* \leqq \lambda < \lambda_1$. *Then there exists a left-branching, nonnegative and continuous solution* $u(x)$ *to* (1.4) *satisfying*

$$\hat{w}(x) \leqq u(x) \leqq \frac{\beta\pi^2}{8f'(0)}, \qquad 0 \leqq x \leqq l, \tag{2.27}$$

for all λ, $\lambda^* \leqq \lambda < \lambda_1$.

THEOREM 6. *Let* $f(z) \leqq \beta$, $z \geqq 0$ *with* $f''(0) > 0$. *Suppose some* $\lambda^* > 0$ *can be found such that*

$$f(\tilde{w}) \geqq f'(0)\tilde{w} + \frac{f''(0)}{2}\tilde{w}^2 \quad \text{and} \quad \tilde{w} \leqq \frac{\beta\pi^2}{8f'(0)}, \qquad 0 \leqq x \leqq l, \tag{2.28}$$

for all λ, $\lambda^* \leqq \lambda < \lambda_1$. *Then there exists a left-branching, nonnegative and continuous solution* $u(x)$ *to* (1.4) *satisfying*

$$\tilde{w}(x) \leqq u(x) \leqq \frac{\beta\pi^2}{8f'(0)}, \qquad 0 \leqq x \leqq l, \tag{2.29}$$

for all λ, $\lambda^* \leqq \lambda < \lambda_1$.

The proof of Theorems 5 and 6 follow as an application of Theorem 4. In (2.12), $U = \beta\pi^2/8f'(0)$ was established as an upper solution. A lower solution, $\hat{w}$ in Theorem 5 and $\tilde{w}$ in Theorem 6, follows from (2.24) and (2.25) when the integrand is nonnegative. This is in fact provided by the conditions (2.26) and (2.27), respectively.

In practice, λ^* is determined as a constraint under which both inequalities in (2.26) or in (2.28) will hold. These conditions are trivially satisfied as $m \to 0$, since $\lambda \to \lambda_1$, while $\hat{w} \to 0$ and $\tilde{w} \to 0$. When some m^*, $0 < m^* < \frac{1}{2}$, can be found such that (2.26) or (2.28) holds, then (2.17) or (2.22), respectively, yields the desired $\lambda^* < \lambda_1$.

The results of Theorems 1, 2, 3 together with either Theorem 5 or 6 yield considerable insight into the nature of the bifurcation diagram, Fig. 1. The local branching behavior near λ_1 is given by Theorem 1. Theorems 2 and 3 yield λ^{**} and U to limit the leftward and upward extent of the branch. Theorem 5 or 6 is intended to provide λ^*, and hence bound the leftward extent of the branch by (1.5).

Finally, it is worthwhile to comment again that many of the conditions imposed on $f(z)$ are only sufficient to obtain the stated results. In many places they can be replaced by weaker conditions. In individual problems, there may also be other, simpler conditions which yield the equivalent results.

3. Application to a nonlinear buckling problem. As an example of a left-branching bifurcation problem, we will examine a model problem for the buckling of a pin-ended slender rod subjected to end loading. The rod material is assumed to have certain nonlinear compressibility properties. The problem considered here is a particular case from a class of buckling problems examined by

Stakgold [9] and Olmstead and Mescheloff [5]. The work in [5] and [9] is based upon a general theory of rods due to Antman [1].

The boundary value problem is posed in terms of the tangent angle $u(x)$ of the deformed rod as a function of the material points x on the rod in its originally straight position. We will consider the example

$$u''(x)+\frac{P}{k}e^{F[P\cos u(x)]-F(0)}\sin u(x)=0, \qquad F(z)\equiv\tfrac{1}{5}\tanh[5(3-5z+2z^2)], \tag{3.1}$$

$$u'(0)=u'(2)=0, \qquad 0\leqq u(0)<\pi. \tag{3.2}$$

Here $P>0$ is the magnitude of the end load, and for convenience we set the original length as $l=2$. The bending law has been assumed linear in this example with $k>0$ as a measure of the stiffness. The compressibility effects are reflected by the exponential function; while the sine function arises from purely geometrical aspects of the formulation.

As it stands, the problem (3.1)–(3.2) is not quite the type to which the procedure of § 2 can be applied. First, the boundary conditions are different, so that any nontrivial solution is somewhere negative. This is easily remedied by considering only the first buckled state of the rod which has a shape that is symmetric with respect to the midpoint. Thus, the first buckled state can be examined for $u(x)\geqq 0$ on $0\leqq x\leqq 1$ with $u(1)=0$. The other difficulty is that the physically natural choice of P as the bifurcation parameter puts this problem into a slightly more complicated class than (1.1)–(1.2). To avoid this technicality, we will examine the problem from another viewpoint. The end load will be fixed at $P=1$, and the stiffness k of the rod material is allowed to vary. Then with $\lambda = k^{-1}P\exp[-F(0)]=(0.818\ldots)k^{-1}$, we can reformulate (3.1)–(3.2) as an appropriate integral equation for the tangent angle $u(x)$ on one-half of the rod. This gives

$$u(x)=\lambda\int_0^1 g(x|\xi)\,e^{F[\cos u(\xi)]}\sin u(\xi)\,d\xi\equiv Au(x). \tag{3.3}$$

Here $g(x|\xi)=1-\xi+(\xi-x)H(x-\xi)\geqq 0$.

With the buckling problem expressed in the form of (3.3), we can follow the general procedure of § 2. For the local branching behavior near the smallest eigenvalue $\lambda_1=\pi^2/4$, case (i) of Theorem 1 applies. The nonnegative solution has the form

$$u(x)=2\left[\left(1-\frac{\lambda}{\lambda_1}\right)\right]^{1/2}\cos\frac{\pi x}{2}+o\left[\left(1-\frac{\lambda}{\lambda_1}\right)^{1/2}\right] \tag{3.4}$$

when λ is near λ_1.

It is easily shown (cf. [5]) that all solutions of physical interest for this type of buckling problem are bounded in absolute value by π. Thus the nonnegative solution which branches from λ_1 must satisfy

$$0\leqq u(x)\leqq\pi. \tag{3.5}$$

This ad hoc result does provide a bound on the upward extent of the solution branch. We cannot however use π as an upper solution of (3.3).

To limit the leftward extent of the branch, we can follow Theorem 2 to find that

$$u(x) \equiv 0, \qquad 0 < \lambda < \lambda^{**} = \lambda_1 e^{-F(0)} = (0.818\ldots)\lambda_1. \tag{3.6}$$

In order to determine some λ^* to which the solution brand must extend, we will employ the results of Theorem 5. Since this does require that the operator A in (3.3) be monotone, we must necessarily restrict our analysis to the cone of nonnegative and continuous functions truncated by $\pi/2$.

In this truncated cone, we in fact find that $V = \pi/2$ is an upper solution of (3.3). That is, for all $\lambda \leqq \lambda_1$,

$$A\frac{\pi}{2} = \lambda e^{F(0)} \int_0^1 g(x|\xi)\, d\xi \leqq \frac{\lambda_1 e^{(1/5)\tanh(15)}}{2} < \frac{\pi}{2}. \tag{3.7}$$

To construct a lower solution of (3.3) we consider $\hat{w}(x)$, the solution of (2.13)–(2.14) with $\Lambda = \lambda$ and $a = 1/3$. Then, to satisfy (2.26) we require that

$$\hat{w}(x) \leqq \frac{\pi}{2}, \qquad 0 \leqq x \leqq 1, \tag{3.8}$$

and

$$f(\hat{w}) \equiv e^{(1/5)\tanh[5(3-5\cos\hat{w}+2\cos^2\hat{w})]} \sin\hat{w} \geqq \hat{w} + \tfrac{1}{3}\hat{w}^3, \qquad 0 \leqq x \leqq 1. \tag{3.9}$$

To have (3.9) hold, it is sufficient to require that

$$0 \leqq \hat{w}(x) \leqq 0.550, \qquad 0 \leqq x \leqq 1. \tag{3.10}$$

We see that (3.10) is more restrictive than (3.8). To satisfy (3.10), we refer to (2.16) and determine that we must take $m \leqq 0.0457\ldots$. In turn from (2.17) this implies that (3.10) holds, and hence the solution branch exists for

$$\lambda \geqq \lambda^* = (0.919\ldots)\lambda_1. \tag{3.11}$$

Thus we conclude from (3.6) and (3.11) that the leftward extent of the nonnegative solution branch is bounded by

$$(0.818\ldots)\lambda_1 \leqq \lambda' \leqq (0.919\ldots)\lambda_1. \tag{3.12}$$

REFERENCES

[1] S. ANTMAN, *The theory of rods*, Handbuch der Physik VI a/2, Springer-Verlag, New York, 1972.

[2] H. B. KELLER AND D. S. COHEN, *Some positone problems suggested by nonlinear heat generation*, J. Math. Mech., 16 (1967), pp. 1361–1376.

[3] J. B. KELLER, *Bifurcation theory for ordinary differential equations*, Bifurcation Theory and Nonlinear Eigenvalue Problems, J. B. Keller and S. Antman, eds., W. A. Benjamin, New York, 1969.

[4] M. A. KRANSNOSEL'SKII, *Positive Solutions of Operator Equations*, P. Noordhoff, Groningen, the Netherlands, 1964.

[5] W. E. OLMSTEAD AND D. J. MESCHELOFF, *Buckling of a nonlinear elastic rod*, J. Math. Anal. Appl., 46 (1974), pp. 609–634.

[6] P. H. RABINOWITZ, *Nonlinear Sturm–Liouville problems for second order ordinary differential equations*, Comm. Pure Appl. Math., 23 (1970), pp. 939–961.

[7] D. H. SATTINGER, *Monotone methods in nonlinear elliptic and parabolic boundary value problems*, Indiana Univ. Math. J., 21 (1972), pp. 979–1000.

[8] ———, *Topics in Stability and Bifurcation Theory*, Lecture Notes in Mathematics No. 309, Springer-Verlag, New York, 1973.

[9] I. STAKGOLD, *Branching of solutions of nonlinear equations*, SIAM Rev., 13 (1971), pp. 289–332.

[10] J. H. WOLKOWISKY, *Nonlinear Sturm–Liouville problems*, Arch. Rational Mech. Anal., 35 (1970), pp. 299–320.

SIAM J. MATH. ANAL.
Vol. 8, No. 3, May 1977

ON AN INTEGRAL TRANSFORM OCCURRING IN THE THEORY OF DIFFRACTION*

D. NAYLOR†

Abstract. This paper considers an integral transform adapted to the solution of certain boundary value problems connected with the Helmholtz equation in cylindrical or spherical polar coordinates when the radial variable r varies over some infinite interval $0<a\leqq r<\infty$. At infinity a radiation type limiting condition is imposed. A formula of inversion is derived which does not involve any summability factor, despite the singular nonself-adjoint nature of the expansion problem.

1. Introduction. In a previous paper [3] the author considered the problem of finding a formula of inversion for the integral transform defined by the equation

$$G(u)=\int_a^\infty [J_u(kr)H_u^{(1)}(ka)-J_u(ka)H_u^{(1)}(kr)]f(r)\frac{dr}{r}, \tag{1}$$

where $a>0$, $k>0$ in which the function $f(r)$ is supposed to belong to a certain class of functions which at infinity satisfy a radiation condition

$$\lim_{r\to\infty} r^{1/2}[f'(r)-ikf(r)]=0. \tag{2}$$

The actual formula obtained in [3] is

$$f(r)=\frac{1}{2}\lim_{\lambda\to 0}\int_W \frac{uH_u^{(1)}(kr)G(u)\,du}{H_u^{(1)}(ka)\cos(\lambda u^{3/2})}. \tag{3}$$

The trigonometric function appearing in this formula is a summability factor, the parameter λ tending to zero through positive values. The path W denotes the wedge $\arg u=\pm\psi$ in the complex u-plane, the angle ψ being chosen small enough to ensure that none of the zeros u_n of $H_u^{(1)}(ka)$, regarded as a function of the order u, lie inside W. The form of the above formula raises the question of the possibility of the existence of an alternative formula of inversion not involving a summability factor. A formula of this type, which is useful in connection with the solution of certain problems associated with the Helmholtz equation, is developed in this paper. Although this formula can be used to construct the kind of expansion involving the eigenfunctions $H_{u_n}^{(1)}(kr)$ which appear in such problems, it is not especially adapted to derive this expansion, and in fact it generates an expansion of a different form.

The formula in question together with a set of conditions sufficient to ensure its validity are stated in the following theorem.

THEOREM. *Suppose that $f(r)$ is twice continuously differentiable for $r\geqq a$, $r^{-1/2}(rf_{rr}+f_r+k^2rf)\in L(a,\infty)$, $\lim_{r\to\infty} r^{1/2}f(r)\,e^{-ikr}$ exists and $\lim_{r\to\infty} r^{1/2}(f_r-ikf)=0$, where k is real and positive. Let $G(u)$ be defined by* (1); *then, if $r>a$,*

$$f(r)=\frac{1}{2}\int_L \frac{uJ_{-u}(kr)G(u)\,du}{J_{-u}(ka)}+i\pi\sum_{u=u_n'}\frac{uJ_{-u}(kr)G(u)}{(\partial/\partial u)J_{-u}(ka)}, \tag{4}$$

* Received by the editors June 16, 1975, and in revised form October 10, 1975.

† Department of Applied Mathematics, University of Western Ontario, London, Ontario, Canada.

where L denotes the imaginary axis of the complex u-plane and the summation is extended over all of the positive zeros u_n' of the function $J_{-u}(ka)$, regarded as a function of u.

2. Proof of the theorem. A proof of the expansion (4) can be obtained by following the procedure developed in [4] to derive a related expansion theorem. Let $f(r)$, the function to be expanded, satisfy the conditions of the theorem and let the function $g(r)$ be defined by means of the equation

$$r^2 f_{rr} + r f_r + (k^2 r^2 - v^2) f = g(r), \tag{5}$$

where $r > a$ and v is some positive number which is not a zero of $J_{-u}(ka)$. It is known [1] that there is an infinite number of such zeros, all real and simple, and [5] that they all lie in the interval $-ka \leqq u < \infty$. Furthermore, the large u-zeros are asymptotic to the positive integers.

The equation (5) is now regarded as a nonhomogeneous differential equation for $f(r)$, which will be inverted in terms of a suitable Green's function. Since all the Bessel functions of real argument are $O(r^{-1/2})$ as $r \to \infty$, there are infinitely many Green's functions that can be used for this purpose, each giving rise to a different representation or expansion formula for $f(r)$. In the author's previous discussion of this problem, that Green's function was chosen which satisfied the radiation condition at infinity. The resulting Green's function contained the Hankel function $H_u^{(1)}(kr)$ as one of its factors, and this gave rise to the expansion involving the eigenfunctions $H_{u_n}^{(1)}(kr)$.

In this paper a different Green's function is adopted, as defined by the equations,

$$G(r,\rho) = \begin{cases} \dfrac{-\pi \Phi_v(k,r) J_{-v}(k\rho)}{2J_{-v}(ka)}, & a \leqq r \leqq \rho, \\ \dfrac{-\pi \Phi_v(k,\rho) J_{-v}(kr)}{2J_{-v}(ka)}, & a \leqq \rho \leqq r, \end{cases} \tag{6}$$

where

$$\Phi_v(k,r) = \frac{J_v(kr)J_{-v}(ka) - J_v(ka)J_{-v}(kr)}{\sin v\pi}. \tag{7}$$

The equation (5) may now be inverted to yield the formula

$$\begin{aligned} f(r) = f(a)\frac{J_{-v}(kr)}{J_{-v}(ka)} + \int_a^\infty G(r,\rho) g(\rho) \frac{d\rho}{\rho} \\ + \lim_{\rho \to \infty} [\rho f(\rho) G_\rho(r,\rho) - \rho f'(\rho) G(r,\rho)]. \end{aligned} \tag{8}$$

The value of the limit appearing in the preceding equation can be obtained as follows. Upon substituting the Hankel asymptotic expressions for the Bessel functions of large argument, it follows that

$$\lim_{r \to \infty} [r f'(r) J_v(kr) - kr f(r) J_v'(kr)] = c\, e^{iv\pi/2}, \tag{9}$$

where

$$c = i(2k/\pi)^{1/2}\, e^{i\pi/4} \lim_{r\to\infty} [r^{1/2} f(r)\, e^{-ikr}]. \tag{10}$$

The limit in (8) may be determined by substituting the first of the expressions appearing in (6) and using the result (9). This gives the formula

$$f(r) = \int_a^\infty G(r,\rho) g(\rho) \frac{d\rho}{\rho} + f(a) \frac{J_{-v}(kr)}{J_{-v}(ka)} + \frac{c\pi \Phi_v(k,r)\, e^{-iv\pi/2}}{2J_{-v}(ka)}. \tag{11}$$

To obtain the desired formulas, it is necessary to represent the Green's function defined by the composite expression (6) by means of a single formula, which will be substituted in (11). The required formula is given by the equation

$$\begin{aligned} G(r,\rho) = {} & \frac{1}{2i} \int_L \frac{u \Phi_u(k,\rho) J_{-u}(kr)\, du}{(u^2 - v^2) J_{-u}(ka)} \\ & - \pi \sum_{u=u_n'} \frac{u J_u(ka) J_{-u}(kr) J_{-u}(k\rho)}{(u^2 - v^2) \sin u\pi (\partial/\partial u)[J_{-u}(ka)]}. \end{aligned} \tag{12}$$

In this formula, L denotes the imaginary axis of the complex u-plane and the summation includes all those zeros u_n' of the function $J_{-u}(ka)$ which are located in the half-plane $\mathrm{Re}\,(u) > 0$. The validity of (12) may be demonstrated with the aid of the calculus of residues, and for this purpose, it is necessary to determine the behavior of the integrand when u is large. This may be estimated with the asymptotic expression

$$J_u(x) = \frac{(x/2)^u}{\Gamma(u+1)} [1 + O(u^{-1})], \tag{13}$$

which holds when u is large compared with x and bounded away from the negative integers. Since the zeros of $J_{-u}(ka)$ tend as $u \to \infty$ to the (large) positive integers, it is convenient to close the contour of the integral in (12) in the right-hand half-plane by means of a sequence of semicircles C_n of radii $(n + \frac{1}{2})$ and let $n \to \infty$. The asymptotic behavior of the function Φ_u defined by (7) can be obtained after estimating the Bessel functions by means of the formula (13) and then using the identity $\Gamma(1+u)\Gamma(1-u) \sin u\pi = u\pi$. This procedure leads to the estimate

$$\Phi_u(k,\rho) = \frac{1}{u\pi} [(\rho/a)^u - (a/\rho)^u][1 + O(u^{-1})]. \tag{14}$$

The expressions (13), (14) show that on the semicircle C_n, the integrand in (12) is $O[u^{-2}(\rho/r)^t]$ where $u = t + is$, $t > 0$, which tends to zero as $u \to \infty$ provided that $a \leqq \rho \leqq r$. For such values of r, ρ the contour integral in (12) may be evaluated by closing the contour on the right and taking the residue at the pole $u = v$ and at all the zeros u_n' lying to the right of L. When this procedure is carried out, the series appearing in (12) cancels out with the series of corresponding residues, and the formula (12) for $G(r,\rho)$ reduces to the second of the two expressions on the right-hand side of (6). The validity of the representation (12) for the complementary range $a \leqq r \leqq \rho$ cannot be established directly by the above method since for

such values of r, ρ the integrand does not tend to zero as $u \to \infty$, except on L itself. In this case, the formula can be verified by decomposing the integral into two parts, corresponding to the two terms in (7), and evaluating the resulting integrals by closing the contour on opposite sides of L as appropriate. Since this procedure introduces the residues at the zeros of the function $\sin u\pi$, it is shorter to proceed indirectly as follows by verifying that the expression proposed in (12) is a symmetric function of r, ρ. By means of (12), (7), it follows that

$$G(r,\rho)-G(\rho,r)=\frac{1}{2i}\int_L \frac{u[J_u(k\rho)J_{-u}(kr)-J_u(kr)J_{-u}(k\rho)]\,du}{(u^2-v^2)\sin u\pi}.$$

The integral appearing on the right-hand side of the above equation is zero since the integrand is an odd function of u. Thus the value of the expression on the right-hand side of (12) when $r<\rho$ may be found by interchanging r, ρ in (12) and evaluating the resulting integral as before, a procedure which leads to the first of the two expressions stated in (6). Thus (12) is established for all values of r, ρ.

3. The integral theorem. The integral formulas sought are obtained by inserting the expression (12) for the Green's function into the formula (11) for $f(r)$. This procedure leads, after an interchange in the order of integration in the repeated integral and of integration and summation in the series, to the equation

$$\begin{aligned} f(r)=&\frac{1}{2i}\int_L \frac{uJ_{-u}(kr)\,du}{(u^2-v^2)J_{-u}(ka)}\int_a^\infty \Phi_u(k,\rho)g(\rho)\frac{d\rho}{\rho}\\ &-\pi\sum_{u=u_n'}\frac{uJ_u(ka)J_{-u}(kr)}{(u^2-v^2)\sin u\pi[\partial/\partial u J_{-u}(ka)]}\int_a^\infty J_{-u}(k\rho)g(\rho)\frac{d\rho}{\rho}\\ &+f(a)\frac{J_{-v}(kr)}{J_{-v}(k\rho)}+\frac{c\pi\Phi_v(k,r)\,e^{-iv\pi/2}}{2J_{-v}(ka)}. \end{aligned} \tag{15}$$

To justify the above formula, it is necessary to verify that the repeated integral and the series are absolutely convergent, and for this purpose a suitable bound must be found for the ρ-integral appearing in (15). Such a bound is obtained in the Appendix where it is shown, equation (A.8), that on the imaginary axis where $u = is$ (s real),

$$\int_a^\infty \left|\Phi_{is}(k,\rho)g(\rho)\right|\left|\frac{d\rho}{\rho}\right| \leqq \frac{C|J_{is}(ka)|}{|\sinh(s\pi/2)|}, \tag{16}$$

where C is a constant. The repeated integral in (15) is in absolute magnitude less than

$$C\int_{-\infty}^{\infty}\frac{|s|\cdot|J_{-is}(kr)|\,ds}{(s^2+v^2)\sinh(|s\pi|/2)}.$$

This integral is absolutely convergent since for s real, $|\Gamma(1-is)| = [\pi s/\sinh(\pi s)]^{1/2}$ so by (13), $|J_{-is}(kr)| = O[\Gamma(1-is)^{-1}] = O\{|s^{-1}\sinh(\pi s)|^{1/2}\}$ as $s\to\infty$.

To discuss the convergence of the series appearing on the right-hand side of (15) care is needed since the large u-zeros of $J_{-u}(ka)$ tend to the positive integers

so that $\sin u\pi \to 0$. The necessary bounds and inequalities that are needed to discuss this series can be deduced from certain representations of products of Bessel functions as integrals. These bounds are obtained in the Appendix to this paper where it is proved that, if u is a zero (positive) of $J_{-u}(ka)$,

$$\left|J_u(ka)\frac{\partial}{\partial u}J_{-u}(ka)\right| \geqq \frac{1}{2\pi}|\sin u\pi|(2/ka)^{2u}[\Gamma(u)]^2,$$

$$|J_u(ka)J_{-u}(kr)| \leqq \left|\frac{1}{\pi u}\sin u\pi\right|(r/a)^u,$$

$$\left|\int_a^\infty J_{-u}(kr)g(r)\frac{dr}{r}\right| = O(u^{3/2}) \quad \text{as } u \to \infty.$$

Also, if u is a zero of $J_{-u}(ka)$, then from (A.10) of the Appendix, it follows that

$$J_u(ka) = Y_u(ka)\tan u\pi \sim -\frac{1}{\pi}(2/ka)^u\Gamma(u)\tan u\pi,$$

as $u \to \infty$. On collecting these results, it follows that for u large, the corresponding term in the series in (15) is

$$O\left[\frac{u^{1/2}(kr/2)^u}{\Gamma(u+1)}\right].$$

Since the u-zeros are asymptotic to the positive integers n, it follows that the series in (15) is absolutely convergent, so that (15) is itself established.

To obtain the form of the expansion theorem quoted in the theorem, it is necessary to insert the following expression:

$$\int_a^\infty \Phi_u(k,r)g(r)\frac{dr}{r} = (u^2-v^2)\int_a^\infty \Phi_u(k,r)f(r)\frac{dr}{r} + \frac{2}{\pi}f(a) + \frac{c[e^{iu\pi/2}J_{-u}(ka) - e^{-iu\pi/2}J_u(ka)]}{\sin u\pi}. \tag{17}$$

This formula can be obtained from (5) by multiplying by $\Phi_u(k,r)$ and integrating by parts. If u is a zero of $J_{-u}(ka)$, then (17) reduces on multiplication by $-\sin u\pi/J_u(ka)$ to

$$\int_a^\infty J_{-u}(kr)g(r)\frac{dr}{r} = (u^2-v^2)\int_a^\infty J_{-u}(kr)f(r)\frac{dr}{r} - \frac{2}{\pi}f(a)\frac{\sin u\pi}{J_u(ka)} + c\,e^{-iu\pi/2}. \tag{18}$$

Upon inserting (17), (18) into (15), there results the equation

$$f(r) = \frac{1}{2i}\int_L \frac{uJ_{-u}(kr)\,du}{J_{-u}(ka)}\int_a^\infty \Phi_u(k,\rho)f(\rho)\frac{d\rho}{\rho} - \pi\sum_{u=u_n'}\frac{uJ_u(ka)J_{-u}(kr)}{\sin u\pi(\partial/\partial u)J_{-u}(ka)}\int_a^\infty J_{-u}(k\rho)f(\rho)\frac{d\rho}{\rho} + Af(a) + Bc, \tag{19}$$

where

$$A=\frac{J_{-v}(kr)}{J_{-v}(ka)}+\frac{1}{i\pi}\int_L \frac{uJ_{-u}(kr)\,du}{(u^2-v^2)J_{-u}(ka)}+2\sum_{u=u_n'}\frac{uJ_{-u}(kr)}{(u^2-v^2)(\partial/\partial u)J_{-u}(ka)},$$

$$B=\frac{\pi\Phi_v(k,r)\,e^{-iv\pi/2}}{2J_{-v}(ka)}+\frac{1}{2i}\int_L \frac{uJ_{-u}(kr)\,e^{+iu\pi/2}\,du}{(u^2-v^2)\sin u\pi} \tag{20}$$

$$-\frac{1}{2i}\int_L \frac{uJ_{-u}(kr)J_u(ka)\,e^{-iu\pi/2}\,du}{(u^2-v^2)\sin u\pi J_{-u}(ka)}$$

$$-\pi\sum_{u=u_n'}\frac{uJ_u(ka)J_{-u}(kr)\,e^{-iu\pi/2}}{(u^2-v^2)\sin u\pi(\partial/\partial u)J_{-u}(ka)}.$$

It will now be shown that $A=B=0$.

The integral appearing in the expression for A can be evaluated by closing the contour on the right by means of a sequence of semicircles of radii $(n+\frac{1}{2})$ where n is a positive integer which tends to infinity. On such a sequence of semicircles, the integrand is $O[(a/r)^n]$ which tends to zero as $n\to\infty$. Upon evaluating the integral by taking the residues at those poles of the integrand which are positioned in the half-plane Re $(u)>0$, it is found that $A=0$ as required.

The quantity B can be shown to be equal to zero by a similar argument, the first integral in (20) being evaluated by closing the contour on the left and the second integral being evaluated by closing the contour on the right.

Since

$$J_{-u}(x)=J_u(x)\,e^{-iu\pi}+iH_u^{(1)}(x)\sin u\pi, \tag{21}$$

the function $G(u)$ defined by (1) can be expressed by means of the equation

$$G(u)=\frac{1}{i\sin u\pi}\int_a^\infty [J_u(kr)J_{-u}(ka)-J_u(ka)J_{-u}(kr)]f(r)\frac{dr}{r}$$

$$=-i\int_a^\infty \Phi_u(k,r)f(r)\frac{dr}{r}.$$

If u is a zero of $J_{-u}(ka)$, the above expression reduces to

$$G(u)=-\frac{J_u(ka)}{i\sin u\pi}\int_a^\infty J_{-u}(kr)f(r)\frac{dr}{r}.$$

When the above substitutions are made in (19), it is found that this formula reduces to the formula (4) stated in the theorem.

Appendix. It remains to derive the bounds used in the paper.

(i) A suitable bound on the product $J_u(ka)J_{-u}(kr)$ valid when u is a zero of

$J_{-u}(ka)$ can be deduced from the following formula due to Dixon and Ferrar [2, p. 207]:

$$J_u(X)J_{-u}(x)=\frac{1}{\pi}\int_0^{\pi} J_0[\sqrt{X^2+x^2+2Xx\cos\theta}]\cos u\theta\,d\theta$$
$$+\frac{1}{\pi}\sin u\pi\int_1^{X/x}\rho^{u-1}J_0[\sqrt{X^2+x^2-Xx(\rho+\rho^{-1})}]\,d\rho$$

for $X\geqq x$. It follows from this result that

$$J_u(X)J_{-u}(x)-J_{-u}(X)J_u(x)$$
$$=\frac{2\sin u\pi}{\pi}\int_0^{\ln(X/x)} J_0[\sqrt{X^2+x^2-2Xx\cosh\theta}]\cosh u\theta\,d\theta.$$

Since $|J_0(x)|\leqq 1$ for any real x and $|\cosh u\theta|\leqq\cosh t\theta$ where $u=t+is$, then it follows that, for $t>0$,

$$|J_u(X)J_{-u}(x)-J_u(x)J_{-u}(X)|\leqq\left|\frac{\sin u\pi}{\pi t}\right|[(X/x)^t-(x/X)^t]. \tag{A.1}$$

In this formula X, x are replaced by kr, ka, respectively, and u is taken to be a zero of $J_{-u}(ka)$. Since all such zeros are real, it follows that

$$|J_u(ka)J_{-u}(kr)|\leqq\left|\frac{\sin u\pi}{u\pi}\right|(r/a)^u. \tag{A.2}$$

(ii) The bound required on the quantity $(\partial/\partial u)J_{-u}(ka)$ may be obtained from the following formulas of Watson [6, p. 444]:

$$J_u(x)^2+Y_u(x)^2=\frac{8}{\pi^2}\int_0^{\infty} K_0(2x\sinh\theta)\cosh 2u\theta\,d\theta, \tag{A.3}$$

$$J_u(x)\frac{\partial Y_u(x)}{\partial u}-Y_u(x)\frac{\partial J_u(x)}{\partial u}=-\frac{4}{\pi}\int_0^{\infty} K_0(2x\sinh\theta)\,e^{-2u\theta}\,d\theta.$$

Upon setting $J_{-u}(x)=J_u(x)\cos u\pi-Y_u(x)\sin u\pi$, it is found after some reduction that

$$J_u(x)\frac{\partial J_{-u}(x)}{\partial u}-J_{-u}(x)\frac{\partial J_u(x)}{\partial u}=-\pi Y_u(x)J_{-u}(x)$$
$$-\frac{4}{\pi}\sin u\pi\int_0^{\infty} K_0(2x\sinh\theta)\,e^{2u\theta}\,d\theta. \tag{A.4}$$

Now $\cosh\theta \leqq e^{\theta}$ and $2\sinh\theta \leqq e^{\theta}$ for $\theta \geqq 0$ so that, if $u > \frac{1}{2}$,

$$\begin{aligned}\int_0^\infty K_0(2x\sinh\theta)\,e^{2u\theta}\,d\theta &= \int_0^\infty K_0(2x\sinh\theta)\,e^{(2u-1)\theta}\,e^{\theta}\,d\theta \\ &\geqq \int_0^\infty K_0(2x\sinh\theta)(2\sinh\theta)^{2u-1}\cosh\theta\,d\theta \\ &= \frac{1}{2}x^{-2u}\int_0^\infty K_0(y)y^{2u-1}\,dy \\ &= \frac{1}{8}(2/x)^{2u}[\Gamma(u)]^2,\end{aligned} \tag{A.5}$$

by [6, p. 388]. If we set $x = ka$ in (A.4), where u is a zero of $J_{-u}(ka)$, and use the inequality (A.5), we find that

$$\left|J_u(ka)\frac{\partial}{\partial u}J_{-u}(ka)\right| \geqq \frac{|\sin u\pi|}{2\pi}(2/ka)^{2u}[\Gamma(u)]^2, \tag{A.6}$$

whenever $J_{-u}(ka) = 0$.

(iii) Finally, it is necessary to establish suitable bounds on the function

$$\int_a^\infty \Phi_u(k, r)g(r)\frac{dr}{r}.$$

A bound valid on the imaginary axis where $u = is$ (s real) may be found from Watson's formula (A.3) which can be written as

$$\begin{aligned}e^{-s\pi}|H_{is}^{(1)}(x)|^2 &= \frac{8}{\pi^2}\int_0^\infty K_0(2x\sinh\theta)\cos 2s\theta\,d\theta \\ &\leqq \frac{8}{\pi^2}\int_0^\infty K_0(2x\sinh\theta)\cosh\theta\,d\theta \\ &= \frac{2}{\pi x}.\end{aligned}$$

Since $2J_{is}(x) = H_{is}^{(1)}(x) + \overline{H_{-is}^{(1)}(x)}$ and $H_{-is}^{(1)}(x) = e^{-s\pi}H_{is}^{(1)}(x)$, it follows from the above inequality that

$$|J_{is}(x)| \leqq (2/x\pi)^{1/2}\cosh(s\pi/2), \tag{A.7}$$

so that the function Φ defined by (7) is such that

$$|\Phi_{is}(k, r)| \leqq [2/(\pi kr)]^{1/2}|J_{is}(ka)\,\operatorname{csch}(s\pi/2)|.$$

It follows that

$$\int_a^\infty \left|\Phi_{is}(k, r)g(r)\right|\frac{dr}{r} \leqq (2/\pi k)^{1/2}\left|J_{is}(ka)\,\operatorname{csch}\left(\frac{s\pi}{2}\right)\right|\int_a^\infty |g(r)|\frac{dr}{r^{3/2}}. \tag{A.8}$$

Since $r^{-3/2}g(r) \in L(a, \infty)$, the inequality (16) is established.

A bound valid when u is a zero of $J_{-u}(ka)$ can be obtained with the aid of (18).

First it is noted that if u is a zero of $J_{-u}(ka)$, then the identities

$$J_u(ka)J'_{-u}(ka)-J_{-u}(ka)J'_u(ka)=-\frac{2\sin u\pi}{\pi ka}, \tag{A.9}$$

$$J_{-u}(ka)=J_u(ka)\cos u\pi - Y_u(ka)\sin u\pi \tag{A.10}$$

show that

$$kaJ'_{-u}(ka)=-\frac{2\sin u\pi}{\pi J_u(ka)}=-\frac{2\cos u\pi}{\pi Y_u(ka)}.$$

This last expression tends to zero as $u\to\infty$ since $Y_u(x)\sim-(1/\pi)(2/x)^u\Gamma(u)$ as $u\to\infty$. It follows from (18) that

$$\int_a^\infty J_{-u}(kr)g(r)\frac{dr}{r}=(u^2-v^2)\int_a^\infty J_{-u}(kr)f(r)\frac{dr}{r}+O(1), \tag{A.11}$$

for $u=u'_n\to\infty$. A bound on the integral involving the function $f(r)$ in (A.11) can be obtained by means of the Schwarz inequality which gives the bound

$$\left|\int_a^\infty J_{-u}(kr)f(r)\frac{dr}{r}\right|\leqq\left[\int_a^\infty J_{-u}(kr)^2\frac{dr}{r}\cdot\int_a^\infty|f(r)|^2\frac{dr}{r}\right]^{1/2}. \tag{A.12}$$

Now the Bessel function integral present in the above inequality can be obtained from Watson [6, p. 135] which gives, when u is a zero of $J_{-u}(ka)$, the formula

$$\begin{aligned}2u\int_a^\infty J_{-u}(kr)^2\frac{dr}{r}&=-1+kaJ'_{-u}(ka)\frac{\partial}{\partial u}J_{-u}(ka)\\&=-1-\frac{2\sin u\pi(\partial/\partial u)J_{-u}(ka)}{\pi J_u(ka)},\end{aligned} \tag{A.13}$$

while from (A.3) and (A.4), since $Y_u(ka)=J_u(ka)\cot u\pi$ (cf. (A.10)),

$$J_u(ka)^2=\frac{8}{\pi^2}\sin^2 u\pi\int_0^\infty K_0(2ka\sinh\theta)\cosh 2u\theta\,d\theta,$$

$$J_u(ka)\frac{\partial}{\partial u}J_{-u}(ka)=-\frac{4}{\pi}\sin u\pi\int_0^\infty K_0(2ka\sinh\theta)\,e^{2u\theta}\,d\theta.$$

Upon substituting these results in the right-hand side of (A.13) and simplifying the resulting expression, it is found that

$$2u\int_a^\infty J_{-u}(kr)^2\frac{dr}{r}=\frac{\int_0^\infty K_0(2ka\sinh\theta)\sinh 2u\theta\,d\theta}{\int_0^\infty K_0(2ka\sinh\theta)\cosh 2u\theta\,d\theta}.$$

Since $\sinh 2u\theta<\cosh 2u\theta$, it follows that

$$\int_a^\infty J_{-u}(kr)^2\frac{dr}{r}<\frac{1}{2u},$$

whenever u is a positive zero of $J_{-u}(ka)$. Upon utilizing this result in (A.12) and then (A.11), it follows that

$$\int_a^\infty J_{-u}(kr)g(r)\frac{dr}{r} = O(u^{3/2}),$$

when u is a large zero of $J_{-u}(ka)$.

REFERENCES

[1] J. COULOMB, *Sur les zéros des fonctions des Bessel considérées comme fontion de l'ordre*, Bull. Sci. Math., 60 (1936), pp. 297–302.

[2] A. L. DIXON AND W. L. FERRAR, *Integrals for the product of two Bessel functions*, Quart. J. Math. Oxford Ser., 4 (1933), pp. 192–208.

[3] D. NAYLOR, *On an integral transform associated with a condition of radiation, Part 2*, Proc. Cambridge Philos. Soc., 77 (1975), pp. 189–197.

[4] ———, *On an eigenfunction expansion associated with a condition of radiation*, Ibid., 67 (1970), pp. 107–121.

[5] ———, *An eigenvalue problem in cylindrical harmonics*, J. Math. and Phys., 44 (1965), pp. 391–402.

[6] G. N. WATSON, *Theory of Bessel Functions*, 2nd Ed., Cambridge University Press, London, 1944.

SIAM J. MATH. ANAL.
Vol. 8, No. 3, May 1977

AN EXPLICIT SOLUTION OF THE CENTRAL CONNECTION PROBLEM FOR AN *n*th ORDER LINEAR ORDINARY DIFFERENTIAL EQUATION WITH POLYNOMIAL COEFFICIENTS*

HERBERT WYRWICH†

Abstract. We consider a differential equation of the form $(-1)^n y^{(n)}(x) = (x^m + Q(x))y(x)$, where $Q(x)$ is a polynomial of degree $m - [m/n] - 2$. For a particular recessive solution $Y(x)$ of this equation, uniquely determined by an asymptotic expansion about $x = \infty$, we derive explicit representations of its n initial values $Y^{(k)}(0)$, $k = 0, \cdots, n-1$. These representations have the form of multiple sums, whose coefficients themselves are hypergeometric sums of several variables. To obtain our results, we apply the Mellin transformation to our differential equation and are led to a system of difference equations, which can be solved explicitly, and whose solution provides the representation of the initial values.

1. Introduction. Solutions of differential equations with polynomial coefficients have recently been studied with special regard to their behavior in the large. A method due to Hsieh and Sibuya [4] and Braaksma [2, pp. 1–15] brings a new idea for solving the central connection problems for such equations (cf. Wasow [8]). In these papers the original connection problem is reduced to the solution of the analogous problems for the components of a recursive system of differential equations.

In [9], we studied the differential equation

$$(-1)^n y^{(n)}(x) = P(x) \cdot y(x), \tag{1.1}$$

where

$$P(x) = x^m + a_1 x^{m-1} + a_2 x^{m-2} + \cdots + a_m \tag{1.2}$$

is a polynomial of degree m. It was our intention to get representations of the first n initial values

$$Y^{(k)}(0), \qquad k = 0, \cdots, n-1, \tag{1.3}$$

for a particular *recessive* solution $Y(x)$ of (1.1). This would, of course, give us knowledge of all Taylor coefficients of $Y(x)$. The central connection problem of $Y(x)$ would then be solved completely. Because of simple symmetry relations, this would imply the solution of all other central connection problems for the equation (1.1) (for the case $n = 2$ cf. [4, p. 87]). In this paper, we consider the special case

$$(-1)^n y^{(n)}(x) = P^*(x) \cdot y(x) \tag{1.4}$$

of (1.1), where

$$P^*(x) = x^m + \sum_{k=[\varkappa]+1}^{m} a_k x^{m-k}, \qquad \varkappa = \frac{m+n}{n}. \tag{1.5}$$

This case is characterized by the absence of parameters in the leading exponential factor of the asymptotic expansion of $Y(x)$, and we shall call it the *undercritical* case of (1.1). In the same sense, we call $\varkappa$ the *critical value* of (1.1) and a

* Received by the editors July 3, 1975.

† Gesamthochschule 41 Duisburg, Mathematik, Lotharstrasse 65, West Germany.

polynomial parameter a_k *undercritical*, if $[\varkappa] < k \leqq m$, and *overcritical* otherwise.

We shall aim at *explicit* representations of the initial values (1.3). This will be achieved by adding some supplementary techniques to the Hsieh–Sibuya–Braaksma (H–S–B)-method.

First we make use of *Mellin* transforms to get a recursive system of difference equations. We then solve this system by summation and determine a special solution so that its components are the Mellin images of the components of the recursive system of differential equations obtained by the H–S–B-method. The image components prove to be meromorphic functions, whose residues are closely interrelated with the coefficients of the power series expansion of the initial values (1.3), considered as entire functions of the undercritical parameters.

We calculate these residues and finally obtain the desired representation of the initial values (1.3) in the form

$$\begin{aligned} &Y^{(k)}(0; a_\lambda, \cdots, a_m) = \sum_{p_\lambda, \cdots, p_m = 0}^{\infty} \rho^{(k)}_{p_\lambda, \cdots, p_m} a_\lambda^{p_\lambda} \cdot \cdots \cdot a_m^{p_m}, \\ &\lambda = [\varkappa] + 1, \qquad k = 0, \cdots, n-1, \end{aligned} \tag{1.6}$$

where the ρ-coefficients are hypergeometric sums of several variables.

2. The Hsieh–Sibuya–Braaksma-method. Our analysis of the central connection problem of (1.4) is based on the following theorem, which is a special case of a theorem of Braaksma [2, pp. 1–15] (cf. also Hsieh and Sibuya [4] for the case $n = 2$).

THEOREM 1. *Let a differential equation*

$$(-1)^n y^{(n)}(x) = P^*(x) \cdot y(x) \tag{2.1}$$

be given, where $2 \leqq n \in \mathbb{N}$,

$$P^*(x) = x^m + a_\lambda x^{m-\lambda} + a_{\lambda+1} x^{m-\lambda-1} + \cdots + a_m, \qquad \lambda = \left[\frac{m}{n}\right] + 2, \tag{2.2}$$

and the a_k *are complex parameters.*

Then (2.1) *possesses a uniquely determined solution*

$$Y(x; a_\lambda, \cdots, a_m) \tag{2.3}$$

such that

(i) $Y(x; a_\lambda, \cdots, a_m)$ *is an entire function of* x *and the parameters* $a_\lambda, \cdots, a_m$,

(ii) $Y(x; a_\lambda, \cdots, a_m)$ *admits the asymptotic representation*

$$Y(x; a_\lambda, \cdots, a_m) \approx \exp\left(-\frac{n}{m+n} x^{(m+n)/n}\right) \cdot x^{-(m/(2n))(n-1)} \sum_{k=0}^{\infty} A_k x^{-k/n}, \tag{2.4}$$

with $A_0 = 1$, *uniformly on each compact set of the* $(a_\lambda, \cdots, a_m)$*-space, as* x *tends to infinity in any closed subsector of*

$$S = \left\{x; |\arg x| < \frac{n+1}{n+m}\pi\right\}. \tag{2.5}$$

The A_k are polynomials of $a_\lambda, \cdots, a_m$ and

$$x^r = \exp\{r(\log|x| + i \arg x)\}$$

for any $r \in \mathbb{C}$.

Following [4], we now represent $Y(x; a_\lambda, \cdots, a_m)$ as a power series of $a_\lambda, \cdots, a_m$ with coefficients that are entire functions of x. We have

$$Y(x; a_\lambda, \cdots, a_m) = \sum_{p_\lambda, \cdots, p_m = 0}^{\infty} \eta_{p_\lambda, \cdots, p_m}(x) \cdot a_\lambda^{p_\lambda} \cdot \cdots \cdot a_m^{p_m}, \tag{2.6}$$

or, using vector and multi index notation with $\underline{a} = (a_\lambda, \cdots, a_m), \underline{p} = (p_\lambda, \cdots, p_m)$,

$$Y(x; \underline{a}) = \sum_{\underline{p}=\underline{0}}^{\infty} \eta_{\underline{p}}(x) \cdot \underline{a}^{\underline{p}}. \tag{2.6'}$$

This series is uniformly and absolutely convergent on each compact set of the $(x, a_\lambda, \cdots, a_m)$-space, so we can differentiate (2.6) termwise. Inserting into (2.1) we get the following system of differential equations for the coefficient functions $\eta_{\underline{p}}(x)$:

$$(-1)^n \eta_{\underline{p}}^{(n)}(x) = x^m \eta_{\underline{p}}(x) + \sum_{\substack{j=\lambda \\ p_j \geqq 1}}^{m} \eta_{\underline{p}-\underline{e}_{j-\lambda+1}}(x) \cdot x^{m-j}. \tag{2.7}$$

Here $\underline{e}_\mu$ is the μth unit vector.

Clearly this system is recursive with respect to the *order*

$$|\underline{p}| = p_\lambda + \cdots + p_m \tag{2.8}$$

of the multi index $\underline{p}$.

We now express $\eta_{\underline{p}}(x)$ by Taylor's coefficient formula in the form

$$\eta_{\underline{p}}(x) = \frac{1}{\underline{p}!} D_{\underline{a}}^{\underline{p}} [Y(x, \underline{a})]_{\underline{a}=\underline{0}}, \tag{2.9}$$

where $D_{\underline{a}}^{\underline{p}}$ means the differential operator

$$D_{\underline{a}}^{\underline{p}} = \frac{\partial^{p_\lambda + \cdots + p_m}}{\partial a_\lambda^{p_\lambda} \cdot \cdots \cdot \partial a_m^{p_m}}$$

and

$$\underline{p}! = p_\lambda ! \cdot \cdots \cdot p_m !.$$

Applying a theorem on differentiation of asymptotic expansions with parameters (cf. Wasow [7, Thm. 9.4] and [9, Lemma 2]) to the representation (2.4), we get the asymptotic expansion

$$\eta_{\underline{p}}(x) \approx \frac{1}{\underline{p}!} x^{-(m/(2n))(n-1)} \exp\left(-\frac{n}{m+n} x^{(m+n)/n}\right) \cdot \sum_{k=0}^{\infty} D_{\underline{a}}^{\underline{p}} (A_k)_{\underline{a}=\underline{0}} \cdot x^{-k/n} \tag{2.10}$$

as $x \to \infty$ in any closed subsector of (2.5).

The system (2.7) of differential equations together with the asymptotic expansions (2.10) of its components now constitute a new central connection problem. This problem is parameter-free and can be solved principally by the method of variation of parameters (cf. [4] for the special case $n = 2$).

The disadvantage of this procedure stems from the following fact: For $n > 2$, a fundamental system of solutions for the homogeneous part of (2.7) is no longer available in terms of Bessel functions (e.g., $n = 2$), but in terms of Meijer's G-functions (cf. Braaksma [1]). Therefore, the successive integrations cannot be carried out explicitly. It is evident that an asymptotic analysis of these integrals to determine the unknown n integration constants (by comparing with expansion (2.10)) would be a hopeless task.

We shall circumvent this difficulty by adding a new idea, which provides some supplementary techniques.

3. Mellin transforms and the system of associated difference equations. Let $M_{a,b}$ be the class of functions on $(0, \infty)$, which are summable in the sense of Lebesgue on each compact set of $(0, \infty)$ and which, with $a < b$, satisfy the two boundary conditions

$$F(t) = O(t^{-a}), \qquad t \to 0,$$
$$F(t) = O(t^{-b}), \qquad t \to +\infty.$$

Then for each $F \in M_{a,b}$, the integral

$$f(s) = \int_0^\infty F(t) t^{s-1}\, dt$$

exists in the strip $a < \operatorname{Re} s < b$ and represents there a holomorphic function. We write

$$f(s) = \mathfrak{M}[F, s]$$

and call f the *Mellin transform* ($\mathfrak{M}$-transform) of F and the mapping $\mathfrak{M}$: $F \to f$ *Mellin transformation* ($\mathfrak{M}$-transformation) (cf. Doetsch [3, vol. I, p. 60]).

Now the solution $Y(x, \underline{a})$ of (2.1) is an element of $M_{0,b}$ for any $b > 0$, so its $\mathfrak{M}$-transform

$$H(s, \underline{a}) = \mathfrak{M}[Y(x, \underline{a}), s] \tag{3.1}$$

exists as a holomorphic function in the right half-plane $\operatorname{Re} s > 0$. Moreover, we have

THEOREM 2. *The $\mathfrak{M}$-transform $H(s, \underline{a})$ of $Y(x, \underline{a})$ has the following properties*:

(i) *It is a meromorphic function of s with at most simple poles in $s = -k$, $k = 0, 1, 2, \cdots$. The residues of H are given by*

$$\operatorname*{Res}_{s=-k} H(s, \underline{a}) = \frac{1}{k!} Y^{(k)}(0, \underline{a}), \qquad k = 0, 1, 2, \cdots. \tag{3.2}$$

(ii) *It is a solution of the difference equation*

$$H(s+n+m) + \sum_{k=\lambda}^{m} a_k H(s+n+m-k) = s(s+1) \cdot \cdots \cdot (s+n-1) \cdot H(s). \tag{3.3}$$

(iii) *For each* $s \in \mathbf{C}$, $s \neq 0, -1, -2, \cdots$, $H(s, \underline{a})$ *is an entire function of the parameters* $a_\lambda, \cdots, a_m$.

Proof. (i) and (ii) are consequences of elementary properties of the $\mathfrak{M}$-transformation (cf. [9, Lemma 10, 11, 12]) and the principle of analytic continuation. (iii) is a consequence of the uniform convergence of the $\mathfrak{M}$-integral for $H(s, \underline{a})$ with respect to the parameters, which follows itself from the uniform validity of the asymptotic expansion (2.4). For details cf. [9, pp. 22–24].

By virtue of formula (3.2), the central connection problem for $Y(x, \underline{a})$ is reduced to the determination of the residues of the function $H(s, \underline{a})$. Now property (iii) permits us to expand $H(s, \underline{a})$ like $Y(x, \underline{a})$ into a Taylor series with respect to the parameters:

$$H(s, \underline{a}) = \sum_{\underline{p}=\underline{0}}^{\infty} \omega_{\underline{p}}(s) \underline{a}^{\underline{p}}. \tag{3.4}$$

For the coefficients of this expansion we have

THEOREM 3. *The coefficient functions* $\omega_{\underline{p}}(s)$ *of* (3.4) *are meromorphic functions. They are connected with the* $\eta_{\underline{p}}(x)$ *by*

$$\omega_{\underline{p}}(s) = \mathfrak{M}[\eta_{\underline{p}}, s] \tag{3.5}$$

and

$$\operatorname*{Res}_{s=-k} \omega_{\underline{p}}(s) = \frac{1}{k!} \eta_{\underline{p}}^{(k)}(0), \qquad k = 0, 1, 2, \cdots, \tag{3.6}$$

and they satisfy the following system of difference equations:

$$\begin{aligned} \omega_{\underline{p}}(s+n+m) &= s(s+1)\cdots(s+n-1) \cdot \omega_{\underline{p}}(s) \\ &\quad - \sum_{\substack{j=\lambda \\ p_j \geqq 1}}^{m} \omega_{\underline{p}-\underline{e}_{j-\lambda+1}}(s+n+m-j). \end{aligned} \tag{3.7}$$

Proof. For (3.5) and (3.6), cf. [9, Lemma 15]. Formula (3.7) readily follows by inserting (3.4) into (3.3).

Like system (2.7), the new system (3.7) is recursive with respect to the order $|p|$.

We shall call (3.3) the associated difference equation (with (2.1)) and the system (3.7) the system of associated difference equations (with (2.7)). Similarly, we shall call connection problems for $H(s, \underline{a})$, resp. for the system (3.7), associated connection problems.

4. An associated connection problem. We will now derive asymptotic representations for the associated coefficient functions $\omega_{\underline{p}}(s)$. We have the following lemma concerning direct Abelian asymptotics for the $\mathfrak{M}$-transformation (cf. Lemma 18 in [9]).

LEMMA 1. *Let* $F(t)$ *be summable in the sense of Lebesgue on each compact subset of* $(0, \infty)$ *and satisfy the two conditions*:

(a) $F(t) = O(t^{-c}), \quad t \to 0, \quad c \in \mathbb{R}$;

(b) $F(t) \sim \exp(-\alpha t^{\beta}) \cdot t^{-\gamma}, \quad t \to \infty, \quad \alpha, \beta > 0, \quad \gamma \in \mathbf{C}$.

Then the $\mathfrak{M}$-transform $f(s)$ of $F(t)$ exists in the half-plane $\operatorname{Re} s > c$ *and satisfies*

$$f(s) \sim \frac{1}{\beta} \alpha^{-(s-\gamma)\beta^{-1}} \cdot \Gamma\left(\frac{s-\gamma}{\beta}\right) \tag{4.1}$$

as $s \to \infty$ in any half-strip

$$\Sigma_{c,d} = \{s \in \mathbb{C},\ \operatorname{Re} s > c,\ |\operatorname{Im} s| < d\}. \tag{4.2}$$

Now we can state

THEOREM 4. *The associated coefficient functions admit the asymptotic representations*

$$\omega_{\underline{p}}(s) = \frac{1}{\underline{p}!}(n\nu)^{1+(m\nu/2)(n-1)-n\nu s} \sum_{j=0}^{N} D_{\underline{a}}^{\underline{p}}[A_j(\underline{a})]_{\underline{a}=\underline{0}} \cdot (n\nu)^{j\nu} \cdot \Gamma(n\nu s - (m\nu/2)(n-1) - j\nu) + R_N(s), \tag{4.3}$$

where

$$R_N(s) = O[(n\nu)^{-n\nu s}\Gamma(n\nu s - (m\nu/2)(n-1) - (N+1)\nu)], \tag{4.4}$$

as $s \to \infty$ in any half-strip $\Sigma_{0,d}$, d arbitrary. Here N is an arbitrary nonnegative integer and $\nu = 1/(m+n)$.

Proof. If we truncate the asymptotic expansion (2.10) for $\eta_p(x)$ after the Nth term, transform termwise and apply Lemma 1 to the remainder of the expansion, we immediately get (4.3) and (4.4).

System (3.7) together with the asymptotic representations (4.3) constitute a connection problem for the associated coefficient functions $\omega_{\underline{p}}(s)$. This problem is well-defined, for we have (cf. [9, Satz 21]).

LEMMA 2. *The system* (3.7) *has a unique solution $\{\omega_{\underline{p}}(s)\}$ whose components $\omega_{\underline{p}}(s)$ are analytic in a common right half-plane and admit the asymptotic representations* (4.3).

5. Solution of the associated connection problem. We start with the equation of order 0 of system (3.7):

$$\omega_{\underline{0}}(s+n+m) = s(s+1) \cdot \cdots \cdot (s+n-1)\omega_{\underline{0}}(s). \tag{5.1}$$

A special solution of this equation is

$$\Omega_{\underline{0}}(s) = \nu^{-n\nu s} \cdot \prod_{j=0}^{n-1} \Gamma[\nu(s+j)], \qquad \nu = \frac{1}{m+n}. \tag{5.2}$$

As the general solution of (5.1) is the product of $\Omega_{\underline{0}}(s)$ and an arbitrary periodic function $p(s)$ of period $n+m$, we can put

$$\omega_{\underline{0}}(s) = \Omega_{\underline{0}}(s)p(s) \tag{5.3}$$

and have to determine $p(s)$ from the asymptotic representation for $\omega_{\underline{0}}(s)$.

Theorem 4 with $N=0$ provides

$$\omega_{\underline{0}}(s)\sim(n\nu)^{1+(m\nu/2)(n-1)-n\nu s}\ \Gamma\left[n\nu s-\frac{m\nu}{2}(n-1)\right] \tag{5.4}$$

as $s\to\infty$ in $\Sigma_{0,d}$. This gives

$$p(s)\sim(n\nu)^{1+(m\nu/2)(n-1)}n^{-n\nu s}\ \Gamma\left[n\nu s-\frac{m\nu}{2}(n-1)\right]$$

$$\cdot\prod_{j=0}^{n-1}\Gamma[\nu(s+j)]^{-1}$$

as $s\to\infty$ in $\Sigma_{0,d}$.

Applying the multiplication theorem of the Γ-function and the asymptotic expansion of $\Gamma(z+a)/\Gamma(z+b)$ (cf. Magnus–Oberhettinger–Soni [5, pp. 3 and 12]), we get

$$p(s)\sim\nu^{1+(m\nu/2)(n-1)}(2\pi)^{(1-n)/2}n^{1/2} \tag{5.5}$$

as $s\to\infty$ in $\Sigma_{0,d}$.

As $p(s)$ was supposed to be periodic, we even have equality in (5.5) and finally obtain

$$\omega_{\underline{0}}(s)=(2\pi)^{(1-n)/2}n^{1/2}\nu^{1+(m\nu/2)(n-1)}\cdot\nu^{-n\nu s}\prod_{j=0}^{n-1}\Gamma[\nu(s+j)], \tag{5.6}$$

where $\nu=1/(m+n)$.

We now introduce new functions $\theta_{\underline{p}}(s)$ and $\Lambda_{\underline{p}}(s)$ by

$$\text{(i)}\qquad \theta_{\underline{p}}(s)=\frac{\omega_{\underline{p}}[s(n+m)]}{\omega_{\underline{0}}[s(n+m)]}\ , \tag{5.7}$$

$$\text{(ii)}\qquad \Lambda_{\underline{p}}(s)=-\sum_{\substack{j=\lambda\\ p_j\geqq 1}}^{m}\frac{\omega_{\underline{p}-\underline{e}_{j-\lambda+1}}[(s+1-j\nu)(n+m)]}{\omega_{\underline{0}}[(s+1)(n+m)]}\ . \tag{5.7}$$

Then we can rewrite (3.7) in the form

$$\theta_{\underline{p}}(s+1)=\theta_{\underline{p}}(s)+\Lambda_{\underline{p}}(s). \tag{5.8}$$

This is an inhomogeneous difference equation of first order and we can apply the following lemma to it.

LEMMA 3. *If $\varphi(s)$ is holomorphic in a half-strip $\Sigma_{c,d}$ defined as in* (4.2), *and if*

$$\varphi(s)=O(s^{-1-\varepsilon})$$

with $\varepsilon>0$, then the difference equation

$$f(s+1)=f(s)+\varphi(s)$$

has a solution

$$\text{(i)}\qquad f_0(s)=-\sum_{k=0}^{\infty}\varphi(s+k),$$

which is holomorphic in $\Sigma_{c,d}$ *and satisfies*

(ii) $$f_0(s) = O(s^{-\varepsilon}).$$

Proof. The series (i) is uniformly convergent on each compact subset of $\Sigma_{c,d}$. For (ii), cf. Meschkowski [6, Satz 8, p. 49).

Now (4.3) provides $\Lambda_{\underline{p}}(s) = O(s^{-n\nu\lambda})$ in $\Sigma_{c,d}$, and as

$$n\nu\lambda = n\nu\left(\left[\frac{m+n}{n}\right]+1\right) \geqq n\nu\left(\frac{m+n}{n}+\frac{1}{n}\right) = 1+\nu,$$

we can apply Lemma 3 with $\varepsilon = \nu$ to (5.8). By this we get a solution $\theta^*_{\underline{p}}(s)$ of (5.8) in the form

$$\theta^*_{\underline{p}}(s) = -\sum_{k=0}^{\infty} \Lambda_{\underline{p}}(s+k),$$

which is holomorphic in $\Sigma_{0,d}$ and satisfies $\theta^*_{\underline{p}}(s) = O(s^{-\nu})$ there. Since the general solution of (5.8) is the sum of the special solution $\theta^*_{\underline{p}}(s)$ and an arbitrary periodic function of period 1, it is clear that $\theta^*_{\underline{p}}(s)$ is the only solution of (5.8) with this asymptotic property. On the other hand, since $A_0(\underline{a}) = 1$, we have from (4.3) for $\underline{p} \neq \underline{0}$, $\theta_{\underline{p}}(s) = O(s^{-\nu})$, $s \in \Sigma_{0,d}$, and this implies

(5.9) $$\theta_{\underline{p}}(s) = \theta^*_{\underline{p}}(s) = -\sum_{k=0}^{\infty} \Lambda_{\underline{p}}(s+k) \quad \text{for} \quad \underline{p} \neq \underline{0}.$$

Rewriting $\theta_{\underline{p}}(s)$ and $\Lambda_{\underline{p}}(s)$ in terms of $\omega_{\underline{p}}(s)$, we have

THEOREM 5. *The associated coefficient functions* $\omega_{\underline{p}}(s)$ *with* $\underline{p} \neq \underline{0}$ *admit the recursive sum representation*

(5.10) $$\omega_{\underline{p}}(s) = \omega_{\underline{0}}(s) \sum_{k=0}^{\infty} \sum_{\substack{j=\lambda \\ p_j \geqq 1}}^{m} \frac{\omega_{\underline{p}-\underline{e}_{j-\lambda+1}}[s+(n+m)(k+1)-j]}{\omega_{\underline{0}}[s+(n+m)(k+1)]},$$

where $\omega_{\underline{0}}(s)$ *is given by* (5.6).

At this point the associated connection problem is solved. What remains to do, is to give an explicit representation for $\omega_{\underline{p}}(s)$ and the initial values of $Y(x, \underline{a})$.

6. An explicit representation of the initial values $Y^{(k)}(0, \underline{a})$. If we replace in (5.10) the functions $\omega_{\underline{p}-\underline{e}_{j-\lambda+1}}$ themselves by their recursive sum representation, we get

(6.1)
$$\omega_{\underline{p}}(s) = \omega_{\underline{0}}(s) \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \sum_{j_1,j_2=\lambda}^{m} {}^{*} \frac{\omega_{\underline{0}}[s+(n+m)(k_1+1-j_1\nu)]}{\omega_{\underline{0}}[s+(n+m)(k_1+1)]} \cdot \frac{\omega_{\underline{q}(j_1,j_2)}[s+(n+m)(k_1+k_2+2-j_1\nu-j_2\nu)]}{\omega_{\underline{0}}[s+(n+m)(k_1+k_2+2-j_1\nu)]}$$

where $\underline{q}(j_1, j_2) = \underline{p} - \underline{e}_{j_1-\lambda+1} - \underline{e}_{j_2-\lambda+1}$ and the *-sign means that the finite sum has to be taken over all j_1, j_2 with $p_{j_1} \geqq 1$, $p_{j_2} - \delta_{j_1,j_2} \geqq 1$, $\delta_{j,k}$ the Kronecker symbol.

Thus we have expressed the coefficient functions of order $|\underline{p}|$ by such of order $|\underline{p}|-2$. We can repeat this process $|\underline{p}|-2$ times and doing this we finally obtain a

$|\underline{p}|$-fold infinite sum over a $|\underline{p}|$-fold finite sum. This final sum contains only terms of $\omega_{\underline{0}}(s)$ and is therefore, by virtue of (5.6), explicitly given in terms of Γ-functions.

Before we state this explicit representation of $\omega_{\underline{p}}(s)$, we introduce the following.

Notation. Let $|\underline{p}|>1$ and $\underline{j}_0(\underline{p})$ be given by

$$j_0(\underline{p})=(\underbrace{\lambda,\cdots,\lambda}_{p_\lambda\text{-times}},\underbrace{\lambda+1,\cdots,\lambda+1}_{p_{\lambda+1}\text{-times}},\cdots,\underbrace{m,\cdots,m}_{p_m\text{-times}})\in\mathbb{N}^{|\underline{p}|}.$$

Then we denote the set $\{\underline{j}=(j_1,j_2,\cdots,j_{|\underline{p}|})\}\subset\mathbb{N}^{|\underline{p}|}$ of all arrangements of $\underline{j}_0(\underline{p})$ as $A(\underline{p})$.

Moreover, we shall use the following abbreviation:

$$\gamma_n(a,b;\nu)=\frac{\Gamma(a)\Gamma(a+\nu)\Gamma(a+2\nu)\cdots\Gamma(a+(n-1)\nu)}{\Gamma(b)\Gamma(b+\nu)\Gamma(b+2\nu)\cdots\Gamma(b+(n-1)\nu)}. \tag{6.2}$$

Now we have

THEOREM 6. *If* $\underline{p}\neq\underline{0}$, *then*

$$\begin{aligned}\omega_{\underline{p}}(s)=C\cdot\nu^{n\nu(\lambda p_\lambda+(\lambda+1)p_{\lambda+1}+\cdots+mp_m)}\cdot\nu^{-n\nu s}\cdot\prod_{r=0}^{n-1}\Gamma[\nu(s+r)]\\ \cdot\sum_{\underline{j}\in A(\underline{p})}\sum_{i_1=1}^{\infty}\sum_{i_2=1}^{\infty}\cdots\sum_{i_{|\underline{p}|}=1}^{\infty}\\ \prod_{k=1}^{|\underline{p}|}\gamma_n[\nu s+(i_1+i_2+\cdots+i_k)-\nu(j_1+j_2+\cdots+j_k),\\ \nu s+(i_1+i_2+\cdots+i_k)-\nu(j_1+j_2+\cdots+j_{k-1});\nu]\end{aligned} \tag{6.3}$$

for all $s\in\mathbb{C}$ *except the points* $s_k=-(n+k)$, $k=0,1,2,\cdots$. *In the case* $\underline{p}=\underline{0}$, *we have*

$$\omega_{\underline{0}}(s)=C\cdot\nu^{-n\nu s}\cdot\prod_{r=0}^{n-1}\Gamma[\nu(s+r)]. \tag{6.4}$$

In both cases, the constant C *is given by*

$$C=(2\pi)^{(1-n)/2}n^{1/2}\nu^{1+(m\nu/2)(n-1)} \tag{6.5}$$

and

$$\nu=\frac{1}{m+n}.$$

We omit the proof which is based on induction. It is not difficult but requires formulas taking up too much space. We refer to [9, pp. 54–56].

We can now immediately state our main result, namely the explicit sum representation of the initial values $Y^{(k)}(0,\underline{a})$. We have the following.

THEOREM 7 (Main theorem). *The solution* $Y(x;a_\lambda,a_{\lambda+1},\cdots,a_m)$ *of the differential equation* (2.1) *which is uniquely defined by the asymptotic property* (2.4) *has the following initial values*:

$$Y^{(L)}(0;a_\lambda,\cdots,a_m)=\sum_{p_\lambda,\cdots,p_m=0}^{\infty}\rho^{(L)}_{p_\lambda,\cdots,p_m}\cdot a_\lambda^{p_\lambda}\cdot\cdots\cdot a_m^{p_m}, \tag{6.6}$$

$L = 0, 1, \cdots, n-1$, *with*

$$\rho^{(L)}_{p_\lambda,\cdots,p_m} = L!(2\pi)^{(1-n)/2} n^{1/2} \nu^{(m\nu/2)(n-1)+n\nu(\lambda p_\lambda+\cdots+m p_m+L)} \cdot \prod_{L\neq r=0}^{n-1} \Gamma[\nu(r-L)] \cdot \sigma^{(L)}_{p_\lambda,\cdots,p_m}. \tag{6.7}$$

The $\sigma^{(L)}_{p_\lambda,\cdots,p_m}$ *are given by*

$$\text{(i)} \qquad \sigma^{(L)}_{0,\cdots,0} = 1, \tag{6.8}$$

$$\text{(ii)} \qquad \sigma^{(L)}_{p_\lambda,\cdots,p_m} = \sum_{\underline{j}\in A(\underline{p})} \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} \cdots \sum_{i_{|\underline{p}|}=1}^{\infty} \prod_{k=1}^{|\underline{p}|} \gamma_n[(i_1+i_2+\cdots+i_k)-\nu(L+j_1+j_2+\cdots+j_k), (i_1+i_2+\cdots+i_k)-\nu(L+j_1+j_2+\cdots+j_{k-1}); \nu] \tag{6.8}$$

for $(p_\lambda, \cdots, p_m) \neq (0, \cdots, 0)$ *and* $\nu = 1/(m+n)$.

Proof. According to (2.6) and (3.6), we have

$$\rho^{(L)}_{p_\lambda,\cdots,p_m} = L! \operatorname*{Res}_{s=-L} \omega_{p_\lambda,\cdots,p_m}(s).$$

These residues can easily be calculated from (6.3) and (6.4) and provide the expressions (6.7) and (6.8).

Remark. The quantities $\sigma^{(L)}_{p_\lambda,\cdots,p_m}$ can be interpreted as $|\underline{p}|$-fold hypergeometric series with unit argument. In particular, we have for

$$|\underline{p}| = 1, \quad \underline{p} = (0, \cdots, \underset{\uparrow}{1}, \cdots, 0), \qquad \mu$$

if we write $\sigma^{(L)}_{\mu}$ instead of $\sigma^{(L)}_{\underline{p}}$:

$$\sigma^{(L)}_{\mu} = \prod_{j=0}^{n-1} \frac{\Gamma[1+\nu(j-L-\mu)]}{\Gamma[1+\nu(j-L)]} \cdot {}_nF_{n-1}[1-\nu(L+\mu), 1-\nu(L+\mu-1), \cdots, 1-\nu(L+\mu-n+1); 1-\nu L, 1-\nu(L-1), \cdots, 1-\nu, 1+\nu, \cdots, 1+\nu(n-L-1); 1] \tag{6.9}$$

(cf. [5, p. 62]).

7. Remarks on the general case. In the more general case of differential equation (1.1), an analogous procedure leads to a similar connection problem for a somewhat larger system of functions $\omega_{\underline{p}}(s)$ which contains the old "undercritical" $\omega_{\underline{p}}(s)$. The additional "overcritical" functions unfortunately have more complicated asymptotic representations than (4.3) and the series (5.10) no longer converges (cf. [9, Satz 19]).

Our farthest reaching result for the overcritical functions is a recursive integral representation obtained by combining certain results from function theory with difference equation methods (cf. [9, Satz 36]).

A further extension towards explicit representations of the initial values in the overcritical case, similar to the main result (6.6)–(6.8) of this paper is, however, not yet in sight.

REFERENCES

[1] B. L. J. BRAAKSMA, *Asymptotic analysis of a differential equation of Turrittin*, this Journal, 2 (1971), pp. 1–16.

[2] ———, *Recessive Solutions of Differential Equations with Polynomial Coefficients*, Lecture Notes in Mathematics, vol. 280, Springer, Berlin–Heidelberg–New York, 1972.

[3] G. DOETSCH, *Handbuch der Laplace-Transformation*, Birkhäuser, Basel–Stuttgart, 1955.

[4] P. F. HSIEH AND Y. SIBUYA, *On the asymptotic integration of second order linear ordinary differential equations with polynomial coefficients*, J. Math. Anal. Appl., 16 (1966), pp. 84–103.

[5] W. MAGNUS, F. OBERHETTINGER AND R. P. SONI, *Formulas and Theorems for the Special Functions of Mathematical Physics*, Springer, Berlin–Heidelberg–New York, 1966.

[6] H. MESCHKOWSKI, *Differenzengleichungen*, Vandenhoeck u. Ruprecht, Göttingen, 1959.

[7] W. WASOW, *Asymptotic Expansions for Ordinary Differential Equations*, Wiley–Interscience, New York–London–Sydney, 1965.

[8] ———, *Connection problems for asymptotic series*, Bull. Amer. Math. Soc., 74 (1968), pp. 831–853.

[9] H. WYRWICH, *Eine explizite Lösung des "Central Connection Problem" für eine gewöhnliche lineare Differentialgleichung n-ter Ordnung mit Polynomkoeffizienten*, Doctoral dissertation, Dortmund, 1974.

SIAM J. MATH. ANAL.
Vol. 8, No. 3, May 1977

POSITIVE SUMS OF THE CLASSICAL ORTHOGONAL POLYNOMIALS*

GEORGE GASPER†

Abstract. An expansion as a sum of squares of Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ is used to prove that if $0 \leqq \lambda \leqq \alpha+\beta$ and $\beta \geqq -1/2$, then

$$(*) \qquad \sum_{k=0}^{n} \frac{(\lambda+1)_{n-k}}{(n-k)!} \frac{(\lambda+1)_k}{k!} \frac{P_k^{(\alpha,\beta)}(x)}{P_k^{(\beta,\alpha)}(1)} \geqq 0, \qquad -1 \leqq x < \infty,$$

and the only cases of equality occur when $x=-1$ for n odd and when $\lambda=0$, $\alpha=-\beta=1/2$. Additional conditions are given under which $(*)$ holds and some special uses, limit cases, and important applications are pointed out. In particular, the case $\lambda=\alpha+\beta$ of $(*)$ is used to prove that if $\alpha, \beta \geqq -1/2$, then the Cesàro $(C, \alpha+\beta+2)$ means of the Jacobi series of a nonnegative function are nonnegative. Also, it is shown that

$$\frac{d}{d\theta} \sum_{k=0}^{n} \frac{(\lambda+1)_{n-k}}{(n-k)!} \frac{(\lambda+1)_k}{k!} \frac{\sin(k+1)\theta}{(k+1)\sin(\theta/2)} < 0, \qquad 0<\theta<\pi, \quad 0 \leqq \lambda \leqq 1,$$

which extends a recent result of Askey and Steinig.

1. Introduction. In [12] Askey and the author conjectured that if $0 \leqq \lambda \leqq \alpha+\beta$ and $\beta \geqq -1/2$, then the Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ satisfy the inequality

$$(1.1) \qquad \sum_{k=0}^{n} \frac{(\lambda+1)_{n-k}}{(n-k)!} \frac{(\lambda+1)_k}{k!} \frac{P_k^{(\alpha,\beta)}(x)}{P_k^{(\beta,\alpha)}(1)} > 0, \qquad -1 < x \leqq 1,$$

except when $\lambda=0$, $\alpha=-\beta=1/2$, when the sum is nonnegative and there are cases of equality. Here $(a)_0=1$ and $(a)_n=a(a+1)\cdots(a+n-1)=\Gamma(a+n)/\Gamma(a)$ for $n=1, 2, \cdots$. It was shown that this conjecture holds for $\beta \geqq \alpha$, for $|\beta| \leqq \alpha \leqq \beta+1$, for $0 \leqq \lambda \leqq \beta$, and for some other special cases; and applications of (1.1) to summability theory, location of zeros of trigonometric polynomials, mechanical quadrature, univalent functions, etc., were pointed out (also see Askey's recent book [9]). In particular, it was pointed out that the case $\lambda=\alpha+\beta$ of (1.1) would yield the result (conjectured in [5]) that the Cesàro $(C, \alpha+\beta+2)$ means of the Poisson kernel for Jacobi series are nonnegative when $\alpha, \beta \geqq -1/2$. It was also conjectured in [12] that if $\lambda \geqq 0$ and $\beta \geqq -1/2$, then the Laguerre polynomials $L_n^{\beta}(x)$ satisfy the inequality

$$(1.2) \qquad \sum_{k=0}^{n} \frac{(\lambda+1)_{n-k}}{(n-k)!} \frac{(\lambda+1)_k}{k!} \frac{(-1)^k L_k^{\beta}(x)}{L_k^{\beta}(0)} \geqq 0, \qquad x \geqq 0,$$

and this inequality was proved for the case $\beta \geqq \lambda \geqq -1/2$. Note that (1.2) is a limit case of (1.1) since

$$(1.3) \qquad \lim_{\alpha \to \infty} P_n^{(\alpha,\beta)}(-1+2x/\alpha) = (-1)^n L_n^{\beta}(x)$$

and $P_n^{(\beta,\alpha)}(1) = L_n^{\beta}(0)$.

* Received by the editors September 24, 1975, and in revised form January 30, 1976.

† Department of Mathematics, Northwestern University, Evanston, Illinois 60201. This work was supported in part by the National Science Foundation under Grant MPS71-03407 A03. The author is an Alfred P. Sloan fellow.

Recently the author [27] used a sum of squares of Bessel functions with nonnegative coefficients to prove that

$$\int_0^x (x-t)^\lambda t^{\lambda+1/2} J_\alpha(t)\,dt \geqq 0, \quad x \geqq 0, \quad 0 \leqq \lambda \leqq \alpha - 1/2, \tag{1.4}$$

which is a limit of the case $\beta = -1/2$ of (1.1), and this suggested that it might be possible to prove (1.1) for $0 \leqq \lambda \leqq \alpha - 1/2$, $\beta = -1/2$, by using a sum of squares of Jacobi polynomials. Then Bateman's integral [11, (3.4)]

$$\frac{P_n^{(\alpha-\mu,\beta+\mu)}(x)}{P_n^{(\beta+\mu,\alpha-\mu)}(1)} = \frac{\Gamma(\beta+\mu+1)}{\Gamma(\beta+1)\Gamma(\mu)} \int_{-1}^{x} \frac{P_n^{(\alpha,\beta)}(y)}{P_n^{(\beta,\alpha)}(1)} \frac{(1+y)^\beta}{(1+x)^{\beta+\mu}} (x-y)^{\mu-1}\,dy, \tag{1.5}$$
$$-1 < x \leqq 1, \quad \mu > 0, \quad \beta > -1,$$

could be used to give (1.1) for the conjectured values of α, β, λ, which would then yield the other conjectures mentioned above.

However, the extension in [28, (8.22)] of the proof of (1.4) for $\lambda = 0$ led to the result

$$\sum_{k=0}^{n} \frac{(1/2)_k(\alpha+1/2)_k(\alpha/2+5/4)_k}{k!(\alpha+1)_k(\alpha/2+1/4)_k} \frac{P_k^{(\alpha,-1/2)}(x)}{P_k^{(-1/2,\alpha)}(1)}$$
$$= \frac{4n+2\alpha+3}{2\alpha+3} \sum_{j=0}^{n} \frac{(1/2)_j(\alpha+3/2)_j(\alpha/2+3/4)_j(\alpha/2-1/4)_j}{j!(n-j)!(j+\alpha+1/2)_j(\alpha+1)_j(\alpha/2+5/4)_j(\alpha/2+7/4)_j}$$
$$\cdot \left(\frac{1-x}{2}\right)^j \left\{ \frac{(n-j)!}{(j+\alpha/2+5/4)_{n-j}} P_{n-j}^{(j+\alpha/2+1/4,\,j+\alpha/2+1/4)}\left(\left(\frac{1+x}{2}\right)^{1/2}\right)\right\}^2$$
$$> 0, \qquad -1 \leqq x \leqq 1, \quad \alpha > 1/2,$$

which, although strong enough to give the positivity of some generalized Cotes' numbers [10], [28], is unfortunately a weaker inequality than the case $\lambda = 0$, $\alpha > 1/2$, $\beta = -1/2$ of (1.1), as a summation by parts shows. Also, there did not seem to be any way to use expansions of the types

$$\sum_{j=0}^{[n/2]} A(n,j)\left\{P_{n-2j}^{(a,a)}\left(\left(\frac{1+x}{2}\right)^{1/2}\right)\right\}^2,$$

$$\sum_{j=0}^{[n/2]} B(n,j)(1-x)^{2j}\left\{P_{n-2j}^{(2j+a,2j+a)}\left(\left(\frac{1+x}{2}\right)^{1/2}\right)\right\}^2,$$

(which have been used to prove (1.1) for some special cases [12], [28]) to prove (1.1) for the case $\lambda = 0$, $\alpha > 1/2$, $\beta = -1/2$, and so a new type of expansion was needed.

By concentrating on the simpler Laguerre polynomial case the author was able to find an expansion which not only gave (1.2) under the less restrictive condition $\lambda, \beta \geqq -1/2$, but could also be extended to give (1.1) for the conjectured values of α, β, λ and for some cases in which $\lambda < 0$ or $\alpha + \beta < 0$.

Since the motivation for our proof of the Jacobi polynomial case comes from the Laguerre polynomial expansions, we first consider the Laguerre polynomial case in § 2 and then the Jacobi polynomial case in §§ 3, 4 and 5. In particular, the

case $\alpha = 3/2$, $\beta = -1/2$, $0 \leqq \lambda \leqq 1$ of (1.1) is used to derive the inequality

$$\frac{d}{d\theta} \sum_{k=0}^{n} \frac{(\lambda+1)_{n-k}}{(n-k)!} \frac{(\lambda+1)_k}{k!} \frac{\sin(k+1)\theta}{(k+1)\sin(\theta/2)} < 0, \qquad 0 < \theta < \pi, \quad 0 \leqq \lambda \leqq 1,$$

which for $0 \leqq \lambda \leqq 1$ is stronger than the result in [12] that

$$\sum_{k=0}^{n} \frac{(\lambda+1)_{n-k}}{(n-k)!} \frac{(\lambda+1)_k}{k!} \frac{\sin(k+1)\theta}{k+1} > 0, \qquad 0 < \theta < \pi, \quad -1 < \lambda \leqq 1.$$

A projection formula derived in § 5 is used in § 6 to prove that

$$\int_0^x (x-t)^{\alpha+2\mu-1/2} t^{\alpha+\mu} J_\alpha(t)\, dt \geqq 0, \qquad x \geqq 0,$$

when $0 \leqq \mu \leqq 1$ and $\alpha + \mu \geqq 1/2$, which was conjectured in [27]. Some new absolutely monotonic functions (i.e., functions with nonnegative power series coefficients) are derived in § 7. Related results and open problems are discussed in § 8, which also includes the observation that if $\alpha \geqq \beta$, $-1 < \beta < -1/2$ and $\lambda > -1$, then inequality (1.1) cannot hold for all n.

2. Laguerre polynomials. The expansions derived in this section arose in trying to find a proof of (1.2) similar to the following proof by Al-Salam and Carlitz [1, (4.6)] of a Turán type inequality for Laguerre polynomials:

$$\begin{aligned} &\left\{\frac{L_n^\alpha(x)}{L_n^\alpha(0)}\right\}^2 - \frac{L_{n-1}^\alpha(x)}{L_{n-1}^\alpha(0)} \frac{L_{n+1}^\alpha(x)}{L_{n+1}^\alpha(0)} \\ &\quad = \frac{(n-1)!}{(\alpha+1)_{n+1}} \sum_{k=1}^{n} \frac{(\alpha+2k)(n-k)!}{(\alpha+1)_{n+k}} x^{2k} \{L_{n-k}^{\alpha+2k}(x)\}^2 \\ &\quad \geqq 0, \qquad -\infty < x < \infty, \quad \alpha > -1. \end{aligned} \tag{2.1}$$

This expansion will clearly have to be modified since it has a double zero at $x = 0$, while the sum

$$S_n(x; \beta, \lambda) = \sum_{k=0}^{n} \frac{(\lambda+1)_{n-k}}{(n-k)!} \frac{(\lambda+1)_k}{k!} \frac{(-1)^k L_k^\beta(x)}{L_k^\beta(0)} \tag{2.2}$$

has a (simple) zero at $x = 0$ only for n odd when $\beta, \lambda > -1$. Also, since [12, (6.5)]

$$\sum_{k=0}^{n} \frac{(1/2)_{n-k}}{(n-k)!} \frac{(1/2)_k}{k!} \frac{(-1)^k L_k^{-1/2}(x)}{L_k^{-1/2}(0)} = \frac{\{H_n((x/2)^{1/2})\}^2}{2^n n!}, \qquad x \geqq 0, \tag{2.3}$$

where $H_n(x)$ is the Hermite polynomial of degree n, any expansion of (2.2) as a sum of squares of Laguerre polynomials will have to reduce to (2.3) as λ and β tend to $-1/2$. Thus, in view of the relations [33, (5.6.1)]

$$\begin{aligned} H_{2n}(x) &= (-1)^n 2^{2n} n! L_n^{-1/2}(x^2), \\ H_{2n+1}(x) &= (-1)^n 2^{2n+1} n! x L_n^{1/2}(x^2), \end{aligned} \tag{2.4}$$

one is led to conjecture that the sums (2.2) have expansions of the forms

$$S_{2n}(x;\beta,\lambda)=\sum_{j=0}^{n} A_{n,j}x^{2j}\{L_{n-j}^{a+2j}(x/2)\}^2,$$
$$S_{2n+1}(x;\beta,\lambda)=\sum_{j=0}^{n} B_{n,j}x^{2j+1}\{L_{n-j}^{b+2j}(x/2)\}^2, \tag{2.5}$$

with $a=a(\beta,\lambda)$ and $b=b(\beta,\lambda)$. Consideration of these expansions for small values of n shows that in order for them to exist it is necessary that $a=\beta$ and $b=\beta+1$; so that on reversing the order of summation of the right-hand sides of (2.5) we are led to consider the combined expansion

$$S_n(x;\beta,\lambda)=\sum_{j=0}^{[n/2]} C_{n,j}x^{n-2j}\{L_j^{\beta+n-2j}(x/2)\}^2. \tag{2.6}$$

In order to show that there is an expansion of the form (2.6) and to compute the coefficients we first use the formula [18, (92)]

$$\{L_n^{\alpha}(x)\}^2=\frac{\Gamma(\alpha+1+n)}{n!}\sum_{r=0}^{n}\frac{x^{2r}L_{n-r}^{\alpha+2r}(2x)}{r!\Gamma(\alpha+1+r)} \tag{2.7}$$

to observe that if there is an expansion of the form (2.6), then there is also one of the form

$$S_n(x;\beta,\lambda)=\sum_{j=0}^{[n/2]} c_{n,j}x^{n-2}L_j^{\beta+n-2j}(x), \tag{2.8}$$

which is computationally much easier to handle than (2.6) since it does not contain any squares of Laguerre polynomials. Multiplying both sides of (2.8) by x^{β} and using the Laplace transform formula [19, 10.12(32)]

$$\int_0^{\infty} e^{-sx}x^{\beta}L_n^{\beta}(x)\,dx=\frac{\Gamma(n+\beta+1)(s-1)^n}{n!s^{n+\beta+1}},\qquad s>0,\quad \beta>-1,$$

we find that (2.8) is equivalent to

$$F\left[\begin{matrix}-n,\lambda+1;\\ -\lambda-n\end{matrix}\,\frac{1-s}{s}\right]=\frac{n!(\beta+1)_n}{(\lambda+1)_n s^n}\sum_{j=0}^{[n/2]}\frac{c_{n,j}(s-s^2)^j}{j!(-\beta-n)_j}, \tag{2.9}$$

where F is the hypergeometric function [19, Chap. II]. From the quadratic transformation formula [19, 2.11(34) and p. 2]

$$F\left[\begin{matrix}-n,\lambda+1;\\ -\lambda-n\end{matrix}\,z\right]=(1+z)^n F\left[\begin{matrix}-n/2,(1-n)/2;\\ -\lambda-n\end{matrix}\,\frac{4z}{(1+z)^2}\right]$$

it follows on setting $z=(1-s)/s$ that (2.9) and hence (2.8) hold with

$$c_{n,j}=\frac{(\lambda+1)_n(-\beta-n)_j(-n)_{2j}}{n!(\beta+1)_n(-\lambda-n)_j},$$

where, as elsewhere, it is assumed that $j=0,1,\cdots,[n/2]$. Then, using the

Burchnall and Chaundy [18, (91)] inverse of (2.7)

$$L_n^{\alpha}(2x)=\sum_{r=0}^{n}\frac{(n-r)!(\alpha+2r)\Gamma(\alpha+r)}{r!\Gamma(\alpha+1+n+r)}(-x^2)^r\{L_{n-r}^{\alpha+2r}(x)\}^2, \tag{2.10}$$

on the right side of (2.8) we find that (2.6) holds with

$$\begin{aligned} C_{n,j}&=\frac{j!(\lambda+1)_n(\beta+1+n-j)_j}{(n-2j)!(\beta+1)_n(\lambda+1+n-j)_j(\beta+1+n-2j)_j}\\ &\quad\cdot {}_3F_2\left[\begin{matrix} j-n/2, j+(1-n)/2, j-n-\beta;\\ j-n-\lambda, 2j-n+1-\beta\end{matrix}\right]\\ &=\frac{2^{2j-2n}j!(2\lambda+2)_n(2\beta)_{n-2j}(\beta+1+n-j)_j}{(n-2j)!(\beta+1)_n(\beta)_{n-2j}(\lambda+3/2)_j(\beta+1+n-2j)_j}\\ &\quad\cdot {}_3F_2\left[\begin{matrix} j-n/2, j+(1-n)/2, \lambda+1/2;\\ j+\lambda+3/2, \beta+1/2\end{matrix}\right], \end{aligned} \tag{2.11}$$

by means of the third transformation formula on page 85 of [16].

As usual, the argument in the hypergeometric series is not displayed when it equals 1.

Clearly $C_{n,j}>0$, $j=0, 1, \cdots, [n/2]$, for $\lambda \geqq -1/2$, $\beta > -1/2$, since then each nonzero term in the second ${}_3F_2$ in (2.11) is positive. Also, consrderation of the limit case $\beta = -1/2$ of (2.11) shows that if $\lambda \geqq -1/2$ and $\beta = -1/2$, then $C_{n,j} \geqq 0$, $j=0, 1, \cdots, [n/2]$, with equality only when $\lambda = \beta = -1/2$ and $j < [n/2]$. Hence, from (2.6) we have

THEOREM 1. *If $\lambda, \beta \geqq -1/2$, then inequality* (1.2) *holds and the only cases of equality are at the endpoint $x=0$ for n odd, and at the zeros of $H_n((x/2)^{1/2})$ when $\lambda = \beta = -1/2$.*

In § 8 we shall show that the restriction $\beta \geqq -1/2$ in this theorem cannot be relaxed and that if $-1 < \lambda < \beta = -1/2$, then inequality (1.2) cannot hold for all n. It should be noted that, in view of (2.4), it follows from the cases $\beta = \pm 1/2$ that for the Hermite polynomials we have, for $\lambda \geqq -1/2$,

$$\sum_{k=0}^{n}\frac{(\lambda+1)_{n-k}}{(n-k)!}\frac{(\lambda+1)_k}{(2k)!}H_{2k}(x)\geqq 0, \qquad -\infty<x<\infty, \tag{2.12}$$

$$\sum_{k=0}^{n}\frac{(\lambda+1)_{n-k}}{(n-k)!}\frac{(\lambda+1)_k}{(2k+1)!}H_{2k+1}(x)\begin{cases}>0, & x>0,\\ =0, & x=0,\\ <0, & x<0,\end{cases} \tag{2.13}$$

and the only cases of equality in (2.12) occur when $\lambda = -1/2$ and for odd n when $x=0$.

Another derivation of the expansion (2.6) can be given by using the formulas [18, (98), (97)]

$$\{L_n^{\alpha}(x)\}^2 = \frac{\Gamma^2(\alpha+1+n)}{\Gamma(\alpha+1+2n)} \sum_{r=0}^{n} \frac{(2n-2r)!x^{2r}}{r!(n-r)!(n-r)!\Gamma(\alpha+1+r)} \cdot L_{2n-2r}^{\alpha+2r}(x), \tag{2.14}$$

$$L_{2n}^{\alpha}(x) = \frac{n!n!\Gamma(\alpha+1+2n)}{(2n)!} \cdot \sum_{r=0}^{n} \frac{(\alpha+2r)\Gamma(\alpha+r)(-x^2)^r}{\Gamma^2(\alpha+1+n+r)} \{L_{n-r}^{\alpha+2r}(x)\}^2 \tag{2.15}$$

in place of (2.7) and (2.10). This reduces the problem to a computation of the coefficients in the expansion

$$S_n(x;\beta,\lambda) = \sum_{j=0}^{[n/2]} c(n,j)x^{n-2j}L_{2j}^{\beta+n-2j}(x/2), \tag{2.16}$$

which can be done by using the above Laplace transform method and the transformation formula [19, use 2.10(1) and 2.11(4)]

$$F\left[\begin{matrix}-n, \lambda+1; z\\ -\lambda-n\end{matrix}\right] = \frac{(2\lambda+2)_n}{(\lambda+1)_n}\left(\frac{1+z}{2}\right)^n F\left[\begin{matrix}-n/2, (1-n)/2; \left(\frac{1-z}{1+z}\right)^2\\ \lambda+3/2\end{matrix}\right]$$

to show that (2.16) holds with

$$c(n,j) = \frac{(2j)!(2\lambda+2)_n(-n)_{2j}}{2^n n!j!(\beta+1)_n(\lambda+3/2)_j}.$$

The identities (2.8) and (2.16) can also be derived by using the series representation

$$L_n^{\beta}(x) = \frac{(\beta+1)_n}{n!} {}_1F_1[-n;\beta+1;x]$$

to write both sides of the identities as polynomials in x, and then comparing corresponding powers of x to show that these identities are equivalent to certain transformation formulas between generalized hypergeometric series. For instance, it then turns out that to compute the coefficients in (2.16) one needs the transformation formula

$$F\left[\begin{matrix}k-n, k+\lambda+1; -1\\ k-n-\lambda\end{matrix}\right] = \frac{(-1)^{k+n}k!(2\lambda+2)_n(1/2)_m}{2^k(\lambda+1)_k(\lambda+1)_{n-k}(\lambda+3/2)_m(k+2m-n)!} \cdot {}_3F_2\left[\begin{matrix}-m+(n-k)/2, -m+(1+n-k)/2, -m-\lambda-1/2;\\ -m+1/2, n-2m+1/2\end{matrix}\right], \tag{2.17}$$

where $m = [n/2]$. Formula (2.17) can be derived by first using a limit case of [16,

4.7(1)] to write the left side of (2.17) as a multiple of a ${}_3F_2(1)$ series and then applying a transformation formula for the ${}_3F_2(1)$ series [16, p. 18].

3. Extension of formula (2.6) to Jacobi polynomials. First observe that since [33]

$$\begin{aligned}\frac{P_n^{(1/2,-1/2)}(\cos\theta)}{P_n^{(-1/2,1/2)}(1)} &= \frac{\sin(n+1/2)\theta}{\sin(\theta/2)},\\ \frac{P_n^{(1/2,1/2)}(\cos\theta)}{P_n^{(1/2,1/2)}(1)} &= \frac{\sin(n+1)\theta}{(n+1)\sin\theta},\end{aligned} \tag{3.1}$$

it follows from the identity

$$\sum_{k=0}^{n}\frac{\sin(2n+1)\theta}{\sin\theta} = \left\{\frac{\sin(n+1)\theta}{\sin\theta}\right\}^2 \tag{3.2}$$

that, denoting the sum in (1.1) by $S_n(x;\alpha,\beta,\lambda)$, we have

$$\begin{aligned}S_{2n}(x;1/2,-1/2,0) &= \left\{\frac{P_n^{(1/2,-1/2)}(x)}{P_n^{(-1/2,1/2)}(1)}\right\}^2,\\ S_{2n+1}(x;1/2,-1/2,0) &= 2(n+1)^2(1+x)\left\{\frac{P_n^{(1/2,1/2)}(x)}{P_n^{(1/2,1/2)}(1)}\right\}^2.\end{aligned} \tag{3.3}$$

Therefore, in view of the identity (2.6) and the limit relation (1.3), one is led to conjecture that there is an expansion of the form

$$S_n(x;\alpha,\beta,\lambda) = \sum_{j=0}^{[n/2]} D_{n,j}(1+x)^{n-2j}\{P_j^{(\sigma,\beta+n-2j)}(x)\}^2 \tag{3.4}$$

with $\sigma=\sigma(\alpha,\beta,\lambda)$. By comparing powers of x on both sides of (3.4) for $n=2$ it is found that a necessary condition for this expansion to exist is that

$$\sigma = \frac{\alpha-\beta+\lambda}{2}. \tag{3.5}$$

We shall use a modification of the methods of § 2 to show that (3.5) is also a sufficient condition for there to exist an expansion of the form (3.4) for $n=0,1,2,\cdots$. To obtain an extension of (2.8) to Jacobi polynomials, observe that in [18] Burchnall and Chaundy derived [18, (92)] as a special case of [18, (43)], which they had derived as a limit case of their formula [17, (51)]

$$\begin{aligned}&F[a,b;c;x]F[a,b;c;y]\\ &\quad=\sum_{r=0}^{\infty}\frac{(a)_r(b)_r(c-a)_r(c-b)_r}{r!(c)_r(c)_{2r}}x^r y^r F[a+r,b+r;c+2r;x+y-xy].\end{aligned} \tag{3.6}$$

Using the series representation [19, 10.8(16)]

$$P_n^{(\alpha,\beta)}(x) = \frac{(-1)^n(\beta+1)_n}{n!}F[-n,n+\alpha+\beta+1;\beta+1;(1+x)/2] \tag{3.7}$$

in (3.6), we obtain an extension of (2.7) to Jacobi polynomials which shows that if

there is an expansion of the form (3.4), then there is one of the form

$$S_n(x;\alpha,\beta,\lambda)=\sum_{j=0}^{[n/2]}E_{n,j}(1+x)^{n-2j}P_j^{(\alpha,\beta+n-2j)}(x+(1-x^2)/2). \tag{3.8}$$

Formula (2.8) is clearly a limit case of (3.8). However, since the Laguerre polynomials on the right side of (2.8) have the variable x and it is difficult to compute the coefficients in (3.8) due to the presence of the quadratic variable $x+(1-x^2)/2$, one is tempted to look for another extension of (2.8) in which the Jacobi polynomials on the right side have the variable x. Such a formula can be discovered by using [18, (23)]

$$\begin{aligned}F[a,b;c;2x-x^2]=&\sum_{r=0}^{\infty}\frac{(a)_r(b)_r(c-a)_r}{r!(c)_{2r}}(-x^2)^r\\&\cdot F[a+r,2b+2r;c+2r;x]\end{aligned} \tag{3.9}$$

and (3.7) to find that if (3.8) holds with $\sigma=(\alpha-\beta+\lambda)/2$, then there is an expansion of the form

$$S_n(x;\alpha,\beta,\lambda)=\sum_{j=0}^{[n/2]}G_{n,j}(1+x)^{n-2j}P_j^{(\alpha+\lambda+1+n-j,\beta+n-2j)}(x), \tag{3.10}$$

which also has (2.8) as a limit case.

Since the Laplace transform method does not seem to be particularly well suited for deriving (3.10), we shall use another method. Reversing the order of summation in the series (3.7) and using [19, 2.9(4)]

$$F[a,b;c;z]=(1-z)^{-b}F[c-a,b;c;z/(z-1)]$$

gives, for $x>3$,

$$\begin{aligned}P_n^{(\alpha,\beta)}(x)=&\frac{(n+\alpha+\beta+1)_n}{n!}\left(\frac{x+1}{2}\right)^n\left(\frac{x-1}{x+1}\right)^{n+\beta}\\&\cdot F\left[\begin{matrix}-n-\alpha-\beta,\,-n-\beta;\\-2n-\alpha-\beta\end{matrix}\,\frac{2}{1-x}\right].\end{aligned} \tag{3.11}$$

Thus, using (3.11) in (3.10) and setting $t=2/(1-x)$, we find that (3.10) is equivalent to

$$\begin{aligned}&\sum_{k=0}^{n}\frac{(\lambda+1)_{n-k}(\lambda+1)_k(k+\alpha+\beta+1)_k}{(n-k)!k!(\beta+1)_k}(-t)^{n-k}F\left[\begin{matrix}-\alpha-\beta-k,\,-\beta-k;\\-\alpha-\beta-2k\end{matrix}\,t\right]\\&\quad=\sum_{j=0}^{[n/2]}\frac{(2n-2j+\alpha+\beta+\lambda+2)_j}{2^{2j-n}j!}G_{n,j}(-t)^j\\&\quad\cdot F\left[\begin{matrix}2j-2n-\alpha-\beta-\lambda-1,\,j-n-\beta;\\j-2n-\alpha-\beta-\lambda-1\end{matrix}\,t\right],\end{aligned}$$

which, by reversing the order of the first sum and comparing powers of t, is

equivalent to

$$
\begin{aligned}
&{}_3F_2\left[\begin{matrix} -n, \lambda+1, -k; \\ -\lambda-n, k-2n-\alpha-\beta \end{matrix}\right] \\
&\quad = \frac{n!(\alpha+\beta+1)_n(\beta+1)_n(-\alpha-\beta-2n)_k}{(\lambda+1)_n(\alpha+\beta+1)_{2n}(-\alpha-\beta-n)_k} \\
&\qquad \cdot \sum_{j=0}^{[n/2]} \frac{2^{n-2j}(-k)_j(n-2j+\beta+1)_j(k-2n-\alpha-\beta-\lambda-1)_j}{j!(-n-\beta)_{2j}} G_{n,j}.
\end{aligned} \tag{3.12}
$$

Now use [16, 4.5(1) and 7.2(1)] to obtain

$$
\begin{aligned}
&{}_3F_2\left[\begin{matrix} -n, \lambda+1, -k; \\ -\lambda-n, k-2n-\alpha-\beta \end{matrix}\right] \\
&\quad = \frac{(k-n-\alpha-\beta)_k}{(k-2n-\alpha-\beta)_k} \\
&\qquad \cdot {}_4F_3\left[\begin{matrix} n-k+\alpha+\beta+1, -n/2, -\lambda-(1+n)/2, -k; \\ -\lambda-n, -k+(n+\alpha+\beta+1)/2, -k+(n+\alpha+\beta+2)/2 \end{matrix}\right] \\
&\quad = \frac{(-\alpha-\beta-2n)_k}{(-\alpha-\beta-n)_k} {}_4F_3\left[\begin{matrix} -n/2, (1-n)/2, -k, k-2n-\alpha-\beta-\lambda-1; \\ -\lambda-n, -n-(\alpha+\beta)/2, -n+(1-\alpha-\beta)/2 \end{matrix}\right]
\end{aligned} \tag{3.13}
$$

which gives (3.12) and hence (3.10) with

$$
G_{n,j} = \frac{2^{2j-n}(\lambda+1)_n(\alpha+\beta+1)_{2n}(-n)_{2j}(n-j+\beta+1)_j}{n!(\beta+1)_n(\alpha+\beta+1)_n(-\alpha-\beta-2n)_{2j}(-\lambda-n)_j}. \tag{3.14}
$$

The main advantage of using the series representation (3.11) instead of (3.7) is that it was easier to derive (3.13) than to derive the transformation formula for the ${}_3F_2(-1)$ series which arises when (3.7) is used, because most of the transformation formulas given in [16] and elsewhere are for series with unit argument.

To return to (3.4) it suffices to observe that the inverses [17, (47)] and [17, (50)] of (3.9) and (3.6) can be combined to give [18, (16)]

$$
F[a, 2b; c; x] = \sum_{r=0}^{\infty} \frac{(a)_r(b)_r(b)_r(c-a)_r}{r!(c+r-1)_r(c)_{2r}} x^{2r} \{F[a+r, b+r; c+2r; x]\}^2, \tag{3.15}
$$

which can then be used in (3.10) to show that (3.4) holds for $\sigma = (\alpha-\beta+\lambda)/2$ with

$$
\begin{aligned}
D_{n,j} = &\frac{2^{2j-n}j!(\lambda+1)_n(\alpha+\beta+1)_{2n}(n-j+\beta+1)_j}{(n-2j)!(\beta+1)_n(\alpha+\beta+1)_n(n-j+\lambda+1)_j} d_{n,j} \\
&\cdot (2n-2j+\alpha+\beta+1)_{2j}(n-2j+\beta+1)_j
\end{aligned} \tag{3.16}
$$

where

$$d_{n,j} = d_{n,j}^{\alpha,\beta,\lambda}$$

$$= {}_5F_4\left[\begin{matrix} j-n/2, j+(1-n)/2, j-n-\beta, j-n-(\alpha+\beta+\lambda)/2, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad j-n+(\alpha+\beta+\lambda)/2; \\ 2j-n+1-\beta, j-n-\lambda, j-n-(\alpha+\beta)/2, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad j-n+(1-\alpha-\beta)/2 \end{matrix}\right]. \tag{3.17}$$

This formula can also be derived by using the Jacobi polynomial analogues of (2.14), (2.15), (2.16) which follows from the Burchnall and Chaundy formulas [17].

4. The cases $\lambda = \alpha+\beta$ and $\lambda = 0$. Since $S_0(x; \alpha, \beta, \lambda) = 1$ and

$$S_1(x; \alpha, \beta, \lambda) = \frac{(\lambda+1)(\alpha+\beta+2)}{\beta+1}\left(\frac{1+x}{2}\right),$$

inequality (1.1) clearly holds for $n = 0, 1$ when $\alpha+\beta > -2$, $\beta > -1$, $\lambda > -1$; and $\alpha+\beta > -2$ is a necessary condition for (1.1) to hold for all n when $\beta, \lambda > -1$. Also note that $S_n(x; \alpha, \beta, -1) = 0$, $n = 1, 2, \cdots$, and that the coefficient of $d_{n,j}$ in (3.16) is positive for $n \geqq 2$ when $\alpha+\beta \geqq -2$, $\beta > -1$, $\lambda > -1$. Therefore we may assume that $n \geqq 2$ and in considering the sign of the coefficients $D_{n,j}$ in (3.4) for $\alpha+\beta \geqq -2$, $\beta > -1$, $\lambda > -1$, it suffices to only consider the sign of $d_{n,j}^{\alpha,\beta,\lambda}$.

For $\lambda = \alpha+\beta$ the terminating Saalschützian (or balanced, see Askey [9, p. 56]) ${}_5F_4$ series in (3.17) reduces to a ${}_4F_3$ series to give

$$d_{n,j}^{\alpha,\beta,\alpha+\beta} = {}_4F_3\left[\begin{matrix} j-n/2, j+(1-n)/2, j-n-\beta, j-n-\alpha-\beta; \\ 2j-n+1-\beta, j-n-(\alpha+\beta)/2, j-n+(1-\alpha-\beta)/2 \end{matrix}\right]. \tag{4.1}$$

In the three cases $\alpha = \beta$, $\alpha = \beta+1$, and $\alpha = -\beta$ this series reduces to a ${}_3F_2$ series which can be summed by Saalschütz's formula [16, 2.2(1)]

$${}_3F_2\left[\begin{matrix} -n, a, b; \\ c, 1+a+b-c-n \end{matrix}\right] = \frac{(c-a)_n(c-b)_n}{(c)_n(c-a-b)_n}. \tag{4.2}$$

To avoid having to consider separately the cases when n is even or odd we let $m = [n/2]$, so that

$$\left(j-\frac{n}{2}\right)_k\left(j+\frac{1-n}{2}\right)_k = (j-m)_k(j+m-n+\tfrac{1}{2})_k$$

and application of (4.2) to (4.1) gives

$$d_{n,j}^{\beta,\beta,2\beta} = \frac{(j+\beta+1)_{m-j}(\beta+1/2)_{m-j}}{(n-m+\beta+1/2)_{m-j}(n-j-m+\beta)_{m-j}}, \tag{4.3}$$

$$d_{n,j}^{\beta+1,\beta,2\beta+1} = \frac{(j+\beta+2)_{m-j}(\beta+1/2)_{m-j}}{(n-m+\beta+3/2)_{m-j}(n-j-m+\beta)_{m-j}}, \tag{4.4}$$

$$d_{n,j}^{-\beta,\beta,0} = \frac{m!(\beta+1/2)_{m-j}}{j!(n-m+1/2)_{m-j}(n-j-m+\beta)_{m-j}}, \tag{4.5}$$

each of which is nonnegative for $\beta \geqq -1/2$, assuming the value zero only when $\beta = -1/2$ and $j < m$. To show that $d_{n,j}^{\alpha,\beta,\alpha+\beta} > 0$ for $\alpha+\beta > 0$, $\beta \geqq -1/2$ and for

$\alpha+\beta>-1$, $\beta>\alpha$ we need but apply Whipple's transformation formula [16, 7.2(1)]

$$(4.6)\qquad {}_4F_3\left[\begin{matrix}x, y, z, -n\\ u, v, w\end{matrix}\right]=\frac{(v-z)_n(w-z)_n}{(v)_n(w)_n}\,{}_4F_3\left[\begin{matrix}u-x, u-y, z, -n;\\ 1-v+z-n, 1-w+z-n, u\end{matrix}\right],$$

where it is assumed that $u+v+w-x-y-z+n=1$, to (4.1) to obtain

$$\begin{aligned} d_{n,j}^{\alpha,\beta,\alpha+\beta} &= \frac{(\alpha+\beta)_{2m-2j}}{(2n-2m+\alpha+\beta+1)_{2m-2j}} \\ &\quad\cdot {}_4F_3\left[\begin{matrix}j-m, j-m-\beta+1/2, j-n-\alpha-\beta, j+1;\\ 2j-n+1-\beta, j-m+(1-\alpha-\beta)/2, j-m+1-(\alpha+\beta)/2\end{matrix}\right] \\ &= \frac{(\beta-\alpha)_{2m-2j}}{(2n-2m+\alpha+\beta+1)_{2m-2j}} \qquad (4.7)\\ &\quad\cdot {}_4F_3\left[\begin{matrix}j-m, j-m-\beta+1/2, j-n-\beta, j+\alpha+1;\\ 2j-n+1-\beta, j-m+(1+\alpha-\beta)/2, j-m+1+(\alpha-\beta)/2\end{matrix}\right]. \end{aligned}$$

Clearly, the nonzero terms in the first ${}_4F_3$ in (4.7) are positive when $\alpha+\beta>0$, $\beta\geqq-1/2$, and those in the second ${}_4F_3$ are positive when $\alpha+\beta>-1$, $\beta>\alpha$. These positivity results could also have been obtained by using Whipple's ${}_7F_6$ representation [16, 7.1(1)]. From (3.4), (3.16) and these observations we have

THEOREM 2. *If $\alpha+\beta\geqq0$, $\beta\geqq-1/2$ or $\alpha+\beta\geqq-1$, $\beta\geqq\alpha$, then*

$$(4.8)\qquad \sum_{k=0}^{n}\frac{(\alpha+\beta+1)_{n-k}}{(n-k)!}\frac{(\alpha+\beta+1)_k}{k!}\frac{P_k^{(\alpha,\beta)}(x)}{P_k^{(\beta,\alpha)}(1)}\geqq0, \qquad -1\leqq x<\infty,$$

and the only cases of equality occur when $x=-1$ for n odd, when $\alpha+\beta=-1$, $n\geqq1$, and when $\lambda=0$, $\alpha=-\beta=1/2$ as given by (3.3).

It should be noted that the case $\alpha=\beta=1/2$ of this theorem gives Lukács' inequality [20, Satz XXIII]

$$(4.9)\qquad \sum_{k=0}^{n}(n+1-k)\sin(k+1)\theta>0, \qquad 0<\theta<\pi,$$

while the case $\alpha=3/2$, $\beta=-1/2$ (which, by means of (4.14) is equivalent to [22, Satz II]) gives the stronger result that

$$(4.10)\qquad \frac{d}{d\theta}\sum_{k=0}^{n}(n+1-k)\frac{\sin(k+1)\theta}{\sin(\theta/2)}<0, \qquad 0<\theta<\pi,$$

since, from (1.5),

$$(4.11)\qquad \frac{d}{dx}(1+x)^{1/2}\frac{P_n^{(1/2,1/2)}(x)}{P_n^{(1/2,1/2)}(1)}=\frac{1}{2}(1+x)^{-1/2}\frac{P_n^{(3/2,-1/2)}(x)}{P_n^{(-1/2,3/2)}(1)}.$$

Now recall that the Cesàro (C,γ) means of a formal series $\sum_{n=0}^{\infty}a_n$ are defined by

$$\frac{n!}{(\gamma+1)_n}\sum_{k=0}^{n}\frac{(\gamma+1)_{n-k}}{(n-k)!}a_k.$$

In his work on the L^p convergence of Lagrange interpolation polynomials at the zeros of Jacobi polynomials Askey [4], [5] was led to conjecture that if $\alpha, \beta \geqq -1/2$, then the $(C, \alpha+\beta+2)$ means of the Poisson kernel

$$\sum_{n=0}^{\infty} t^n h_n^{(\alpha,\beta)} P_n^{(\alpha,\beta)}(x) P_n^{(\alpha,\beta)}(y), \tag{4.12}$$

where

$$h_n^{(\alpha,\beta)} = \left\{ \int_{-1}^{1} [P_n^{(\alpha,\beta)}(x)]^2 (1-x)^\alpha (1+x)^\beta \, dx \right\}^{-1}$$
$$= \frac{2^{-\alpha-\beta-1} n! \Gamma(n+\alpha+\beta+1)(2n+\alpha+\beta+1)}{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)},$$

are nonnegative for $-1 \leqq x, y \leqq 1, 0 \leqq t \leqq 1$. This conjecture is known [5] to be best possible in the sense that not all of the (C, γ) means of (4.12) are nonnegative when $\gamma < \alpha+\beta+2$. We shall show in § 8 that it is also best possible in the sense that the restrictions on α, β cannot be relaxed. Using the relation $P_n^{(\alpha,\beta)}(x) = (-1)^n P_n^{(\beta,\alpha)}(-x)$, the positivity of the sum (4.12), and the convolution structure for Jacobi series [23], [24], Askey showed [5] that to prove his conjecture it is sufficient to prove it for $\alpha \geqq \beta \geqq -1/2$ with $t = y = 1$, i.e. it suffices to prove that

$$\sum_{k=0}^{n} \frac{(\alpha+\beta+3)_{n-k}}{(n-k)!} \frac{(2k+\alpha+\beta+1)(\alpha+\beta+1)_k}{k!(\alpha+\beta+1)} \frac{P_k^{(\alpha,\beta)}(x)}{P_k^{(\beta,\alpha)}(1)} \geqq 0, \tag{4.13}$$
$$-1 \leqq x \leqq 1,$$

for $\alpha \geqq \beta \geqq -1/2$. The cases $\alpha = \beta = -1/2$ and $\alpha = -\beta = 1/2$ of (4.13) are, respectively, Fejér's results [21], [22] that the $(C, 1)$ means of $1 + 2\sum_{k=1}^{\infty} \cos k\theta$ and the $(C, 2)$ means of $\sum_{n=0}^{\infty} (2n+1) \sin (2n+1)\theta$ are nonnegative; and the ultraspherical polynomial case $\alpha = \beta > -1/2$ is due to Kogbetliantz [29]. These classical results, generating functions, and Bateman's integral (1.5) were used in [5] to prove (4.13) for some other special cases, and in [12] the identity

$$\sum_{k=0}^{n} \frac{(\alpha+\beta+3)_{n-k}}{(n-k)!} \frac{(2k+\alpha+\beta+1)(\alpha+\beta+1)_k}{k!(\alpha+\beta+1)} \frac{P_k^{(\alpha,\beta)}(x)}{P_k^{(\beta,\alpha)}(1)}$$
$$= \sum_{k=0}^{n} \frac{(\alpha+\beta+2)_{n-k}}{(n-k)!} \frac{(\alpha+\beta+2)_k}{k!} \frac{P_k^{(\alpha+1,\beta)}(x)}{P_k^{(\beta,\alpha+1)}(1)} \tag{4.14}$$

was used to point out the equivalence of the nonnegativity of the sum (4.13) for $\alpha+\beta \geqq -1$, $\beta \geqq -1/2$ with the nonnegativity of the sum (4.8) for $\alpha+\beta \geqq 0$, $\beta \geqq -1/2$. Hence, Theorem 2 gives (4.13) for $\alpha+\beta \geqq -1$, $\beta \geqq -1/2$; so that in addition to the application of (4.13) in [6, Thm. 1] to the construction of nonnegative measures we also have

THEOREM 3. *Let $\alpha, \beta \geqq -1/2$. Then the $(C, \alpha+\beta+2)$ means and hence the (C, γ) means for $\gamma \geqq \alpha+\beta+2$ of the Poisson kernel* (4.12) *are nonnegative for $-1 \leqq x, y \leqq 1, 0 \leqq t \leqq 1$. Thus if $\gamma \geqq \alpha+\beta+2$ and $f(x)$, $-1 \leqq x \leqq 1$, is a nonnegative measurable function with $\int_{-1}^{1} f(x)(1-x)^\alpha (1+x)^\beta \, dx < \infty$, then the (C, γ)*

means of the Jacobi series of $f(x)$ *are nonnegative, i.e.,*

$$\sum_{k=0}^{n} \frac{(\gamma+1)_{n-k}}{(n-k)!} \hat{f}_k h_k^{(\alpha,\beta)} P_k^{(\alpha,\beta)}(x) \geqq 0, \qquad -1 \leqq x \leqq 1,$$

where $\hat{f}_k = \int_{-1}^{1} f(x) P_k^{(\alpha,\beta)}(x)(1-x)^{\alpha}(1+x)^{\beta}\,dx$.

Let us now consider (1.1) for $\lambda = 0$. Proceeding as in the case $\lambda = \alpha+\beta$ we have

$$d_{n,j}^{\alpha,0,0} = \frac{(1/2)_{m-j}(j+1+\alpha/2)_{m-j}}{(n-j-m)_{m-j}(n-m+(\alpha+1)/2)_{m-j}}, \tag{4.15}$$

$$d_{n,j}^{\beta+1,\beta,0} = \frac{(j+3/2)_{m-j}(\beta+1/2)_{m-j}}{(n-m+1)_{m-j}(n-j-m+\beta)_{m-j}}, \tag{4.16}$$

$$\begin{aligned} d_{n,j}^{\alpha,\beta,0} &= \frac{(1/2)_{m-j}((\alpha+\beta)/2)_{m-j}}{(n-m+1)_{m-j}(n-m+(\alpha+\beta+1)/2)_{m-j}} \\ &\quad \cdot {}_4F_3\left[\begin{matrix} j-m, j-m-\beta+1/2, j-n-(\alpha+\beta)/2, j+1; \\ 2j-n+1-\beta, j-m+1/2, j-m+1-(\alpha+\beta)/2 \end{matrix}\right] \\ &= \frac{(\beta)_{m-j}(j+1)_{m-j}}{(n-m+1)_{m-j}(n-j-m+\beta)_{m-j}} \qquad (4.17) \\ &\quad \cdot {}_4F_3\left[\begin{matrix} j-m, j-n-\beta, -m-(\alpha+\beta)/2, 1/2; \\ j-n+(1-\alpha-\beta)/2, j-m+1-\beta, -m \end{matrix}\right] \\ &\quad \text{to } m-j+1 \text{ terms,} \end{aligned}$$

where $m = [n/2]$. From (4.5), (4.15), (4.16) and (4.17) it follows that if $\alpha+\beta \geqq 0$, $\beta \geqq -1/2$ or $\alpha+\beta \geqq -2$, $\beta \geqq 0$, then $d_{n,j}^{\alpha,\beta,0} \geqq 0$ and the value zero is assumed for $n \geqq 2$ only when $\alpha = -\beta = 1/2$, $j < m$ or $\alpha = -2$, $\beta = 0$, $j = 0$. Hence using (3.4), (3.16) and, when $\alpha = -2$, $\beta = 0$, the fact that [33, (4.22.2)]

$$P_j^{(-1,n-2j)}(x) = \frac{n-j}{j}\left(\frac{x-1}{2}\right) P_{j-1}^{(1,n-2j)}(x), \qquad j \geqq 1,$$

we obtain

THEOREM 4. *If* $\alpha+\beta \geqq 0$, $\beta \geqq -1/2$ *or* $\alpha+\beta \geqq -2$, $\beta \geqq 0$, *then*

$$\sum_{k=0}^{n} \frac{P_k^{(\alpha,\beta)}(x)}{P_k^{(\beta,\alpha)}(1)} \geqq 0, \qquad -1 \leqq x < \infty, \tag{4.18}$$

and the only cases of equality occur when $x = -1$ *for n odd, when* $\alpha = -\beta = 1/2$, *when* $\alpha+\beta = -2$, $n = 1$, *and when* $\alpha = -2$, $\beta = 0$, $x = 1$, $n \geqq 1$.

Note that the case $\alpha = \beta = 1/2$ of (4.18) gives the Fejér-Jackson-Gronwall inequality [20, Satz XXVI]

$$\sum_{k=0}^{n} \frac{\sin(k+1)\theta}{k+1} > 0, \qquad 0 < \theta < \pi, \tag{4.19}$$

and the case $\alpha = 3/2$, $\beta = -1/2$ gives, by means of (4.11), the stronger inequality

of Askey and Steinig [15]

$$\frac{d}{d\theta}\sum_{k=0}^{n}\frac{\sin(k+1)\theta}{(k+1)\sin\theta/2}<0, \qquad 0<\theta<\pi. \tag{4.20}$$

The case $\alpha=5/2$, $\beta=-1/2$ was proved in [12] and it was shown that this case is equivalent to

$$(n+1)\frac{\sin(n-1)\theta}{\sin\theta}-(n-1)\frac{\sin(n+1)\theta}{\sin\theta}\leqq(3+\cos\theta)\left(n-\frac{\sin n\theta}{\sin\theta}\right),$$

which is stronger than the inequality

$$(n+1)\frac{\sin(n-1)\theta}{\sin\theta}-(n-1)\frac{\sin(n+1)\theta}{\sin\theta}\leqq 4\left(n-\frac{\sin n\theta}{\sin\theta}\right)$$

that Robertson proved and used in his work [32] on univalent functions. For other known special cases of Theorem 4 and their applications see [2], [3], [9], [28].

The ${}_5F_4$ representation for $d_{n,j}^{\alpha,\beta,\lambda}$ also reduces to a ${}_4F_3$ in the three cases $\alpha=\beta$, $\alpha=\beta+1$ and $\lambda=\beta$, so that (4.6) then gives

$$d_{n,j}^{\alpha,\alpha,\lambda}=\frac{(\alpha-\lambda/2)_{m-j}((\lambda+1)/2)_{m-j}}{(n-m+\lambda+1)_{m-j}(n-m+\alpha+1/2)_{m-j}} \cdot {}_4F_3\left[\begin{matrix} j-m, j-m-\alpha+1/2, j-n-\alpha-\lambda/2, j+1+\lambda/2; \\ 2j-n+1-\alpha, j-m+(1-\lambda)/2, j-m+1-\alpha+\lambda/2 \end{matrix}\right], \tag{4.21}$$

$$d_{n,j}^{\beta+1,\beta,\lambda}=\frac{(\lambda/2)_{m-j}(\beta+(1-\lambda)/2)_{m-j}}{(n-m+\lambda+1)_{m-j}(n-m+\beta+3/2)_{m-j}} \cdot {}_4F_3\left[\begin{matrix} j-m, j-m-\beta+1/2, j-n-\beta-(\lambda+1)/2, j+(\lambda+3)/2; \\ 2j-n-\beta+1, j-m+1-\lambda/2, j-m-\beta+(\lambda+1)/2 \end{matrix}\right], \tag{4.22}$$

$$d_{n,j}^{\alpha,\beta,\beta}=\frac{(\beta)_{2m-2j}}{(2n-2m+\alpha+\beta+1)_{2m-2j}} \cdot {}_4F_3\left[\begin{matrix} j-m, j-n-\beta-\alpha/2, j-m-\beta+1/2, j+1+\alpha/2; \\ 2j-n+1-\beta, j-m+(1-\beta)/2, j-m+1-\beta/2 \end{matrix}\right], \tag{4.23}$$

where $m=[n/2]$. From these representations it follows that

$$\begin{aligned} d_{n,j}^{\alpha,\alpha,\lambda}&>0, && -1<\lambda<2\alpha, \\ d_{n,j}^{\beta+1,\beta,\lambda}&>0, && 0<\lambda<2\beta+1, \\ d_{n,j}^{\alpha,\beta,\beta}&>0, && \alpha\geqq-2,\ \beta>0, \end{aligned} \tag{4.24}$$

which can be combined with (3.4), (3.16) and our above results for the cases $\lambda=\alpha+\beta$ and $\lambda=0$ to show that inequality (1.1) holds for the three cases

(i) $\alpha=\beta$, $-1<\lambda\leqq 2\alpha$,
(ii) $\alpha=\beta+1$, $0\leqq\lambda\leqq 2\beta+1$, $\beta>-1/2$,
(iii) $\lambda=\beta$, $\alpha\geqq-2$, $\beta>0$,

each of which was proved in [12, Thms. 1,2,7] by other methods. It was also pointed out in [12] that the case $\alpha = 1/2$ of (i) above gives

$$\sum_{k=0}^{n} \frac{(\lambda+1)_{n-k}}{(n-k)!}\frac{(\lambda+1)_k}{k!}\frac{\sin(k+1)\theta}{k+1} > 0, \tag{4.25}$$

$$0<\theta<\pi, \quad -1<\lambda\leqq 1,$$

which includes inequalities (4.9) and (4.19).

5. The case $0<\lambda<\alpha+\beta$. To prove (1.1) for the case $0<\lambda<\alpha+\beta$, $\beta \geqq -1/2$, it suffices to prove the positivity of the ${}_5F_4$ series in (3.17). Unfortunately, there do not seem to be any transformation formulas for ${}_5F_4$ series in the literature which will show this positivity. The method of projection formulas described in [26], [28] suggests that we look for a formula of the type

$$d_{n,j}^{\alpha,\beta,\lambda} = \sum_{j=0}^{n} a_{n,j}F_{n,j},$$

where $F_{n,j}$ is a Saalschützian ${}_4F_3$ series to which the transformation formula (4.6) could be applied. To find it the author observed that by starting with the formula [16, 4.3(3)] and reversing the proof given in Bailey [16, § 4.3] of Whipple's transformation formula [16, 4.3(4)] we obtain

$${}_{p+2}F_{p+1}\left[\begin{matrix}-n, a_1, a_2, \cdots, a_{p+1};\\ b_1, b_2, \cdots, b_{p+1}\end{matrix}\right]$$

$$= \sum_{k=0}^{n} \frac{(-1)^k(-n)_k(b_1+b_2-a_1-1)_k(b_1-a_1)_k \cdot (b_2-a_1)_k(a_2)_k\cdots(a_{p+1})_k}{k!\,(b_1+b_2-a_1-1)_{2k}(b_1)_k(b_2)_k\cdots(b_{p+1})_k} \cdot {}_{p+1}F_p\left[\begin{matrix}k-n, k+a_2, k+a_3, \cdots, k+a_{p+1};\\ 2k+b_1+b_2-a_1, k+b_3, k+b_4, \cdots, k+b_{p+1}\end{matrix}\right], \tag{5.1}$$

which can also be obtained as a limit case of [16, 4.3(6)]. This formula has the property that if the ${}_{p+2}F_{p+1}$ series is Saalschützian, then so are the ${}_{p+1}F_p$ series.

Application of (5.1) to (3.17) gives

$$d_{n,j}^{\alpha,\beta,\lambda} = \sum_{k=0}^{m-j} \frac{(-1)^k(j-m)_k(j-n+(\lambda-\alpha-\beta-1)/2)_k}{k!(j-n+(\lambda-\alpha-\beta-1)/2)_{2k}(j-n-(\alpha+\beta)/2)_k} \cdot \frac{(\lambda/2)_k((\lambda+1)/2)_k(j+m-n+1/2)_k(j-n-\beta)_k(j-n-(\alpha+\beta+\lambda)/2)_k}{(j-n+(1-\alpha-\beta)/2)_k(2j-n+1-\beta)_k(j-n-\lambda)_k} \cdot {}_4F_3\left[\begin{matrix}k+j-m, k+j+m-n+1/2, k+j-n-\beta, k+j-n-(\alpha+\beta+\lambda)/2;\\ 2k+j-n+(1+\lambda-\alpha-\beta)/2, k+2j-n+1-\beta, k+j-n-\lambda\end{matrix}\right] \tag{5.2}$$

with $m=[n/2]$. By (4.6) the above ${}_4F_3$ equals

$$\frac{(k+\lambda+1/2)_{m-j-k}((\alpha+\beta-\lambda)/2)_{m-j-k}}{(2k+j-n+(1+\lambda-\alpha-\beta)/2)_{m-j-k}(k+j-n-\lambda)_{m-j-k}} \cdot {}_4F_3\left[\begin{matrix}j-m-\beta+1/2, j+1, k+j-n-(\alpha+\beta+\lambda)/2, k+j-m;\\ j-m-\lambda+1/2, k+j-m+1+(\lambda-\alpha-\beta)/2, k+2j-n+1-\beta\end{matrix}\right]. \tag{5.3}$$

When $0<\lambda<\alpha+\beta$, $\beta \geqq -1/2$, each nonzero term in this ${}_4F_3$ is positive and the positivity of $d_{n,j}^{\alpha,\beta,\lambda}$ follows by using (5.2), (5.3) and the relation $(a)_n = (-1)^n(1-a-n)_n$. This, combined with our observations in § 4, gives the following result.

THEOREM 5. *If* $0 \leqq \lambda \leqq \alpha+\beta \geqq -1/2$, *then*

$$\sum_{k=0}^{n} \frac{(\lambda+1)_{n-k}}{(n-k)!}\frac{(\lambda+1)_k}{k!}\frac{P_k^{(\alpha,\beta)}(x)}{P_k^{(\beta,\alpha)}(1)} \geqq 0, \qquad -1 \leqq x < \infty, \tag{5.4}$$

and the only cases of equality occur when $x=-1$ *for* n *odd and when* $\lambda=0$, $\alpha=-\beta=1/2$.

In particular, it should be observed that from (4.11) and the case $\alpha=3/2$, $\beta=-1/2$ of this theorem we have

$$\frac{d}{d\theta}\sum_{k=0}^{n} \frac{(\lambda+1)_{n-k}}{(n-k)!}\frac{(\lambda+1)_k}{k!}\frac{\sin(k+1)\theta}{(k+1)\sin(\theta/2)} < 0, \tag{5.5}$$

$$0<\theta<\pi, \quad 0 \leqq \lambda \leqq 1,$$

which includes both (4.10) and (4.20) and for $0 \leqq \lambda \leqq 1$ is stronger than (4.25).

From (4.6) we also have that the ${}_4F_3$ in (5.2) equals

$$\frac{(\beta-\lambda)_{m-j-k}(j+1)_{m-j-k}}{(k+j-n-\lambda)_{m-j-k}(k+2j-n+1-\beta)_{m-j-k}} \cdot {}_4F_3\left[\begin{matrix} k-m+(\lambda-\alpha-\beta)/2,\ k+\lambda+1/2,\ k+j-n-\beta,\ k+j-m; \\ k+j-m+1+\lambda-\beta,\ k-m,\ 2k+j-n+(1+\lambda-\alpha-\beta)/2 \end{matrix}\right],$$

which, with (5.2) and (4.23), shows that $d_{n,j}^{\alpha,\beta,\lambda}>0$ when $0<\lambda \leqq \min(\beta, \alpha+\beta+2)$. This and our previous observations give the previously known result [12, Thm. 8] that if $0 \leqq \lambda \leqq \min(\beta, \alpha+\beta+2)$, then inequality (1.1) holds except when $\alpha+\beta=-2$ and either $n=1$ or $x=1$, $\lambda=\beta=0$, $n \geqq 1$.

It was proved in [12, Thm. 1] that (1.1) also holds when $\beta \geqq \alpha$, $-1<\lambda<\alpha+\beta$. To obtain this result from (3.4) and (3.16), it suffices to note that the positivity of $d_{n,j}^{\alpha,\beta,\lambda}$ in this case follows from

$$\begin{aligned} d_{n,.j}^{\alpha,\beta,\lambda} &= \sum_{k=0}^{m-j} \frac{(-1)^k(j-m)_k(j-n-(\lambda+1)/2)_k((\alpha+\beta-\lambda)/2)_k((\lambda+1)/2)_k}{k!(j-n-(\lambda+1)/2)_{2k}(j-n-\lambda)_k(j-n+(1-\alpha-\beta)/2)_k} \\ &\quad \cdot \frac{(j-n-(\alpha+\beta+\lambda)/2)_k(j-n-\beta)_k(j+m-n+1/2)_k}{(2j-n+1-\beta)_k(j-n-(\alpha+\beta)/2)_k} \\ &\quad \cdot {}_4F_3\left[\begin{matrix} k+j-m,\ k+j-n-(\alpha+\beta+\lambda)/2,\ k+j-n-\beta,\ k+j+m-n+1/2; \\ 2k+j-n+(1-\lambda)/2,\ k+2j-n+1-\beta,\ k+j-n-(\alpha+\beta)/2 \end{matrix}\right] \\ &= \sum_{k=0}^{m-j} \frac{(n-2j-2k+1)_{2k}(n-j-k+(\lambda+3)/2)_k((\alpha+\beta-\lambda)/2)_k((\lambda+1)/2)_k}{k!(n-j-2k+(\lambda+3)/2)_{2k}(n-2j-k+\beta)_k(n-j-k+\lambda+1)_k} \\ &\quad \cdot \frac{(n-j-k+(\alpha+\beta+2+\lambda)/2)_k(n-j-k+\beta+1)_k(j+1)_{m-j-k}((\beta-\alpha)/2)_{m-j-k}}{(2n-2j-2k+\alpha+\beta+1)_{2k}(n-m+1+(\alpha+\beta)/2)_{m-j-k}(n-m-j+\beta)_{m-j-k}} \\ &\quad \cdot {}_4F_3\left[\begin{matrix} k-m-\lambda/2,\ k+(\alpha+\beta+1)/2,\ k+j-n-\beta,\ k+j-m; \\ 2k+j-n+(1-\lambda)/2,\ k+j-m+1+(\alpha-\beta)/2,\ k-m \end{matrix}\right]. \end{aligned}$$

Similarly, it follows from

$$d_{n,j}^{\alpha,\beta,\lambda}$$

$$= \sum_{k=0}^{m-j} \frac{(-1)^k (j-m)_k (2j-n-\beta+(\lambda+1)/2)_k (j+1+(\alpha-\beta+\lambda)/2)_k}{k!(2j-n-\beta+(\lambda+1)/2)_{2k}(2j-n+1-\beta)_k}$$

$$\cdot \frac{((\lambda+1)/2)_k (j-n-(\alpha+\beta+\lambda)/2)_k (j-n-\beta)_k (j+m-n+1/2)_k}{(j-n+(1-\alpha-\beta)/2)_k (j-n-\lambda)_k (j-n-(\alpha+\beta)/2)_k}$$

$$\cdot {}_4F_3\left[\begin{matrix} k+j-m,\ k+j-n-(\alpha+\beta+\lambda)/2,\ k+j-n-\beta,\ k+j+m-n+1/2; \\ 2k+2j-n-\beta+(\lambda+3)/2,\ k+j-n-\lambda,\ k+j-n-(\alpha+\beta)/2 \end{matrix}\right]$$

$$= \sum_{k=0}^{m-j} \frac{(n-2j-2k+1)_{2k}(n-2j-k+\beta+(1-\lambda)/2)_k \cdot (j+1+(\alpha-\beta+\lambda)/2)_k ((\lambda+1)/2)_k}{k!(n-2j-2k+\beta+(1-\lambda)/2)_{2k}(n-2j-k+\beta)_k (n-j-k+\lambda+1)_k}$$

$$\cdot \frac{(n-j-k+\beta+1)_k (n-j-k+1+(\alpha+\beta+\lambda)/2)_k (\beta-\lambda)_{m-j-k} ((\beta-\alpha)/2)_{m-j-k}}{(2n-2j-2k+\alpha+\beta+1)_{2k}(n-m+\lambda+1)_{m-j-k}(n-m+1+(\alpha+\beta)/2)_{m-j-k}}$$

$$\cdot {}_4F_3\left[\begin{matrix} k+j-m-\beta+1+\lambda/2,\ k+j+\lambda+(\alpha-\beta+3)/2,\ k+j-n-\beta,\ k+j-m; \\ k+j-m+1+\lambda-\beta,\ k+j-m+1+(\alpha-\beta)/2,\ 2k+2j-n-\beta+(\lambda+3)/2 \end{matrix}\right]$$

and the cases $\alpha = \beta$ and $\lambda = \beta$ considered in § 4 that we also have

THEOREM 6. *Let $\beta \geqq \alpha$, $\beta \geqq \lambda > -1$ and $2\beta \geqq \lambda \geqq \beta - \alpha - 2$. Then inequality* (5.4) *holds and the only cases of equality occur when $x = -1$ for n odd, when $\alpha = -2$, $\beta = \lambda = 0$, $n = 1$, and when $\alpha = -2$, $\beta = \lambda = 0$, $x = 1$, $n \geqq 1$.*

Note that this theorem gives cases in which (5.4) holds with $\beta < 0$, $\alpha + \beta < -1$ which are not covered by the previously known results. It should also be noted that from [33, (4.1.8)]

$$\frac{P_n^{(-1/2,1/2)}(\cos\theta)}{P_n^{(1/2,-1/2)}(1)} = \frac{\cos(n+1/2)\theta}{(2n+1)\cos\theta/2}$$

and the case $\alpha = -\beta = -1/2$ of Theorem 6 we have

$$\sum_{k=0}^{n} \frac{(\lambda+1)_{n-k}}{(n-k)!} \frac{(\lambda+1)_k}{k!} \frac{\cos(k+1/2)\theta}{2k+1} > 0, \tag{5.6}$$

$$0 \leqq \theta < \pi, \quad -1 < \lambda \leqq 1/2,$$

which can also be obtained by combining Theorems 1 and 8 in [12].

Since Jacobi polynomials with $\alpha = \pm 1/2$ or $\beta = \pm 1/2$ can be expressed in terms of ultraspherical polynomials $P_n^{(\alpha,\alpha)}(x)$ by means of a quadratic transformation [33, Thm. 4.1], our results also give the following inequalities:

$$\sum_{k=0}^{n} \frac{(\lambda+1)_{n-k}}{(n-k)!} \frac{(\lambda+1)_k}{k!} \frac{(\alpha+1)_k}{(1/2)_k} \frac{P_{2k}^{(\alpha,\alpha)}(x)}{P_{2k}^{(\alpha,\alpha)}(1)} \geqq 0 \tag{5.7}$$

for $-\infty < x < \infty$ when $0 \leqq \lambda \leqq \alpha - 1/2$.

$$\sum_{k=0}^{n} \frac{(\lambda+1)_{n-k}}{(n-k)!} \frac{(\lambda+1)_k}{k!} \frac{(-1)^k P_{2k}^{(\alpha,\alpha)}(x)}{P_{2k}^{(\alpha,\alpha)}(1)} \geqq 0 \tag{5.8}$$

for $-1 \leqq x \leqq 1$ when $-1 < \lambda \leqq \alpha - 1/2$, when $0 \leqq \lambda \leqq \alpha$, and when $-1 < \lambda \leqq \alpha$

and $\alpha - 3/2 \leqq \lambda \leqq 2\alpha$.

$$\sum_{k=0}^{n} \frac{(\lambda+1)_{n-k}}{(n-k)!} \frac{(\lambda+1)_k}{k!} \frac{(\alpha+1)_k}{(3/2)_k} \frac{P_{2k+1}^{(\alpha,\alpha)}(x)}{P_{2k+1}^{(\alpha,\alpha)}(1)} \geqq 0 \tag{5.9}$$

for $x \geqq 0$ when $0 \leqq \lambda \leqq \alpha + 1/2$, when $-1 < \lambda \leqq \alpha + 1/2 \leqq 1$, and when $\alpha \leqq 1/2$, $-1 < \lambda \leqq 1/2$ and $\lambda \geqq -\alpha - 3/2$.

$$\sum_{k=0}^{n} \frac{(\lambda+1)_{n-k}}{(n-k)!} \frac{(\lambda+1)_k}{k!} \frac{(-1)^k P_{2k+1}^{(\alpha,\alpha)}(x)}{P_{2k+1}^{(\alpha,\alpha)}(1)} \geqq 0 \tag{5.10}$$

for $0 \leqq x \leqq 1$ when $0 \leqq \lambda \leqq \alpha + 1/2$ and when $\alpha \geqq 1/2$ and $-1 < \lambda \leqq \alpha + 1/2$.

6. Positivity of an integral of a Bessel function. The projection formula (5.1) also enables us to prove the conjecture in [27] that

$$\int_0^x (x-t)^{\alpha+2\mu-1/2} t^{\alpha+\mu} J_\alpha(t)\, dt \geqq 0, \qquad x > 0, \tag{6.1}$$

when $0 \leqq \mu \leqq 1$, $\alpha + \mu \geqq 1/2$. An expression of this integral as a sum of squares of Bessel functions was used in [27] to show that in order to prove the nonnegativity (positivity) of the integral (6.1) it suffices to prove the nonnegativity (positivity) of the Saalschützian series

$$ {}_5F_4\left[\begin{matrix} -n,\ n+2\alpha+2\mu,\ \alpha+\mu+1,\ \alpha+(\mu+1)/2,\ \alpha+1+\mu/2; \\ \alpha+\mu+1/2,\ \alpha+1,\ (3\alpha+3\mu+3/2)/2,\ (3\alpha+3\mu+5/2)/2 \end{matrix}\right] \tag{6.2}$$

for $n = 0, 1, \cdots$. For $\mu = 0$ and $\mu = 1$ this series reduces to ${}_3F_2$ which was summed in [27] to show that it is positive when $\mu = 0$, $\alpha > 1/2$ and when $\mu = 1$, $\alpha > -1/2$, and that it equals zero when $n \geqq 1$ and either $\mu = 0$, $\alpha = 1/2$ or $\mu = 1$, $\alpha = -1/2$. The series (6.2) reduces to a ${}_4F_3$ when $0 < \mu < 1$, $\alpha + \mu = 1/2$, and so in [27] we were able to apply a limit case of Whipple's formula [16, 4.3(4)] to prove its positivity for this case.

To handle the remaining case $0 < \mu < 1$, $\alpha + \mu > 1/2$, we need but observe that from (5.1) and (4.6) the series (6.2) equals

$$\begin{aligned}
&\sum_{k=0}^{n} \frac{(-1)^k (-n)_k (2\alpha+2\mu)_k ((\alpha+\mu-1/2)/2)_k ((\alpha+\mu+1/2)/2)_k}{k!(2\alpha+2\mu)_{2k} ((3\alpha+3\mu+5/2)/2)_k ((3\alpha+3\mu+3/2)/2)_k} \\
&\quad \cdot \frac{(n+2\alpha+2\mu)_k (\alpha+(\mu+1)/2)_k (\alpha+1+\mu/2)_k}{(\alpha+\mu+1/2)_k (\alpha+1)_k} \\
&\quad \cdot {}_4F_3\left[\begin{matrix} k-n, k+n+2\alpha+2\mu,\ k+\alpha+(\mu+1)/2,\ k+\alpha+1+\mu/2; \\ 2k+2\alpha+2\mu+1,\ k+\alpha+\mu+1/2,\ k+\alpha+1 \end{matrix}\right] \\
&\quad = \sum_{k=0}^{n} \frac{2^{-2k} n!(2\alpha+2\mu)_k (\alpha+\mu-1/2)_{2k} (n+2\alpha+2\mu)_k}{k!(n-k)!(2\alpha+2\mu)_{2k} (3\alpha+3\mu+3/2)_{2k}} \\
&\qquad \cdot \frac{(2\alpha+\mu+1)_{2k} (\mu/2)_{n-k} ((1-\mu)/2)_{n-k}}{(\alpha+\mu+1/2)_n (\alpha+1)_n} \\
&\qquad \cdot {}_4F_3\left[\begin{matrix} k-n+1,\ k+\alpha+3\mu/2,\ k+\alpha+(\mu+1)/2,\ k-n; \\ k-n+1-\mu/2,\ k-n+(\mu+1)/2,\ 2k+2\alpha+2\mu+1 \end{matrix}\right],
\end{aligned}$$

which is obviously positive when $0<\mu<1$, $\alpha+\mu>1/2$. Therefore we have

THEOREM 7. *If* $0\leqq\mu\leqq 1$ *and* $\alpha+\mu\geqq 1/2$, *then inequality* (6.1) *holds and the only cases of equality occur when* $\mu=0$, $\alpha=-1/2$ *or* $\mu=1$, $\alpha=-1/2$.

Additional inequalities for integrals of Bessel functions and for sums and integrals of orthogonal polynomials which follow from our method will be considered elsewhere.

7. Absolutely monotonic functions. Since absolutely monotonic functions have figures so prominently in previous research on inequalities [7], [12], [13] and they can be multiplied and added together to obtain additional absolutely monotonic functions, it is of interest to see which absolutely monotonic functions follow from our results.

From Theorem 1 it follows that the function

$$\begin{aligned} f(t;x,\beta,\lambda) &\equiv \sum_{n=0}^{\infty} t^n \sum_{k=0}^{n} \frac{(\lambda+1)_{n-k}}{(n-k)!}\frac{(\lambda+1)_k}{k!}\frac{(-1)^k L_k^{\beta}(x)}{L_k^{\beta}(0)} \\ &= (1-t^2)^{-\lambda-1}\,{}_1F_1[\lambda+1;\beta+1;xt/(1+t)] \end{aligned} \tag{7.1}$$

is absolutely monotonic for $x\geqq 0$ when $\beta,\lambda\geqq -1/2$. In particular, setting $\lambda=\beta$, we have that

$$(1-t^2)^{-\beta-1}\,e^{xt/(1+t)}$$

is absolutely monotonic for $x\geqq 0$ when $\beta\geqq -1/2$.

For Jacobi polynomials it follows from the generating function [19, 19.10(26)] that

$$\begin{aligned} f(t;x,\alpha,\beta,\lambda) &\equiv \sum_{n=0}^{\infty} t^n \sum_{k=0}^{n} \frac{(\lambda+1)_{n-k}}{(n-k)!}\frac{(\lambda+1)_k}{k!}\frac{P_k^{(\alpha,\beta)}(x)}{P_k^{(\beta,\alpha)}(1)} \\ &= (1-t)^{-\lambda-1}F_4[\lambda+1,\alpha+1;\alpha+1,\beta+1;t(x-1)/2,t(x+1)/2], \end{aligned}$$

where F_4 is the fourth type of Appell's functions of two variables and, for convergence, it is assumed that

$$\left|\frac{t(x-1)}{2}\right|^{1/2}+\left|\frac{t(x+1)}{2}\right|^{1/2}<1.$$

This F_4 function reduces to a simpler function in some special cases [16, p. 102].

For $\lambda=\alpha+\beta$ it was shown in [12] that

$$\begin{aligned} &f(t;x,\alpha,\beta,\alpha+\beta) \\ &\quad=(1-t^2)^{-\alpha-\beta-1}F\left[\frac{\alpha+\beta+1}{2},\frac{\alpha+\beta+2}{2};\beta+1;\frac{2t(1+x)}{(1+t)^2}\right], \end{aligned} \tag{7.3}$$

which, by Theorem 2, is absolutely monotonic for $x\geqq -1$ when $\alpha+\beta\geqq 0$, $\beta\geqq -1/2$ and when $\alpha+\beta\geqq -1$, $\beta\geqq\alpha$. For $\beta=-1/2$ the hypergeometric series in (7.3) can be summed by means of [19, 2.8(5)] to give the absolute monotonicity

of the function

$$
\begin{aligned}
&f(t; x, \alpha, -1/2, \alpha - 1/2) \\
&\quad = \tfrac{1}{2}(1-t)^{-\alpha-1/2}\{(1+t+(2t+2xt)^{1/2})^{-\alpha-1/2} \\
&\qquad + (1+t-(2t+2xt)^{1/2})^{-\alpha-1/2}\}
\end{aligned} \tag{7.4}
$$

for $x \geqq -1$, $\alpha \geqq 1/2$. Replacing t by t^2 and x by $2x^2-1$, this gives the absolute monotonicity of

$$(1-t^2)^{-\alpha-1/2}\{(1+2xt+t^2)^{-\alpha-1/2}+(1-2xt+t^2)^{-\alpha-1/2}\}$$

for $-\infty < x < \infty$, $\alpha \geqq 1/2$. The function (7.3) also reduces to an elementary function in the three cases $\alpha = \beta$, $\alpha = \beta + 1$, $\alpha = -1$ and the function (7.2) reduces in the case $\lambda = \beta$ [12] to give the (previously known) absolute monotonicity of the functions

$$
\begin{aligned}
&(1-t)^{-2\alpha-1}(1-2xt+t^2)^{-\alpha-1/2}, && \alpha \geqq -1/2, \\
&(1+t)(1-t)^{-2\beta-2}(1-2xt+t^2)^{-\beta-3/2}, && \beta \geqq -1/2, \\
&(1-t)^{-\beta}\{1+t+(1-2xt+t^2)^{1/2}\}^{-\beta}, && \beta \geqq 0, \\
&(1-t)^{-\beta-1}\rho^{-1}(1-t+\rho)^{-\alpha}(1+t+\rho)^{-\beta}, && \alpha \geqq -2, \quad \beta \geqq 0,
\end{aligned}
$$

for $x \geqq -1$, where $\rho = (1-2xt+t^2)^{1/2}$.

Two additional reductions of (7.2) can be obtained by using [19, 19.10(11) and 19.10(12)] to write

$$
\begin{aligned}
&f(t; x, \alpha, \alpha, \lambda) \\
&\quad = (1-t)^{-\lambda-1}(1-xt)^{-\lambda-1}F\left[\frac{\lambda+1}{2}, \frac{\lambda+2}{2}; \alpha+1; \frac{t^2(x^2-1)}{(1-xt)^2}\right] \\
&\quad = (1-t)^{-\lambda-1}(1-xt)\rho^{-\lambda-2}F\left[\frac{2\alpha+1-\lambda}{2}, \frac{\lambda+2}{2}; \alpha+1; \frac{t^2(1-x^2)}{\rho^2}\right]
\end{aligned} \tag{7.5}
$$

and

$$
\begin{aligned}
f(t; x, \alpha, \beta, \beta-1) &= 2^{\beta}(1-t)^{-\beta}(1+t+\rho)^{-\beta} \\
&\quad \cdot F\left[\alpha+1, \beta; \beta+1; \frac{1+t-\rho}{2}\right],
\end{aligned} \tag{7.6}
$$

where ρ is as defined above. The function (7.5) is absolutely monotonic for $x \geqq -1$ when $-1 < \lambda \leqq 2\alpha$, and (7.6) is absolutely monotonic for $x \geqq -1$ when $\alpha \geqq -3$, $\beta \geqq 1$.

For $\lambda = 2\alpha + 1$, (7.5) reduces to

$$f(t; x, \alpha, \alpha, 2\alpha+1) = \frac{1-xt}{(1-t)^{2\alpha+2}(1-2xt+t^2)^{\alpha+3/2}}. \tag{7.7}$$

It is the relative simplicity of this function which enabled Askey [7] to prove its absolute monotonicity for $-1 \leqq x \leqq 1$, $\alpha = 1/2$ and then extend this result to $\alpha > 1/2$. Then a standard argument gave (5.4) for $-1 \leqq x \leqq 1$, $2 \leqq \lambda \leqq \alpha + \beta + 1$, and hence the absolute monotonicity of (7.5) for $-1 \leqq x \leqq 1$, $2 \leqq \lambda \leqq 2\alpha + 1$.

When $\alpha = 1/2$ the second hypergeometric function in (7.5) can be factored by using Orr's formula [16, 10.1(5)], and it might be possible to use this factorization to prove the absolute monotonicity of (7.5) for $-1 \leqq x \leqq 1$, $1 < \lambda < 2$, and hence the nonnegativity of the sum (4.25) for $0 < \theta < \pi$, $1 < \lambda < 2$, which is the only open case, [8].

8. Additional observations and open problems. Theorem 3 is best possible in the sense that the $(C, \alpha+\beta+2)$ means of the Poisson kernel (4.12) are not all nonnegative for $-1 \leqq x, y \leqq 1$, $0 \leqq t \leqq 1$, when $\alpha > -1$, $-1 < \beta < -1/2$ or $-1 < \alpha < -1/2$, $\beta > -1$. In view of (4.14) it is enough to show that inequality (4.8) fails for some n when $\alpha > 0$, $-1 < \beta < -1/2$, which is a special case of

THEOREM 8. *Let* $\alpha > -1$, $\lambda > \max(-1, \beta-\alpha-1)$, *and either* $-1 < \beta < -1/2$ *or* $-1 < \beta < 1/2$, $\lambda = \alpha+\beta+1 > 0$. *Then the inequality*

$$\sum_{k=0}^{n} \frac{(\lambda+1)_{n-k}}{(n-k)!} \frac{(\lambda+1)_k}{k!} \frac{P_k^{(\alpha,\beta)}(x)}{P_k^{(\beta,\alpha)}(1)} \geqq 0, \quad -1 \leqq x \leqq 1, \quad n = 0, 1, \cdots, \tag{8.1}$$

fails to hold, and the integral

$$\int_0^x (x-t)^{\lambda} t^{\lambda-\beta} J_\alpha(t)\, dt, \qquad x > 0, \tag{8.2}$$

changes sign infinitely often as $x \to \infty$.

Since the integral (8.2) is a limit case [8] of the sums in (8.1) it suffices to consider the sign of the integral. From the series representation [27, (2.2)] and the asymptotic formula [30, 5.11.4(4)] we have

$$\begin{aligned} & x^{\beta-\alpha-2\lambda-1} \int_0^x (x-t)^{\lambda} t^{\lambda-\beta} J_\alpha(t)\, dt \\ & = \frac{\Gamma(\lambda+1)\Gamma(\lambda+\alpha-\beta+1)2^{-\alpha}}{\Gamma(\alpha+1)\Gamma(2\lambda+\alpha-\beta+2)} {}_3F_2\left[\begin{matrix} \frac{\lambda+\alpha-\beta+1}{2}, \frac{\lambda+\alpha-\beta+2}{2}; \frac{-x^2}{4} \\ \alpha+1, \lambda+\frac{\alpha-\beta+2}{2}, \lambda+\frac{\alpha-\beta+3}{2} \end{matrix}\right] \\ & \sim z^{-(\alpha+\lambda+3/2)/2} \left\{ \frac{A z^{(\beta+1/2)/2}}{\Gamma((\alpha+\beta+1-\lambda)/2)} \right. \\ & \qquad \left. - \frac{B z^{(\beta-1/2)/2}}{\Gamma((\alpha+\beta-\lambda)/2)\Gamma(\lambda/2)} + C\cos\left(2z^{1/2} + O(1)\right) \right\}, \end{aligned} \tag{8.3}$$

for $\lambda > \max(-1, \beta-\alpha-1)$, $\alpha > -1$, where $z = x^2/4$ and A, B, C are positive constants depending only on α, β, λ. The power of z in the first term in braces in (8.3) is negative when $\beta < -1/2$ and this term is zero when $\lambda = \alpha+\beta+1$. Also the power of z in the second term in braces is negative when $\beta < 1/2$ and the coefficient of z in this term is not zero when $\lambda = \alpha+\beta+1 > 0$. Therefore it follows from (8.3) that under the conditions of Theorem 8 the integral (8.2) must change sign infinitely often as $x \to \infty$.

Askey [8] showed that if $\lambda > \alpha + \beta + 1$ and $\alpha > -1$, then the integral (8.2) is not nonnegative for all $x > 0$ and hence (8.1) fails to hold. His results in [8] suggest that inequality (8.1) probably holds for $2\alpha < \lambda < 2\alpha + 1$ when $\alpha = \beta \geqq 1/2$, but our method cannot be used to show this without some modifications since $d_{n,j}^{\alpha,\alpha,\lambda}$ assumes negative values for sufficiently large n when $-1 < 2\alpha < \lambda < 2\alpha + 1$. One way to see this is to reverse the series (4.21) to obtain

$$d_{n,j}^{\alpha,\alpha,\lambda} = \frac{(\alpha+1/2)_{m-j}(n-m+1+\alpha+\lambda/2)_{m-j}(j+1+\lambda/2)_{m-j}}{(n-m-j+\alpha)_{m-j}(n-m+\alpha+1/2)_{m-j}(n-m+\lambda+1)_{m-j}}$$
$$\cdot {}_4F_3\left[\begin{matrix} j-m, n-m-j+\alpha, (\lambda+1)/2, \alpha-\lambda/2; \\ \alpha+1/2, n-m+1+\alpha+\lambda/2, -m-\lambda/2 \end{matrix}\right]$$

and then to observe that for fixed j this ${}_4F_3$ series tends to minus infinity as $n \to \infty$, when $-1 < 2\alpha < \lambda < 2\alpha + 2$. Similarly, it can be shown that $d_{n,j}^{\alpha,\beta,\lambda} < 0$ for sufficiently large n when $\alpha > \beta > -1$, $\alpha + \beta < 0$, $\lambda = \alpha + \beta$, and when $-2 < \alpha + \beta < 0$, $-1 < \beta < 0$, $\lambda = 0$; so that even by using all of the known transformation formulas one could not extend Theorems 2 and 4 to other α, β by the method of this paper. It seems likely that (4.18) also holds for $-1 < \alpha < 1/2$, $\beta(\alpha) \leqq \beta < 0$, where $\beta(\alpha)$ is the unique solution to the equation (see [14], [31])

$$\int_0^{j_{\alpha,2}} t^{-\beta(\alpha)} J_\alpha(t)\, dt = 0, \qquad -1/2 < \beta(\alpha) < 0,$$

$j_{\alpha,2}$ being the second positive zero of $J_\alpha(t)$; and so another method will have to be developed to prove this.

For Laguerre polynomials it should be pointed out that the restriction $\beta \geqq -1/2$ in Theorem 1 cannot be relaxed for any $\lambda > -1$. For, setting $n = 2m + 1$, reversing the order of summation on the right side of (2.6), replacing x by $2x/m$ and using the limit relation

$$\lim_{n\to\infty} m^{-\beta} L_m^\beta(x/m) = x^{-\beta/2} J_\beta(2x^{1/2}),$$

we obtain a sum of squares of Bessel functions which, by [27, (3.3)], is a positive multiple of

(8.4) $$ {}_1F_2[1; 2, \beta+2; -4x]$$

for $x > 0$, $\beta > -1$. Then the asymptotic formula [30, 5.11.4(4)] shows that (8.4) changes sign infinitely often as $x \to \infty$ when $-1 < \beta < -1/2$. Similarly the case $n = 2m$ leads to the function

$$1 - \frac{2x}{\beta+1} {}_1F_2[1; 2, \beta+2; -4x]$$

which also changes sign infinitely often as $x \to \infty$ when $-1 < \beta < -1/2$. It is not known whether the restriction $\lambda \geqq -1/2$ in Theorem 1 can be relaxed when $\beta > -1/2$. However, when $n = 2$ and $\beta = -1/2$ the inequality (1.2) cannot hold for any λ, $-1 < \lambda < -1/2$, since then

$$S_2\left(\frac{\beta+2}{\lambda+2}; \beta, \lambda\right) = (\lambda+1)\left(1 - \frac{\beta+2}{2(\lambda+2)(\beta+1)}\right),$$

which is the minimum value of $S_2(x; \beta, \lambda)$, is negative.

For Jacobi polynomials a computation shows that the minimum value of $S_2(x;\alpha,\beta,\lambda)$ is

$$(8.5)\qquad (\lambda+1)\left(1-\frac{(\beta+2)(\lambda+\alpha+\beta+4)^2}{2(\lambda+2)(\beta+1)(\alpha+\beta+3)(\alpha+\beta+4)}\right),$$

whose nonnegativity is a necessary condition for (8.1) to hold. From (8.5) we obtain the particularly interesting observation that if $\beta=-1/2$, then (8.1) cannot hold for any $\lambda>-1$ when $-1/2\leqq\alpha<1/2$, for any $\lambda\neq 0$ when $\alpha=1/2$, and for any λ in the interval $-1<\lambda\leqq -1/2$ when $\alpha>1/2$. Theorem 1 and (8.5) suggest that inequality (8.1) probably holds for some $\lambda<0$ when $\alpha>1/2$, $\beta=-1/2$. It should also be observed that since the variable x appears to the power $n-2j$ in the expansion (2.6) it follows from our results for $C_{n,j}$ that if $\lambda,\ \beta\geqq -1/2$, then $S_{2n}(x;\beta,\lambda)\geqq 0$ and $S_{2n+1}(x;\beta,\lambda)\leqq 0$ when $x<0$. Similarly, $S_{2n}(x;\alpha,\beta,\lambda)\geqq 0$ and $S_{2n+1}(x;\alpha,\beta,\lambda)\leqq 0$ for $x<-1$ whenever the coefficients in (3.4) are nonnegative for all n.

Among the open problems and conjectures pointed out in [12] is the conjecture that if $\alpha+\beta+\mu\geqq 0$, $0\leqq\mu\leqq 1$, $\beta\geqq -1/2$, then

$$(8.6)\qquad \sum_{k=0}^{n}\frac{(\alpha+\beta+1+2\mu)_{n-k}}{(n-k)!}\frac{(\alpha+\beta+1)_k}{k!}\frac{\left(\frac{\alpha+\beta+1+2\mu}{2}\right)_k}{\left(\frac{\alpha+\beta+1}{2}\right)_k}\frac{P_k^{(\alpha,\beta)}(x)}{P_k^{(\beta,\alpha)}(1)}\geqq 0$$

for $-1\leqq x\leqq 1$. This sum is a common generalization of the two sums in (4.14), and it has the integral (6.1) as a limit case. Since (8.6) is not zero at $x=-1$ for all odd n when $0<\mu<1$, it could not have an expansion of the form (3.4); so another expansion or method will be needed to prove this conjecture for $0<\mu<1$. Another generalization of the two sums in (4.14) which also has the integral (6.1) as a limit case and might be nonnegative under the above conditions is the sum

$$(8.7)\qquad \sum_{k=0}^{n}\frac{(\alpha+\beta+1+2\mu)_{n-k}}{(n-k)!}\frac{(\alpha+\beta+1)_k}{k!}\frac{\left(\frac{\alpha+\beta+2+\mu}{2}\right)_k}{\left(\frac{\alpha+\beta+2-\mu}{2}\right)_k}\frac{P_k^{(\alpha,\beta)}(x)}{P_k^{(\beta,\alpha)}(1)}.$$

Computationally, this sum should be easier to handle since, unlike (8.6), it at least reduces to a nearly-poised ${}_3F_2(-1)$ series when $x=-1$.

The sum

$$(8.8)\qquad \sum_{k=0}^{n}\frac{(\gamma+1)_{n-k}}{(n-k)!}\frac{(2k+\alpha+\beta+1)(\alpha+\beta+1)_k}{k!(\alpha+\beta+1)}\frac{P_k^{(\alpha,\beta)}(x)}{P_k^{(\beta,\alpha)}(1)}$$

also reduces to a nearly-poised ${}_3F_2(-1)$ series when $x=-1$. So, if one could prove the above conjecture for the sum (8.7), then it might be possible to use the same method to solve the problem of finding for which α,β with $-1<\beta<-1/2$ the sums (8.8) are nonnegative for $-1\leqq x\leqq 1$ for some $\gamma=\gamma(\alpha,\beta)>0$. This would then yield an extension of Askey's results [5, Thm. 10] on the L^p-convergence of Lagrange interpolation polynomials at the zeros of Jacobi polynomials.

Another interesting problem is to find an extension of the expansion (2.1) to Jacobi polynomials which would give the Turán type inequality

$$\Delta_n(x;\alpha,\beta)\equiv\left\{\frac{P_n^{(\alpha,\beta)}(x)}{P_n^{(\alpha,\beta)}(1)}\right\}^2-\frac{P_{n-1}^{(\alpha,\beta)}(x)}{P_{n-1}^{(\alpha,\beta)}(1)}\frac{P_{n+1}^{(\alpha,\beta)}(x)}{P_{n+1}^{(\alpha,\beta)}(1)}\geqq 0,\qquad -1\leqq x\leqq 1, \tag{8.9}$$

for $\beta\geqq\alpha\geqq -1$, $n\geqq 1$. This inequality was proved in [25] by a method which actually gave the stronger result that if $\alpha,\beta>-1$ and $n\geqq 1$, then

$$\Delta_n(x;\alpha,\beta)-\frac{(\beta-\alpha)(1-x)}{2(n+\alpha+1)(n+\beta)}\left\{\frac{P_n^{(\alpha,\beta)}(x)}{P_n^{(\alpha,\beta)}(1)}\right\}^2\geqq 0,\qquad -1\leqq x\leqq 1, \tag{8.10}$$

with equality only for $x=\pm 1$. Thus it is preferable (and probably easier) to prove (8.9) by first proving (8.10). This suggests that to prove (8.6), (8.7), or (8.8) by use of the Burchnall and Chaundy formulas [17] it might be necessary to first subtract an appropriate nonnegative function and actually prove a stronger result.

REFERENCES

[1] W. A. Al-Salam and L. Carlitz, *Generalized Turán expressions for certain hypergeometric series*, Portugal. Math., 16 (1957), pp. 119–127.

[2] R. Askey, *Positive Jacobi polynomial sums*, Tôhoku Math. J., 24 (1972), pp. 109–119.

[3] ———, *Positivity of the Cotes numbers for some Jacobi abscissas*, Numer. Math., 19 (1972), pp. 46–48.

[4] ———, *Mean convergence of orthogonal series and Lagrange interpolation*, Acta Math. Acad. Sci. Hungar., 23 (1972), pp. 71–85.

[5] ———, *Summability of Jacobi series*, Trans. Amer. Math. Soc., 179 (1973), pp. 71–84.

[6] ———, *Refinements of Abel summability for Jacobi series*, Harmonic Analaysis on Homogeneous Spaces, Proc. Symp. Pure Math., vol. 26, C. Moore, ed., American Mathematical Society, Providence, R.I., 1973, pp. 335–338.

[7] ———, *Some absolutely monotonic functions*, Studia Sci. Math. Hungar., 9 (1974), pp. 51–56.

[8] ———, *Positive Jacobi polynomial sums*, III, Linear Operators and Approximation, II, P. L. Butzer and B. Sz. Nagy, eds., ISNM 25 Birkhäuser Verlag, Basel, 1974, pp. 305–312.

[9] ———, *Orthogonal Polynomials and Special Functions*, Regional Conference Series in Applied Mathematics, Society for Industrial and Applied Mathematics, Philadelphia, 1975.

[10] R. Askey and J. Fitch, *Positivity of the Cotes numbers for some ultraspherical abscissas*, SIAM J. Numer. Anal., 5 (1968), pp. 199–201.

[11] ———, *Integral representations for Jacobi polynomials and some applications*, J. Math. Anal. Appl., 26 (1969), pp. 411–437.

[12] R. Askey and G. Gasper, *Positive Jacobi polynomial sums, II*, Amer. J. Math. to appear.

[13] R. Askey and H. Pollard, *Some absolutely and completely monotonic functions*, this Journal, 5 (1974), pp. 58–63.

[14] R. Askey and J. Steinig, *Some positive trigonometric sums*, Trans. Amer. Math. Soc., 187 (1974), pp. 295–307.

[15] ———, *A monotonic trigonometric sum*, Amer. J. Math., 98 (1976), pp. 356–365.

[16] W. N. Bailey, *Generalized Hypergeometric Series*, Cambridge University Press, Cambridge, England, 1935.

[17] J. L. Burchnall and T. W. Chaundy, *Expansions of Appell's double hypergeometric functions*, Quart. J. Math. (Oxford), 11 (1940), pp. 249–270.

[18] ———, *Expansions of Appell's double hypergeometric functions (II)*, Quart. J. Math. (Oxford), 12 (1941), pp. 112–128.

[19] A. ERDÉLYI, W. MAGNUS, F. OBERHETTINGER AND F. G. TRICOMI, *Higher Transcendental Functions*, vols. I, II, III, McGraw-Hill, New York, 1953.
[20] L. FEJER, *Einige Sätze, . . .*, Monatsh. Math. und Physik, 35 (1928), pp. 305–344; Gesammelte Arbeiten, II, pp. 202–237.
[21] ———, *Über die Laplacesche Reihe*, Math. Ann., 67 (1909), pp. 76–109; Gesammelte Arbeiten, I, pp. 503–537.
[22] ———, *Neue Eigenschaften der Mittelwerte bei den Fourierreihen*, J. London Math. Soc., 8 (1933), pp. 53–62; Gesammelte Arbeiten II, pp. 493–501.
[23] G. GASPER, *Positivity and the convolution structure for Jacobi series*, Ann. of Math., 93 (1971), pp. 122–118.
[24] ———, *Banach algebras for Jacobi series and positivity of a kernel*, Ibid., 95 (1972), pp. 261–280.
[25] ———, *An inequality of Turán type for Jacobi series*, Proc. Amer. Math. Soc., 32 (1972), pp. 435–439.
[26] ———, *Projection formulas for orthogonal polynomials of a discrete variable*, J. Math. Anal. Appl., 45 (1974), pp. 176–198.
[27] ———, *Positive integrals of Bessel Functions*, this Journal, 6 (1975), pp. 868–881.
[28] ———, *Positivity and special functions*, Theory and Applications of Special Functions, R. Askey, ed., Academic Press, New York, 1975, pp. 375–434.
[29] E. KOGBETLIANTZ, *Recherches sur la sommabilité des séries ultrasphériques par la méthode des moyennes arithmétiques*, J. Math. Pures Appl., (9) 3 (1924), pp. 107–187.
[30] Y. L. LUKE, *The Special Functions and their Approximations*, vol. 1, Academic Press, New York, 1969.
[31] E. MAKAI, *An integral inequality satisfied by Bessel functions*, Acta Math. Acad. Sci. Hungar., 25 (1974), pp. 387–390.
[32] M. S. ROBERTSON, *The coefficients of univalent functions*, Bull. Amer. Math. Soc., 51 (1945), pp. 733–738.
[33] G. SZEGÖ, *Orthogonal Polynomials*, Colloquium Publications, vol. 23, third ed., American Mathematical Society, Providence, R.I., 1967.

SIAM J. MATH. ANAL.
Vol. 8, No. 3, May 1977

SOLUTIONS OF NONLINEAR OPERATOR EQUATIONS*

PETER LANCASTER AND JON G. ROKNE†

Abstract. Some theorems concerning existence and uniqueness of zeros of operator polynomials are given. Under certain hypotheses we show the existence of complete pairs of zeros. A numerical example is given applying the theorems.

1. Introduction. Let $\mathscr{X}$ be a Banach space over the complex field $\mathscr{C}$ and let $L = L(\mathscr{X}, \mathscr{X})$ be the noncommutative Banach algebra (with density I) of all bounded linear operators from $\mathscr{X}$ to itself. It is assumed that the norm on L is such that $\|I\| = 1$, and we denote the spectrum of $A \in L$ by $\sigma(A)$.

We consider functions $R: L \to L$, and we are interested in operators $X \in L$ for which $R(X) = 0$. Such an operator will be called a zero of R.

In particular, we consider polynomial functions F defined by

$$F(X) = \sum_{i=0}^{\infty} A_i X^i, \tag{1}$$

where $X, A_0, A_1, \cdots \in L$. We trust that no confusion will be created by using the same symbol F for the associated function from $\mathscr{C}$ to L defined for all $\lambda \in \mathscr{C}$ by

$$F(\lambda) = \sum_{i=0}^{\infty} A_i \lambda^i. \tag{2}$$

For polynomial operators we define the degree to be the largest index i for which $A_i \neq 0$. If the degree of a polynomial operator F is n and X is a zero of F, then it follows that

$$F(\lambda) = \sum_{i=0}^{n} A_i \lambda^i = Q(\lambda)(I\lambda - X) \tag{3}$$

for a polynomial Q of degree $n-1$ and all $\lambda \in \mathscr{C}$. It is important to note that the order of the factors on the right cannot generally be inverted.

The case when the degree of the polynomial operator is two is of particular interest in mechanics. We will study this case in detail. For this purpose, we write the second degree equation as

$$F(X) \equiv AX^2 + BX + C = 0. \tag{4}$$

If, for the moment, we assume that $A = I$ and that $(B^2 - 4C)^{1/2}$ exists and commutes with B, then (4) has two solutions,

$$Z_{1,2} = -\tfrac{1}{2}B \pm \sqrt{\tfrac{1}{4}B^2 - C}.$$

These are seen to be zeros of (4) by substitution. Furthermore, $Z_1 - Z_2 = (B^2 - 4C)^{1/2}$. If B^{-1} exists as well, then $Z_1 - Z_2 = B(I - 4B^{-2}C)^{1/2}$. A "strong damping" hypothesis implies that B is "large" compared to A and C. For example, if $4\|B^{-2}C\| < 1$, then the Banach lemma leads to the conclusion that

* Received by the editors May 22, 1975.

† Department of Mathematics, University of Calgary, Calgary, Alberta, Canada T2N IN4.

$Z_1 - Z_2$ is invertible. We say that if $Z_1 - Z_2$ has an inverse, then $Z_{1,2}$ is a complete pair of roots.

The existence of a complete pair is important for the following reason (Eisenfeld [2], Krein and Langer [5] and Lancaster [6]): If Z_1, Z_2 form a complete pair for F, then we have the factorization

$$F(\lambda) = A(Z_1 - Z_2)(\lambda I - Z_2)(Z_1 - Z_2)^{-1}(\lambda I - Z_1).$$

From this we deduce the relation between resolvent operators:

$$F(\lambda)^{-1}A(Z_1 - Z_2) = (\lambda I - Z_1)^{-1} - (\lambda I - Z_2)^{-1},$$

which indicates clearly the division of the spectrum of F between the spectra of Z_1 and Z_2.

The strongest results on existence known to date relate to cases in which A, B, C are symmetric and are due to Krein and Langer [5] and Langer [7]. In this paper we focus on some results which involve minimal assumptions with regard to the symmetry of the coefficient operators.

The case of infinite-dimensional Banach spaces is of considerable interest in the analysis of some physical problems formulated as differential equations ([5], for example). For the purpose of numerical analysis, the continuous problem is replaced by a discrete one using variational or finite difference methods. In this way, we arrive at a problem on a finite-dimensional space $\mathscr{X}$. If, in this case, a matrix solution of the (now matrix) equation $F(\mathscr{X}) = 0$ can be found, this yields a "packet" of information on n eigenvalues and associated eigenspaces of $F(\lambda)$ if $\mathscr{X}$ has dimension n. This situation is illustrated in our numerical example.

The key to the theorems that we will prove is the Newton–Kantorovich theorem. We state the theorem as it appears in [9] where a simple proof may be found.

THEOREM 1. *Let $\mathscr{X}$, $\mathscr{Y}$ be Banach spaces, $\mathscr{D} \subset \mathscr{X}$ and suppose $G: \mathscr{D} \to \mathscr{Y}$. Assume that on an open set $\mathscr{D}_0 \subset \mathscr{D}$, G is Fréchet differentiable and that*

$$\|G'_X - G'_Y\| \leqq \kappa \|X - Y\|, \qquad X, Y \in \mathscr{D}_0.$$

Given $X_0 \in \mathscr{D}_0$, assume that $\Gamma_0 = [G'_{X_0}]^{-1}$ is defined on all of $\mathscr{Y}$. Let $\|\Gamma_0\| \leqq \beta$ and $\|\Gamma_0 H(X_0)\| \leqq \eta$. Suppose $h = \eta\beta\kappa \leqq \frac{1}{2}$ and set

$$t^* = \frac{1}{\beta\kappa}(1 - \sqrt{1 - 2h}),$$

$$t^{**} = \frac{1}{\beta\kappa}(1 + \sqrt{1 - 2h}),$$

and suppose $S = \{X \mid \|X - X_0\| \leqq t^*\} \subset \mathscr{D}_0$. Then the interates

$$X_{k+1} = X_k - [G'_{X_k}]^{-1}G(X_k), \qquad k = 0, 1, \cdots,$$

are well-defined, lie in S and converge to a solution X^* of $G(X) = 0$ which is unique in $\mathscr{D}_0 \cap \{X \mid \|X - X_0\| < t^{**}\}$. If $h < \frac{1}{2}$, we obtain the rapid convergence generally associated with Newton's method [3].

The usefulness of this theorem is usually confined to the properties of the algorithm. Here, we also have special interest in the statements regarding *existence* of zeros of $F(X)$.

2. Existence theorems for equations of second degree with strong damping. We consider the second degree equation as written in (4). In addition we will assume that B^{-1} exists and is "small" in an appropriate sense. We are going to apply Theorem 1 to this equation and for this we need the Fréchet derivative of F. This may be written as

$$F'_X(H) = AXH + AHX + BH,$$

and we immediately find that

$$\|F'_X - F'_Y\| \leqq 2\|A\|\|X - Y\|$$

for all $X, Y \in L$. This gives $2\|A\|$ as a bound for κ needed in Theorem 1.

Since B^{-1} exists, $B^{-1}F(X) = B^{-1}AX^2 + X + B^{-1}C$ has the same zeros as F. Our first result concerns the existence of a "small" zero of F.

THEOREM 2. *Let F be as defined in* (4) *and B be invertible. Suppose* $h = 2\|B^{-1}A\|\|B^{-1}C\| \leqq \frac{1}{2}$, *and define*

$$t^* = \frac{1}{2\|B^{-1}A\|}(1 - \sqrt{1-2h}),$$

$$t^{**} = \frac{1}{2\|B^{-1}A\|}(1 + \sqrt{1-2h}).$$

Then $F(X) = 0$ has a solution in the sphere $S = \{X \mid \|X\| \leqq t^*\}$ which is unique in the sphere $T = \{X \mid \|X\| < t^{**}\}$.

Proof. We apply the Newton–Kantorovich theorem to $B^{-1}F$ with $X_0 = 0$. Then $B^{-1}F(X_0) = B^{-1}C$ and $B^{-1}F'_{X_0}(H) = H$. This means that we may take $\beta = \|I\| = 1$ and $\eta = \|B^{-1}C\|$. From above we get (replacing F by $B^{-1}F$) $\kappa = 2\|B^{-1}A\|$. With the hypotheses made in this theorem, the Newton–Kantorovich hypotheses are satisfied and the conclusions of Theorem 2 follow.

The Newton–Kantorovich theorem will also furnish convergence results for the iteration. We omit the statement of these, however, both for this theorem and for subsequent theorems for the sake of brevity.

In Theorem 2 we needed the invertibility of B. We now assume, in addition, that A is invertible and write

$$A^{-1}F(X) = X^2 + A^{-1}BX + A^{-1}C.$$

We can now assert the existence of a "large" solution near $-A^{-1}B$.

THEOREM 3. *Let F be defined as in* (4) *and assume A and B to be invertible. Suppose*

$$h = 2\|B^{-1}A\|\|A^{-1}CB^{-1}A\| \leqq \tfrac{1}{2}$$

and

$$t^* = \frac{1}{2\|B^{-1}A\|}(1-\sqrt{1-2h}),$$

$$t^{**} = \frac{1}{2\|B^{-1}A\|}(1+\sqrt{1-2h}).$$

Then $F(X)=0$ *has a solution in the sphere* $S=\{X|\|X+A^{-1}B\|\leqq t^*\}$ *which is unique in the sphere* $T=\{X|\|X+A^{-1}B\|<t^{**}\}$.

Proof. Let $X_0=-A^{-1}B$ in the Newton–Kantorovich theorem applied to $G\equiv A^{-1}F$. Then $G(X_0)=A^{-1}C$, and this gives bounds $\beta=\|\Gamma_0\|=\|B^{-1}A\|$ and $\eta=\|A^{-1}CB^{-1}A\|$. The hypotheses of Theorem 1 applied to G are now satisfied and the conclusions follow.

We now make a stronger "strong damping" hypothesis to deduce that the roots obtained from Theorems 2 and 3 form a complete pair.

THEOREM 4. *Let* F *be defined by* (4) *with* $A=I$ *and* B *invertible. If* $h=2\|B^{-1}\|\max(\|B^{-1}C\|, \|CB^{-1}\|)<\frac{1}{2}$, *then the roots* X_1, X_2 *of* F *of Theorems* 2 *and* 3 *form a complete pair.*

Proof. By Theorem 2 there is a root X_1 of F within the closed sphere S_1 of radius $t=\frac{1}{2}\|B^{-1}\|^{-1}(1-\sqrt{1-2h})$ and center $X=0$. Theorem 3 implies that there is a root X_2 of F within the closed sphere S_2 of the same radius t with center at $-B$. Since $\|B^{-1}\|^{-1}\leqq\|B\|$, we have $2t<\|B\|$, and hence the two spheres have no points in common. Thus $X_1\neq X_2$.

Now we write

$$-X_2=B-(B+X_2)=B[I-B^{-1}(B+X_2)].$$

Since $B+X_2\in S_2$, we have $\|B+X_2\|\leqq t$ and so

$$\|B^{-1}(B+X_2)\|\leqq\|B^{-1}\|t<\tfrac{1}{2}.$$

The Banach lemma then implies that X_2 is invertible and

$$\|X_2^{-1}\|\leqq\|B^{-1}\|(1-\|B^{-1}\|t)^{-1}.$$

Then $\|X_1\|<t$ implies

$$\|X_2^{-1}X_1\|\leqq\|B^{-1}\|t(1-\|B^{-1}\|t)^{-1}=\frac{1-\sqrt{1-2h}}{1+\sqrt{1-2h}}<1.$$

Since we may write $X_2-X_1=X_2^{-1}(I-X_2^{-1}X_1)$, the Banach lemma implies that X_2-X_1 is invertible and

$$\|(X_2-X_1)^{-1}\|\leqq\|X_2^{-1}\|(1-\|X_2^{-1}X_1\|)^{-1}\leqq\frac{\|B^{-1}\|}{\sqrt{1-2h}}.$$

The theorem is now proved, and we note that as h approaches $\frac{1}{2}$, the inverse may become unbounded. We also note that the spectrum of X_2 strictly dominates the spectrum of X_1. First, it is easily verified that, if $A\in L$ and $\|A^{-1}\|\leqq k$, then $\lambda\in\sigma(A)$ implies $1/k\leqq|\lambda|\leqq\|A\|$. Then the bound on $\|X_2^{-1}\|$ implies that if

$\lambda \in \sigma(X_2)$, then

$$|\lambda| \geqq \frac{1+\sqrt{1-2h}}{2\|B^{-1}\|} \geqq \|X_1\| \geqq \sup\{|\mu|: \mu \in \sigma(X_1)\}.$$

This property is required of some suggested interative methods for the solution of equations in matrices ([1] and [8]).

Following the line of argument of Eisenfeld [2], we can arrive at a complete pair under apparently weaker conditions than those of Theorem 4, but without the implicit convergence results of Theorem 3. Thus under the hypotheses of Theorem 2 let $G(X) = X^2(B^{-1}A) + X + (B^{-1}C)$. By the method of proof of that theorem, one easily establishes the existence of an $\hat{X} \in S$ which is unique in T and which satisfies $G(\hat{X}) = 0$. Now, under the further hypothesis that A^{-1} exists, we define $Y = -A^{-1}B - A^{-1}B\hat{X}B^{-1}A$, and it is easily seen that $B^{-1}F(Y) = G(\hat{X}) = 0$. Clearly,

$$\|Y + A^{-1}B\| \leqq \|A^{-1}B\|\|B^{-1}A\|\|\hat{X}\|,$$

and since $\hat{X} \in S$, it follows that

$$\|Y + A^{-1}B\| \leqq \tfrac{1}{2}\|A^{-1}B\|(1-\sqrt{1-2h})$$

and h is as defined in Theorem 2. To see that $\hat{X}$ and Y form a complete pair observe that $B^{-1}AY = -I - \hat{X}B^{-1}A$, whence

$$B^{-1}A(\hat{X} - Y) = I - D,$$

where $D = -B^{-1}A\hat{X} - \hat{X}B^{-1}A$, and so $\|D\| \leqq 2\|B^{-1}A\|t^* = 1 - \sqrt{1-2h} < 1$ provided $h < \frac{1}{2}$. It then follows that $\hat{X} - Y$ is invertible as we set out to prove.

Associated with the function F: $C \to L$ we have the derivative with respect to λ : $F^{(1)}(\lambda) = 2A\lambda + B$. The following result generalizes Theorem 2 which is obtained on putting $a = 0$.

THEOREM 5. *If for a complex number a it holds that $F^{(1)}(a)^{-1}$ exists and*

$$h = 2\|F^{(1)}(a)^{-1}F(a)\|\|F^{(1)}(a)^{-1}A\| \leqq \tfrac{1}{2},$$

then with

$$t^* = \frac{1}{2\|F^{(1)}(a)^{-1}A\|}(1-\sqrt{1-2h}),$$

$$t^{**} = \frac{1}{2\|F^{(1)}(a)^{-1}A\|}(1+\sqrt{1-2h}),$$

a zero of F exists in the sphere $S = \{X | \|X - aI\| \leqq t^{**}\}$.

The proof is obtained by applying Theorem 1 to the function $G = F^{(1)}(a)^{-1}F$.

3. Existence theorems for equations of second degree with weak damping. The terminology weak damping is taken from mechanics where the norm of B in (4) is small. If $B = 0$, then (4) reduces to

$$AX^2 + C = 0,$$

which has a solution if A^{-1} exists and if $(-A^{-1}C)^{1/2}$ exists. For the next theorem we assume $A=I$ and that $C^{1/2}$ exists, i.e., there exists $C^{1/2}\in L$ such that $(C^{1/2})^2=C$. We then show that if B is small in an appropriate sense, we still get zeros for the function F of (4).

For the purpose of this theorem only we assume that $\mathscr{X}$ is a Hilbert space and denote by A^* the (Hilbert) adjoint of $A\in L$.

THEOREM 6. *Suppose* $\gamma=\inf\sigma(C^{1/2}+(C^{1/2})^*)/2>0$, $\|B\|<2\gamma$, *and* $h=2\|BC^{1/2}\|/(2\gamma-\|B\|)^2\leqq\frac{1}{2}$. *Define*

$$t^*=\frac{2\gamma-\|B\|}{2}(1-\sqrt{1-2h}),$$

$$t^{**}=\frac{2\gamma-\|B\|}{2}(1+\sqrt{1-2h}).$$

Then (4) *has a solution in the sphere* $S=\{X\,|\,\|X\mp iC^{1/2}\|<t^*\}$ *which is unique in the sphere* $T=\{X\,|\,\|X\mp iC^{1/2}\|<t^{**}\}$.

Proof. We only prove the case when the spheres S and T are centered around $iC^{1/2}$. Let $X_0=iC^{1/2}$ in the Newton–Kantorovich theorem. Then $F(X_0)=iBC^{1/2}$ and

$$F'_{X_0}(H)=H(iC^{1/2})+(iC^{1/2}+B)H. \tag{5}$$

We now need estimates for $\|\Gamma_0\|$ and $\|\Gamma_0F(X_0)\|$. Consider first the equation

$$\Delta C^{1/2}+(C^{1/2}-iB)\Delta=P, \tag{6}$$

and define $T(\Delta)=\Delta C^{1/2}+C^{1/2}\Delta$. Then $T\in L$. Since $\gamma>0$ it follows that if $\lambda\in\sigma(C^{1/2})$ $(\mu\in\sigma(C^{1/2})^*)$ then $\operatorname{Re}\lambda>0$ $(\operatorname{Re}\mu>0)$. Hence $\sigma(C^{1/2})\cap\sigma(-C^{1/2})=\phi$. By a corollary of Rosenblum [10], T^{-1} exists and we may define $\varphi: L\to L$ by $\varphi(X)=T^{-1}(P+iBX)$. Obviously, if Δ satisfies (6), then Δ is a fixed point of φ and conversely. Thus to prove the existence of a solution of (6) we may apply the contraction mapping principle to φ. We have

$$\varphi(X)-\varphi(Y)=T^{-1}(iB(X-Y)),$$

which gives

$$\|\varphi(X)-\varphi(Y)\|\leqq\|T^{-1}\|\|B\|\|X-Y\|.$$

Using the definition of γ, a lemma of Heinz [4] yields $\|T^{-1}\|\leqq 1/(2\gamma)$, and hence the contraction property follows from the fact that $\|B\|<2\gamma$.

Since $T(\Delta)=P+iB\Delta$, we have

$$\|\Delta\|\leqq\|T^{-1}\|(\|P\|+\|B\|\|\Delta\|)\leqq\frac{\|P\|}{2\gamma}+\frac{\|B\|}{2\gamma}\|\Delta\|$$

from which

$$\|\Delta\|\leqq\frac{\|P\|}{2\gamma-\|B\|}.$$

Since (5) implies that $-iF'_{X_0}(H)=HC^{1/2}+(C^{1/2}-iB)H$, it follows that

$$\|\Gamma_0F(X_0)\|\leqq\frac{\|BC^{1/2}\|}{2\gamma-\|B\|}=\eta$$

and

$$\|\Gamma_0\| \leqq \frac{1}{2\gamma - \|B\|} = \beta.$$

Applying Theorem 1 to F, we have $\kappa = 2$ and so the hypothesis on h yields $h = \eta\beta\kappa \leqq \frac{1}{2}$. Our conclusions therefore follow Theorem 1.

We note that a similar theorem holds if we define $X_0 = \pm iC^{1/2} - B$.

4. Existence theorems for more general equations. Assume now that our operator-valued function can be written in the form

$$F(X) = A_0 + A_1X + P(X). \tag{7}$$

In the next theorem we prove a natural generalization of Theorem 2 on hypotheses which ensure that $P(X)$ behaves sufficiently like a power of X higher than the first.

THEOREM 7. *Suppose that, in* (7), A_1^{-1} *exists and that P satisfies*

(i) *P is Fréchet differentiable in an open set* $\mathscr{D}_0$,

(ii) $P(0) = 0$,

(iii) $P_0'(H) = 0$ *for all* $H \in L$,

(iv) $\exists \rho \geqq 0$ *such that* $\|A_1^{-1}(P_X' - P_Y')\| \leqq \rho\|X - Y\| \quad \forall X, Y \in \mathscr{D}_0$.

If $h = \rho\|A_1^{-1}A_0\| \leqq \frac{1}{2}$ *and we define*

$$t^* = \frac{1}{\rho}(1 - \sqrt{1-2h}),$$

$$t^{**} = \frac{1}{\rho}(1 + \sqrt{1-2h}),$$

then the function (7) *has a zero in the sphere* $S = \{X \mid \|X\| \leqq t^*\}$ *unique in the sphere* $T = \{X \mid \|X\| < t^{**}\}$.

Proof. We apply the Newton–Kantorovich theorem to $G(X) = A_1^{-1}A_0 + X + A_1^{-1}P(X)$. Since

$$G_X'(H) - G_Y'(H) = A_1^{-1}\{P_X'(H) - P_Y'(H)\},$$

we may use (iv) and take $\kappa = \rho$. Then if $X_0 = 0$, $\Gamma_0 = (G_{X_0}')^{-1} = I$ and $\Gamma_0 G(X_0) = A_1^{-1}A_0$. Thus we can take $\beta = 1$, $\eta = \|A_1^{-1}A_0\|$ and the theorem follows. We note that Theorem 2 is obtained with $P(X) = A_2X^2$.

We can apply Theorem 7 to prove the following result for a cubic polynomial. We write

$$F(X) = A_0 + A_1X + A_2X^2 + A_3X^3. \tag{8}$$

Note that, when $A_3 = 0$, condition (9) reduces to the (strict) strong damping condition of Theorem 2.

THEOREM 8. *If, in* (8), A_1^{-1} *exists and*

$$4\|A_1^{-1}A_0\|(6\|A_1^{-1}A_3\|\|A_1^{-1}A_0\| + \|A_1^{-1}A_2\|) < 1, \tag{9}$$

then F has a zero in the sphere with center at the origin (in L) and radius $2\|A_1^{-1}A_0\|$.

Proof. Comparing (8) with (7) we write $P(X)=A_2X^2+A_3X^3$, and it is a straightforward calculation to show that, if $X, Y\in\{X\in L:\|X\|\leqq r\}$, then in hypothesis (iv) of Theorem 7, we may take

$$\rho = 6r\|A_1^{-1}A_3\|+2\|A_1^{-1}A_2\|.$$

Using inequality (9) we may choose

$$r=\frac{1-4\|A_1^{-1}A_0\|\|A_1^{-1}A_2\|}{12\|A_1^{-1}A_0\|\|A_1^{-1}A_3\|}>0,$$

in which case it is found that $\rho\|A_1^{-1}A_0\|=\frac{1}{2}$. Thus the parameter $h=\frac{1}{2}$ in Theorem 7, and we can conclude the existence and uniqueness of a solution of $F(X)=0$ in the sphere of radius $t^{**}=\rho^{-1}=2\|A_1^{-1}A_0\|$ and center at the origin.

5. A numerical example. From § 9.6 of [6] we get the following numerical example of a lightly damped system

$$F(\lambda)=I\lambda^2+B\lambda+C,$$

where

$$B=\begin{bmatrix} .00827963 & .00176811 & -.00327865 \\ .00176811 & .00741734 & .00067080 \\ -.00327865 & .00067080 & .05761640 \end{bmatrix}$$

and

$$C=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

We apply Theorem 6 to this equation. For this purpose we compute the following quantities

$$X_0=iC^{1/2}=i\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{bmatrix},$$

$$\gamma=\inf\sigma\left(\frac{C^{1/2}+(C^{1/2})^*}{2}\right)=1$$

and $\|B\|=.06156585$ using the maximum row sum norm. Then the hypothesis $\|B\|<2\gamma$ is certainly satisfied. We compute:

$$h=\frac{2\|BC^{1/2}\|}{(2\gamma-\|B\|)}=\frac{(2\sqrt{2})(.06156585)}{(2-.0615658)^2}=.04634285<\tfrac{1}{2}.$$

Thus our hypotheses are satisfied. Now

$$t^*=\frac{2\gamma-\|B\|}{2}(1-\sqrt{1-2h})=.046008288,$$

$$t^{**}=\frac{2\gamma-\|B\|}{2}(1+\sqrt{1-2h})=1.8924258,$$

and we have the result that

$$S = \{X \mid \|X - iC^{1/2}\|\} \leqq .046008288$$

contains one root that is unique in

$$T = \{X \mid \|X - iC^{1/2}\|\} < 1.8924258.$$

We compute this root using the Newton iteration method programmed on the CDC 6400 at the University of Calgary. With X_0 defined as above we compute $H_n = X_{n+1} - X_n$ from the equation

$$(X_n + B)H_n + H_n X_n = -(X_n^2 + BX_n + C), \qquad n = 0, 1, \cdots,$$

and get

$$\|H_1\| = .0304,$$
$$\|H_2\| = .000316,$$
$$\|H_3\| = .0000000335.$$

X_3 is now

$$\begin{bmatrix} -.41398150\text{E}-02 + i.99999012\text{E}+00, & -.88405493\text{E}-03 - i.32805807\text{E}-05, & .19206292\text{E}-02 + i.19356897\text{E}-04 \\ -.88405507\text{E}-03 + i.32805807\text{E}-05, & -.37086700\text{E}-02 + i.99999269\text{E}+00, & -.39295072\text{E}-03 - i.32162145\text{E}-05 \\ .13580209\text{E}-02 + i.19356899\text{E}-04, & -.27784928\text{E}-03 - i.32162148\text{E}-05, & -.28808200\text{E}-01 + i.14139188\text{E}+01 \end{bmatrix}$$

Let the eigenvalues of X_3 be e_1, e_2 and e_3. We compute these using routines from Eispack and get

$$e_1 = -.28808438\text{E}-01 + i.14139122\text{E}+01,$$
$$e_2 = -.48341075\text{E}-02 + i.99999065\text{E}+00,$$
$$e_3 = -.30141399\text{E}-02 + i.99999871\text{E}+00,$$

which compare favorably with numerical results claimed in [6].

REFERENCES

[1] J. E. DENNIS, J. F. TRAUB AND R. P. WEBER, *On the matrix polynomial, lambda-matrix and block eigenvalue problems*, Tech. Rep. 71-109, Dept. of Computer Science, Cornell Univ., Ithaca, N.Y.

[2] J. EISENFELD, *Operator equations and nonlinear eigenparameter problems*, J. Functional Analysis, 12 (1973), pp. 475–490.

[3] W. B. GRAGG AND R. A. TAPIA, *Optimal error bounds for the Newton–Kantorovich theorem*, SIAM J. Numer. Anal., 11 (1974), pp. 10–14.

[4] E. HEINZ, *Beiträge zur Störungstheorie der Spektralzerlegung*, Math. Ann., 123 (1951), pp. 415–438.

[5] M. G. KREIN AND H. LANGER, *On some mathematical principles of linear theory of damped vibrations of continua*, Proc. Int'l. Symp. in Applications of the Theory of Functions in Continuum Mechanics, Moscow, 1965, pp. 283–322.

[6] P. LANCASTER, *Lambda-Matrices and Vibrating Systems*, Pergamon Press, Oxford, 1966.

[7] H. LANGER, *On strongly damped bundles in Hilbert space*, J. Math. Mech., 17 (1968), pp. 685–706.

[8] M. I. MAVLYANOVA, *On a method for constructing the matrix solution for a polynomial matrix*, Automatic programming and numerical methods of anälysis, Steklov Seminar vol. 18, Consultants Bureau, New York 1972, pp. 71–79.

[9] J. M. ORTEGA, *The Newton–Kantorovich theorem*, Amer. Math. Monthly, 78 (1968), pp. 658–660.

[10] M. ROSENBLUM, *On the operator equation* $BX - XA = Q$, Duke Math. J., 23 (1956), pp. 263–270.

SIAM J. MATH. ANAL.
Vol. 8, No. 3, May 1977

DECOMPOSITIONS OF A HILBERT SPACE AND FACTORIZATION OF A W–A DETERMINANT*

R. J. HANGELBROEK† AND C. G. LEKKERKERKER‡

Abstract. The dispersion function $\Lambda(\lambda)$ which occurs in linear transport theory can be introduced as the W–A determinant of a certain pair of operators B_1, B_2 defined in $L^2[-1, 1]$. Each of the two operators is reduced by a complementary pair of subspaces of $L^2[-1, 1]$. In this paper the factorization $\Lambda(\lambda) = X(\lambda)X(-\lambda)$ is shown to correspond with a factorization of the operator $(VB_2V^{-1} - \lambda E) \cdot (B_1 - \lambda E)^{-1}$ into the product of two operators with determinants $X(\pm\lambda)$. Here V is an automorphism of $L^2[-1, 1]$ which is defined in terms of the projections associated with the two pairs of subspaces. The results are brought into a general setting.

1. Introduction. This paper deals with a question originating in linear transport theory. In a specific type of problems in that theory, viz., the so-called half-space problems, one has to deal with a situation which, within the framework of a Hilbert space, can be described as follows. There are given two bounded linear operators B_1 and B_2 which differ by a finite-dimensional operator. Each of the two operators is reduced by a complementary pair of closed subspaces, say, $\{H_+, H_-\}$ and $\{H_p, H_m\}$, respectively. Then the question is whether the projection P_+ onto H_+ along H_- induces an isomorphism from H_p onto H_+.

In transport theory (cf. [1], [2], [3]) the above question is currently answered in a constructive way. There, for an arbitrary $f_+ \in H_+$, an element $f \in H_p$ is constructed such that $P_+f = f_+$. The various methods to obtain f, as proposed so far, have in common that they use an a priori factorization of the W–A determinant associated with the operators B_1 and B_2, i.e., the determinant of $(B_2 - \lambda E) \cdot (B_1 - \lambda E)^{-1}$. ($E$ is the identity operator.) If this determinant is denoted by $\Lambda(\lambda)$, then the factorization takes the form $\Lambda(\lambda) = X(\lambda)X(-\lambda)$. The function $\Lambda(\lambda)$ is known as the dispersion function and its factorization is called a Wiener–Hopf factorization. One could hope that the factorization corresponds to some factorization of the operator $(B_2 - \lambda E)(B_1 - \lambda E)^{-1}$. This hope seems to be vain. In this paper, an operator V is introduced which is defined in terms of the projections associated with the pairs of subspaces $\{H_+, H_-\}$ and $\{H_p, H_m\}$. It is shown that H_+ and H_p are isomorphic under P_+ if V is an automorphism in H. The operator VB_2V^{-1} is reduced by $\{H_+, H_-\}$. Moreover, the operator $(VB_2V^{-1} - \lambda E) \cdot (B_1 - \lambda E)^{-1}$ does admit a factorization as a product of two operators having determinants $X(\lambda), X(-\lambda)$.

Since we believe that they may be of use in other fields than transport theory, we bring our results into an abstract setting in § 2. We shall make a generalization by considering two bounded operators B_1 and B_2 which differ by a nuclear operator. On the other hand, we shall restrict ourselves to a (separable) Hilbert space in spite of the fact that Hilbert space concepts do not enter directly into the problem as formulated above in the first paragraph. The reason for this restriction

* Received by the editors August 23, 1974, and in revised form May 15, 1975.

† Applied Mathematics Division, Argonne National Laboratory, Argonne, Illinois 60439. This author's work was performed under the auspices of the U.S. Energy Research and Development Administration.

‡ Institute of Mathematics, University of Amsterdam, Amsterdam, the Netherlands.

is that we want to make an unlimited use of [4] for reference. In § 3 a simple half-space problem of transport theory is discussed by way of example. The results of § 2 as applied in this example furnish a new approach to the half-space problems. This approach shows that questions of existence and uniqueness of a solution to that type of problem are not directly linked to the possibility of a Wiener–Hopf factorization of the dispersion function. On the other hand, the example indicates clearly why such a factorization is needed if one wants to determine a solution explicitly. The approach in § 3 can be seen as complementary to the theory developed in [2], in particular, the part of that publication concerned with half-space problems.

Notation. In this paper the identity operator is indicated by E. For A a linear operator $\rho(A)$ will denote its resolvent set and $A|G$ its restriction to a subspace G.

2. Three theorems. Let H be a Hilbert space with closed linear subspaces H_+, H_-, H_p, H_m such that we have the direct sum decompositions

$$H = H_+ \oplus H_- = H_p \oplus H_m.$$

We do not assume any orthogonality properties of these subspaces. Let $P_\pm$ denote the linear projection operators in H that project onto $H_\pm$ along $H_\mp$, respectively. Similarly, P_p and P_m are the linear projections onto H_p, H_m along H_m, H_p, respectively.

We define a bounded linear operator V in H by

$$V = P_+P_p + P_-P_m. \tag{2.1}$$

From the definition it will be clear that V restricted to H_p, H_m coincides with P_+, P_-, respectively. In other words,

$$V|H_p = P_+|H_p, \qquad V|H_m = P_-|H_m.$$

With this remark the first theorem becomes almost trivial.

THEOREM 1. *If the operator V is an automorphism in H, then $V^{-1} = P + Q$, where $P = V^{-1}P_+$ is a bounded projection onto H_p along H_- and $Q = V^{-1}P_-$ is a bounded projection onto H_m along H_+. The restrictions $P|H_+$ and $P_+|H_p$ are inverse to each other. The same is true for the restrictions $Q|H_-$ and $P_-|H_m$.*

Proof. From the definition of the operator V it is clear that $VH_p \subset H_+$. Similarly, $VH_m \subset H_-$. Since, by our hypothesis, V is a bijection from H onto itself, it follows that $VH_p = H_+$. Thus $V|H_p$ is a bijection from H_p onto H_+.

We put $P = V^{-1}P_+$ and $Q = V^{-1}P_-$. Then $V^{-1} = P + Q$. The restrictions of V^{-1} and P to the subspace H_+ coincide. The same is true for the restrictions of V and P_+ to the subspace H_p. Therefore, $P|H_+$ and $P_+|H_p$ are inverse to each other. Moreover, we have the relations

$$PP_+ = P \quad \text{and} \quad P_+P = P_+(P_pP) = (P_+P_p)P = VP = P_+.$$

Finally, we show that P is a projection by using a familiar argument: $P^2 = (PP_+)P = P(P_+P) = PP_+ = P$. The assertions with respect to Q follow in a similar way.

In addition to the definitions and assumptions preceding Theorem 1, we assume that in H two bounded linear operators B_1 and B_2 exist such that H_+, H_- are invariant subspaces under B_1, and H_p, H_m are invariant under B_2.

First we collect some simple properties of the projections P_+, P_-, P, Q, P_p and P_m:

$$P_+P = P_+, \quad PP_+ = P, \quad P_-Q = P_-, \quad QP_- = Q, \tag{2.2}$$

$$P_pP = P, \quad P_pQ = 0, \quad P_mP = 0, \quad P_mQ = Q, \tag{2.3}$$

$$PB_2P = B_2P, \qquad QB_2Q = B_2Q, \tag{2.4}$$

$$P_+B_1P = P_+B_1, \qquad P_-B_1Q = P_-B_1. \tag{2.5}$$

The first two relations of (2.2) were already obtained in proving Theorem 1. The latter pair is of the same type as the first pair; they follow by replacing P_+, P by P_-, Q, respectively. The relations in (2.3) are obvious since P and Q have ranges H_p and H_m, respectively. The relations in (2.4) follow from the assumption that H_p and H_m are invariant under B_2. The projections P_+ and P_- commute with B_1 as a consequence of the assumption that B_1 is reduced by the complementary pair $\{H_+, H_-\}$. The relations (2.5) then follow from (2.2).

THEOREM 2. *Let the operator V defined by* (2.1) *be an automorphism in H and let B_1, B_2 be two bounded linear operators in H which are reduced by the complementary pairs of subspaces $\{H_+, H_-\}$ and $\{H_p, H_m\}$, respectively. Then the operator VB_2V^{-1} is reduced by $\{H_+, H_-\}$. For each $\lambda \in \rho(B_1)$, $\lambda \neq 0$, the operator $\Omega(\lambda) = (VB_2V^{-1} - \lambda E)(B_1 - \lambda E)^{-1}$ can be factorized as a product $(E + R_+(\lambda)) \cdot (E + R_-(\lambda))$, where $R_\pm(\lambda)$ are bounded linear operators which vanish on $H_\mp$. Furthermore,*

$$\begin{aligned} E + R_+(\lambda) &= (P_+B_2P - \lambda E)(P_+B_1 - \lambda E)^{-1}, \\ E + R_-(\lambda) &= (P_-B_2Q - \lambda E)(P_-B_1 - \lambda E)^{-1}. \end{aligned} \tag{2.6}$$

If $B_2 - B_1$ is a nuclear operator, then $R_\pm(\lambda)$ are nuclear.

Proof. Let us investigate the operator VB_2V^{-1}. According to Theorem 1 and (2.4),

$$B_2V^{-1} = B_2P + B_2Q = PB_2P + QB_2Q.$$

Hence by (2.2) and (2.3),

$$\begin{aligned} VB_2V^{-1} &= (P_+P_p + P_-P_m)(PB_2P + QB_2Q) \\ &= P_+PB_2P + P_-QB_2Q = P_+B_2P + P_-B_2Q. \end{aligned}$$

Substituting $B_2 = B_1 + D$ and using (2.5), we obtain

$$VB_2V^{-1} = B_1 + P_+DP + P_-DQ. \tag{2.7}$$

The penultimate formula also furnishes a decomposition of VB_2V^{-1} as the sum of two operators leaving invariant H_+, H_- and annihilating H_-, H_+, respectively. In other words, VB_2V^{-1} is reduced by the pair $\{H_+, H_-\}$. Clearly, B_1 admits a similar decomposition, viz., $B_1 = P_+B_1 + P_-B_1$. These two facts enable us to factorize $\Omega(\lambda)$.

Let $\lambda \in \rho(B_1)$, $\lambda \neq 0$. Since $P_+P_- = P_-P_+ = 0$ and P_+, P_- commute with B_1, we have

$$E - \lambda^{-1}B_1 = (E - \lambda^{-1}P_+B_1)(E - \lambda^{-1}P_-B_1),$$
$$P_+(E - \lambda^{-1}P_-B_1)^{-1} = (E - \lambda^{-1}P_-B_1)^{-1}P_+ = P_+.$$

The analogous relations obtained by interchanging P_+, P_- are also true. Using (2.2) we get

$$P(E - \lambda^{-1}P_-B_1)^{-1} = P, \qquad Q(E - \lambda^{-1}P_+B_1)^{-1} = Q.$$

So for $\Omega(\lambda)$ we find, taking into account (2.7),

$$\begin{aligned}\Omega(\lambda) &= (E - \lambda^{-1}VB_2V^{-1})(E - \lambda^{-1}B_1)^{-1} \\ &= E - \lambda^{-1}(P_+DP + P_-DQ)(E - \lambda^{-1}B_1)^{-1} \\ &= E - \lambda^{-1}P_+DP(E - \lambda^{-1}P_+B_1)^{-1} - \lambda^{-1}P_-DQ(E - \lambda^{-1}P_-B_1)^{-1} \\ &= (E - \lambda^{-1}P_+DP(E - \lambda^{-1}P_+B_1)^{-1})(E - \lambda^{-1}P_-DQ(E - \lambda^{-1}P_-B_1)^{-1}).\end{aligned}$$

Thus our result is that for $\lambda \in \rho(B_1)$, $\lambda \neq 0$, $\Omega(\lambda)$ can be written as

$$\Omega(\lambda) = E + R(\lambda) = (E + R_+(\lambda))(E + R_-(\lambda)),$$

where $R(\lambda)$, $R_+(\lambda)$ and $R_-(\lambda)$ are given by

$$R(\lambda) = (P_+DP + P_-DQ)(B_1 - \lambda E)^{-1},$$

$$R_+(\lambda) = P_+DP(P_+B_1 - \lambda E)^{-1}, \qquad R_-(\lambda) = P_-DQ(P_-B_1 - \lambda E)^{-1}. \tag{2.8}$$

Also, the relations (2.6) hold.

The operators $R_\pm(\lambda)$ determined by (2.8) satisfy

$$R_+(\lambda) = R_+(\lambda)P_+ = R(\lambda)P_+, \qquad R_-(\lambda) = R_-(\lambda)P_- = R(\lambda)P_-,$$

since $P = PP_+$, $Q = QP_-$ and $P_\pm$ commute with $(P_\pm B_1 - \lambda E)^{-1}$. Consequently $R_\pm(\lambda)$ vanish on $H_\mp$, so that $E + R_\pm(\lambda)$ coincide with $\Omega(\lambda)$ on the subspaces $H_\pm$.

For $D = B_2 - B_1$ a nuclear operator, the last assertion of the theorem follows from (2.8).

In Theorem 3 we assume that $B_2 - B_1$ is nuclear. We also have to be more specific about the Hilbert space H and the operator V: H is assumed to be separable and $E - V$ is required to be a Hilbert–Schmidt operator.

First we bring some definitions and results about determinants from [4], which we need in Theorem 3. As is known, the determinant $\det(E - K)$ and the (spectral) trace $\operatorname{tr} K$ can be defined for any nuclear operator in a separable Hilbert space by

$$\det(E - K) = \prod_{j=1}^{\nu(K)} (1 - \lambda_j(K)) \tag{2.9}$$

and

$$\operatorname{tr} K = \sum_{j=1}^{\nu(K)} \lambda_j(K),$$

where $\lambda_j(K)$ are the nonzero eigenvalues of K and $\nu(K)$ is the sum of the (algebraic) multiplicities of the $\lambda_j(K)$. For $\nu(K)=\infty$, the convergence of the sum and the product are a direct consequence of K being nuclear. In fact, we have the inequality

$$\sum_{j=1}^{\nu(K)} |\lambda_j(K)| \leqq \sum_{j=1}^{r(K)} s_j(K) \tag{2.10}$$

with $s_j(K)=\lambda_j((K^*K)^{1/2})$; $r(K)$ is the rank of K. The right-hand member of the inequality converges by definition if K is a nuclear operator.

If K is finite-dimensional (i.e., if $r(K)<\infty$), then $\det(E-K)$ can be defined equivalently as the determinant of the restriction of $E-K$ to any finite-dimensional subspace which contains the range of the operator K.

For two bounded linear operators A_1, A_2 that differ by a nuclear operator, the W–A determinant (Weinstein–Aronszajn or perturbation determinant) is defined by

$$\det\{(A_2-\lambda E)(A_1-\lambda E)^{-1}\}=\det\{E+(A_2-A_1)(A_1-\lambda E)^{-1}\}, \qquad \lambda\in\rho(A_1).$$

In particular, if $0\in\rho(A_1)$, then $A_2A_1^{-1}=E+(A_2-A_1)A_1^{-1}$ and the determinant $\det\{A_2A_1^{-1}\}$ exists.

In this section we use the following properties of the determinant and the trace,

$$\det\{B(E-K)B^{-1}\}=\det(E-K), \tag{2.11}$$

$$\det\{(E-K)(E-K_1)\}=\det(E-K)\det(E-K_1), \tag{2.12}$$

$$\operatorname{tr}(K+K_1)=\operatorname{tr}K+\operatorname{tr}K_1,$$

where B is a bounded automorphism and K, K_1 are nuclear operators in H. The relations (2.11) and (2.12) are the results 6 and 7 in [4, Chap. IV, § 1]. The last relation is a consequence of the fact that the spectral trace is equal to the matrix trace [4, Chap. III, Thm. 8.4].

The perturbation determinant of two bounded linear operators A_1, A_2 differing by a nuclear operator K is a holomorphic function on the resolvent set $\rho(A_1)$ [4, Chap. IV, § 1 (result 8) and § 3].

For K a Hilbert–Schmidt operator, a regularized determinant $\tilde{\det}(E-K)$ can be introduced by

$$\tilde{\det}(E-K)=\prod_{j=1}^{\nu(K)}[(1-\lambda_j(K))\exp\lambda_j(K)]. \tag{2.13}$$

If $\nu(K)=\infty$, then the product converges since

$$\sum_{j=1}^{\nu(K)} |\lambda_j(K)|^2 \leqq \sum_{j=1}^{r(K)} (s_j(K))^2$$

and the right-hand member converges by definition.

For A_1, A_2 bounded linear operators, $0\in\rho(A_1)$ and A_2-A_1 a Hilbert–

Schmidt operator, we have

$$\det(A_2A_1^{-1}) = \det(E+(A_2-A_1^{-1}).$$

Finally, we need the property [4, IV. 3]: If K_1, K_2 are Hilbert–Schmidt, K_2-K_1 is nuclear and $0\in\rho(E-K_1)$, then

$$\det\{(E-K_2)(E-K_1)^{-1}\} = \frac{\det(E-K_2)}{\det(E-K_1)}\exp[\operatorname{tr}(K_1-K_2)]. \tag{2.14}$$

THEOREM 3. *Let B_1, B_2 be bounded linear operators in the separable Hilbert space H such that B_1 and B_2 differ by a nuclear operator and are reduced by the complementary pairs of closed subspaces $\{H_+, H_-\}$ and $\{H_p, H_m\}$, respectively. Let the operator V, defined by* (2.1) *be an automorphism in H such that $E-V$ is a Hilbert–Schmidt operator. Then we have the following relations between* W–A *determinants*:

$$\begin{aligned}\det\{(VB_2V^{-1}-\lambda E)(B_1-\lambda E)^{-1}\} &= \det\{(B_2-\lambda E)(B_1-\lambda E)^{-1}\}\\ &= \det\{(P_+B_2P-\lambda E)(P_+B_1-\lambda E)^{-1}\}\det\{(P_-B_2Q-\lambda E)\\ &\qquad\cdot(P_-B_1-\lambda E)^{-1}\}\end{aligned} \tag{2.15}$$

for $\lambda\in\rho(B_1)$, $\lambda\neq 0$.

If $0\in\rho(B_1)$, then the point $\lambda=0$ is a removable singularity of each of the two factors in the right-hand side of (2.15).

Proof. Since $B_2 = B_1+D$, we can write relation (2.7) in the form

$$VB_1V^{-1}-B_1 = -VDV^{-1}+P_+DP+P_-DQ.$$

We denote the right-hand member, which is a nuclear operator, by K and deduce

$$V(B_1-\lambda E)V^{-1}-(B_1-\lambda E) = K.$$

After a multiplication by V^{-1} from the left and by $(B_1-\lambda E)^{-1}V$ from the right, the latter relation yields

$$(B_1-\lambda E)V^{-1}(B_1-\lambda E)^{-1}V-E = V^{-1}K(B_1-\lambda E)^{-1}V, \quad \lambda\in\rho(B_1). \tag{2.16}$$

The right-hand member of this formula is again a nuclear operator. So we may write (using (2.11) in the first step)

$$\begin{aligned}\det\{(VB_2V^{-1}-\lambda E)(B_1-\lambda E)^{-1}\} &= \det\{(B_2-\lambda E)V^{-1}(B_1-\lambda E)^{-1}V\}\\ &= \det\{(B_2-\lambda E)(B_1-\lambda E)^{-1}\}\det\{(B_1-\lambda E)V^{-1}(B_1-\lambda E)^{-1}V\}.\end{aligned} \tag{2.17}$$

In order to prove the first equality in (2.15), we have to show that the second determinant in the last member of (2.17) equals 1.

The operator V^{-1} can be written as

$$V^{-1} = (P_+P+P_-Q)+P_-P+P_+Q = E+H,$$

where $H = P_-P+P_+Q$. Substituting this for V^{-1} in the relation $(B_1-\lambda E)V^{-1}(B_1-\lambda E)^{-1}-V^{-1} = V^{-1}K(B_1-\lambda E)^{-1}$ (cf. (2.16)), we obtain

$$(B_1-\lambda E)H(B_1-\lambda E)^{-1}-H = V^{-1}K(B_1-\lambda E)^{-1}, \qquad \lambda\in\rho(B_1). \tag{2.18}$$

The assumption that $E-V$ is Hilbert–Schmidt implies that $H=V^{-1}(E-V)$ also is Hilbert–Schmidt. This means that we can apply formula (2.14) for $K_1=-H$, $K_2=-(B_1-\lambda E)H(B_1-\lambda E)^{-1}$. We obtain

$$\det\{(B_1-\lambda E)V^{-1}(B_1-\lambda E)^{-1}V\}$$
$$=\frac{\tilde{\det}\{(B_1-\lambda E)V^{-1}(B_1-\lambda E)^{-1}\}}{\tilde{\det}\,V^{-1}}\exp[\operatorname{tr}\{-H+(B_1-\lambda E)H(B_1-\lambda E)^{-1}\}].$$

From the definition of a regularized determinant as given in (2.13), it will be clear that

$$\tilde{\det}\{(B_1-\lambda E)V^{-1}(B_1-\lambda E)^{-1}\}=\tilde{\det}\,V^{-1}.$$

It remains to show that the exponential function in (2.19) equals 1. We apply the projection P_- to both members of (2.18), bearing in mind that the operators B_1 and P_- commute and that $P_-H=P_-P$ since $P_-P_+=0$. We obtain

$$(B_1-\lambda E)P_-P(B_1-\lambda E)^{-1}-P_-P=P_-V^{-1}K(B_1-\lambda E)^{-1},$$

where the right-hand member is nuclear. The square power of the left-hand member is 0 since $PP_-=0$. That means that the left-hand member is nilpotent and hence a nuclear Volterra operator. The trace of such an operator equals 0, i.e.,

$$\operatorname{tr}\{(B_1-\lambda E)P_-P(B_1-\lambda E)^{-1}-P_-P\}=0.$$

Similarly, we have

$$\operatorname{tr}\{(B_1-\lambda E)P_+Q(B_1-\lambda E)^{-1}-P_+Q\}=0.$$

Addition of the last two formulas yields

$$\operatorname{tr}\{(B_1-\lambda E)H(B_1-\lambda E)^{-1}-H\}=0,$$

which completes the proof of the first equality in (2.15). The second equality in (2.15) is a direct consequence of Theorem 2 and (2.12).

From the proof of Theorem 2, we know that

$$(E+R_+(\lambda))|H_-=E|H_-,\qquad (E+R_+(\lambda))|H_+=\Omega(\lambda)|H_+,$$

for $\lambda\in\rho(B_1)$, $\lambda\neq 0$. Similar relations hold for $E+R_-(\lambda)$. If one uses the definition of the determinant as given in (2.9), it will be clear that

$$\det(E+R_+(\lambda))=\det\{\Omega(\lambda)|H_+\},\qquad \lambda\in\rho(B_1),\quad \lambda\neq 0.$$

If $0\in\rho(B_1)$, then the right-hand member in this relation is holomorphic in a neighborhood of $\lambda=0$. Hence $\det(E+R_+(\lambda))$ can be extended to $\lambda=0$ by analytic continuation, which proves the last assertion of the theorem.

3. Application. In this section we apply the results of § 2 to a half-space problem from neutron transport theory. The problem was studied a.o. by Case and Zweifel [1]. A functional analytic approach was given by the first author of this paper [2] and by Larsen and Habetler [3]. The problem requires the solution

of the equation

$$\mu\frac{\partial\psi}{\partial x}(x,\mu)+\psi(x,\mu)=\frac{c}{2}\int_{-1}^{1}\psi(x,\mu')\,d\mu', \qquad 0<x<\infty, \quad -1\leqq\mu\leqq+1, \tag{3.1}$$

under the conditions

$$\psi(0,\mu)=f_+(\mu), \qquad 0<\mu\leqq+1, \tag{3.2}$$

and

$$\lim_{x\to\infty}\psi(x,\mu)=0, \qquad -1\leqq\mu\leqq+1.$$

The parameter c in (3.1) is a constant strictly between 0 and 1. The condition at $x=0$ has the peculiarity that the given function f_+ is defined only for values of $\mu>0$.

Equation (3.1) and its conditions can be written in a more concise form as follows. Consider the Hilbert space $H=L^2(I)$ ($I=\{-1\leqq\mu\leqq+1\}$) with the inner product

$$(f,g)=\int_{-1}^{+1}f(\mu)\overline{g(\mu)}\,d\mu.$$

Denote by e the function on I that is identically equal to 1. Next, define two bounded linear operators $T=T_I$ and A in H by putting

$$(Tf)(\mu)=\mu f(\mu),$$

$$(Af)(\mu)=f(\mu)-\frac{c}{2}\int_{-1}^{+1}f(\mu')\,d\mu', \quad f\in H, \quad \mu\in I. \tag{3.3}$$

In other words, T is the operator of multiplication by the independent variable μ, and A is obtained from the identity operator E in H by adding a certain one-dimensional perturbation. The relation (3.3) may be written as

$$Af=f-\frac{c}{2}(f,e)e.$$

It is easily seen that A can be inverted and that

$$A^{-1}g=g+\gamma(g,e)e, \qquad g\in H, \quad \gamma=\frac{c}{2(1-c)}. \tag{3.4}$$

We use the following notation:

$$I_+=[0,1], \quad I_-=[-1,0], \quad H_+=L^2(I_+), \quad H_-=L^2(I_-).$$

H_+ and H_- can be considered as closed subspaces of H. They are invariant under the operator T. We denote the projections onto $H_\pm$ along $H_\mp$ by $P_\pm$.

Applying A^{-1} and interpreting ψ as a function of x with values in $L^2(I)$, we can write (3.1) as a first order linear differential equation:

$$A^{-1}T\frac{d\psi}{dx}(x)=-\psi(x). \tag{3.5}$$

The conditions at $x = 0$ and for $x \to \infty$ can be written as

$$P_+\psi(0) = f_+, \qquad f_+ \in H_+, \tag{3.6}$$

and

$$\lim_{x\to\infty} \psi(x) = 0, \tag{3.6'}$$

where the limit has to be taken in H.

We indicate some of the basic properties of the operator $A^{-1}T$. Full proofs are given in [2]. $A^{-1}T$ is bounded and is obtained from T by adding a certain perturbation. In fact,

$$A^{-1}T = T + B_0T, \tag{3.7}$$

where B_0 is the one-dimensional operator given by

$$B_0f = \gamma(f, e)e, \qquad f \in H. \tag{3.8}$$

The W–A determinant of the pair of operators T, $A^{-1}T$ is denoted by $\Lambda(\lambda)$. Since B_0T is one-dimensional, this determinant is obtained as the eigenvalue of the operator $(A^{-1}T - \lambda E)(T - \lambda E)^{-1}$, which corresponds to the eigenvector e. Thus

$$\Lambda(\lambda) = \det\{(A^{-1}T - \lambda E)(T - \lambda E)^{-1}\} = 1 + \gamma \int_{-1}^{+1} \frac{\mu}{\mu - \lambda} d\mu, \qquad \lambda \notin I. \tag{3.9}$$

The function $\Lambda(\lambda)$ is holomorphic outside the segment I. It has two simple zeros $\pm \nu_0$, where $\nu_0 > 0$. Moreover, the spectrum N of the operator $A^{-1}T$ just consists of I and the two points $\pm \nu_0$, which are eigenvalues of multiplicity 1.

The operator $A^{-1}T$ is self-adjoint if H is endowed with the inner product (equivalent to the originally given inner product $(\cdot, \cdot)$)

$$(f, g)_A = (Af, g), \qquad f, g \in H.$$

Another property is that the function e is a cyclic element for the operator $A^{-1}T$. This can be deduced from the formulas (3.7) and (3.8). It follows then from the spectral theorem that there exists a unitary isomorphism F from the space H endowed with the inner product $(\cdot, \cdot)_A$ onto the Hilbert space $L^2(N, \sigma)$, σ a finite positive Borel measure on N, such that F diagonalizes the operator $A^{-1}T$. By this we mean that

$$A^{-1}T = F^{-1}T_NF,$$

where T_N is the operator of multiplication by the independent variable $\nu \in N$. The transformation F is so chosen that $Fe = \tilde{e}$, where $\tilde{e}$ is the function on N that is identically equal to 1. The inner product in $L^2(N, \sigma)$ is defined by

$$(\tilde{f}, \tilde{g}) = \int_N \tilde{f}(\nu)\overline{\tilde{g}(\nu)}\, d\sigma(\nu).$$

The isomorphism F and the measure σ can be determined explicitly. For functions f, $\tilde{f}$ that (with the exception of possibly a finite jump at the point 0) are

Hölder continuous on I, N, respectively, F and F^{-1} are given by

$$(Ff)(\nu)=\begin{cases}-\gamma\int_{-1}^{+1}\dfrac{f(\mu)-f(\nu)}{\mu-\nu}\mu\,d\mu+f(\nu), & -1\leqq\nu\leqq+1,\\ -\gamma\int_{-1}^{+1}\dfrac{f(\mu)}{\mu-\nu}\mu\,d\mu, & \nu=\pm\nu_0,\end{cases}$$

$$(3.10)\quad (F^{-1}\tilde{f})(\mu)=\gamma\int_N\frac{\tilde{f}(\nu)-\tilde{f}(\mu)}{\nu-\mu}\nu\,d\sigma(\nu)+\tilde{f}(\mu),\qquad \mu\in I.$$

The measure σ is absolutely continuous on $-1\leqq\nu\leqq+1$; one has

$$d\sigma(\nu)=\frac{d\nu}{\Lambda^{+}\Lambda^{-}(\nu)},\qquad -1\leqq\nu\leqq+1,\quad \sigma\{\pm\nu_0\}=\rho(\nu_0),$$

where $\Lambda^{\pm}$ denote the boundary values of Λ, i.e., $\Lambda^{\pm}(\nu)=\lim_{\eta\downarrow 0}\Lambda(\nu\pm i\eta)$, and where $\gamma\nu_0\rho(\nu_0)$ is the residue of $1/\Lambda(\lambda)$ at the point ν_0.

Next, we consider half-ranges. We decompose N into two parts by putting

$$N_{+}=[0,1]\cup\{\nu_0\},\qquad N_{-}=[-1,0]\cup\{-\nu_0\}.$$

The spaces $L^2(N_{\pm},\sigma)$ may be looked upon as subspaces of $L^2(N,\sigma)$. We denote the projections onto $L^2(N_{\pm},\sigma)$ along $L^2(N_{\mp},\sigma)$ by $\tilde{P}_{\pm}$ and the inverse images $F^{-1}L^2(N_{\pm},\sigma)$ by H_p, H_m. It is clear that H_p, H_m are closed subspaces of H which are orthogonal relative to the inner product $(\cdot,\cdot)_A$ and that we have the direct sum decomposition

$$H=H_p\oplus H_m.$$

The two summands in the last decomposition are invariant under $A^{-1}T$ and the restrictions of $A^{-1}T$ to these two subspaces have spectra N_{+}, N_{-}, respectively. We thus have an example of a dissection of (a connected part of) the spectrum N, with a corresponding decomposition of the space H. The projection operators associated with this decomposition are denoted by P_p and P_m, so

$$(3.11)\qquad P_p=F^{-1}\tilde{P}_{+}F,\qquad P_m=F^{-1}\tilde{P}_{-}F.$$

The transformation F makes it possible to study (3.5) in a simpler form and to obtain explicit solutions. The equation is transformed by F into

$$(3.12)\qquad T_N\frac{d\tilde{\psi}}{dx}(x)=-\tilde{\psi}(x),\qquad \tilde{\psi}(x)\in L^2(N,\sigma),$$

where $\tilde{\psi}(x)=F\psi(x)$ for all $0<x<\infty$. In [2], it has been shown that $\tilde{\psi}(0)\in L^2(N_{+},\sigma)$ is a necessary and sufficient condition for (3.12) to have a solution satisfying $\lim_{x\to\infty}\tilde{\psi}(x)=0$. Moreover, for a given $\tilde{\psi}(0)$, the solution is unique and $\tilde{\psi}(x)\in L^2(N_{+},\sigma)$ for all $0\leqq x<\infty$. These facts imply the truth of the following assertion: *equation* (3.5) *has a unique solution* $\psi(x)$ *such that* (3.6) *and* (3.6′) *hold if and only if* f_{+} *is the* H_{+}*-component of a unique element* h *of* H_p. *The element* h *coincides with the initial value* $\psi(0)$.

This result induces us to consider the following questions about the restriction of P_{+} to H_p: (i) is $P_{+}|H_p$ an injective map into H_{+}; (ii) what is the range of

$P_+|H_p$? It will be clear that, once the answers to these questions have been found, the analogous questions for $P_-|H_m$ can be answered by a symmetry argument. (Such symmetry arguments will also be used later on. They can be made precise by the use of a transformation S defined by $(Sf)(\mu)=f(-\mu)$, $\mu \in I$. The transformation S is unitary with respect to $(\cdot\,,\cdot)$ as well as $(\cdot\,,\cdot)_A$; P_+, H_+ and H_p are transformed by S into P_-, H_- and H_m, respectively. Moreover, $STS=-T$ and $SA^{-1}TS=-A^{-1}T$.)

It is not difficult to show that $P_+|H_p$ is injective. The proof will be given as a part of the proof of the following lemma.

LEMMA 1. *The operator $V=P_+P_p+P_-P_m$ is an automorphism in H. The operator $E-V$ is Hilbert–Schmidt.*

Proof. Let $g\neq 0$ be an arbitrary element of H_p. Then $(A^{-1}Tg, g)_A = \int_{N_+} \nu|(Fg)(\nu)|^2\, d\sigma(\nu)>0$. Hence, $(Tg, g)=\int_I \mu|g(\mu)|^2\, d\mu>0$ and so $P_+g\neq 0$. This means that $P_+|H_p$ and, by symmetry, $P_-|H_m$ are injective. Next, take an arbitrary $f\neq 0$ of H. f can be decomposed as $f=f_p+f_m$ with $f_p\in H_p$, $f_m\in H_m$, and at least one of the components has to be different from 0. Then at least one of the summands in $VF=P_+f_p+P_-f_m$, hence also Vf itself, has to differ from 0. So V is injective.

In order to show that V is surjective, we write $V=E-P_-P_p-P_+P_m$. Since V is injective, it suffices to prove that $P_-P_p+P_+P_m$ is compact. We consider $P_-P_p=P_-F^{-1}\tilde{P}_+F$. For Hölder continuous functions $\tilde{f}$ defined on N_+, we obtain from (3.10)

$$(P_-F^{-1}\tilde{f})(\mu)=\gamma\int_{N_+}\frac{\tilde{f}(\nu)}{\nu-\mu}\nu\, d\sigma(\nu), \qquad \mu\in I_-. \tag{3.13}$$

The kernel of the integral operator in the right-hand member is bounded and hence square integrable on $I_-\times N_+$ with respect to the product measure $d\mu\times d\sigma(\nu)$. Therefore the integral operator is a Hilbert–Schmidt map of $L^2(N_+,\sigma)$ into H_-. Since the Hölder continuous functions on N_+ form a dense subset of $L^2(N_+,\sigma)$ and since P_-F^{-1} is continuous on $L^2(N_+,\sigma)$, it follows that relation (3.13) holds for all $\tilde{f}\in L^2(N_+,\sigma)$. Then P_-P_p and, by symmetry, P_+P_m are Hilbert–Schmidt. The same is true for their sum $P_-P_p+P_+P_m=E-V$.

Taking $B_1=T$, $B_2=A^{-1}T$ and using the results of Lemma 1, we are now in the position to apply the theorems in § 2. From Theorem 1 we know that the projection P onto H_p along H_- exists. Thus we arrive at the following result (obtained in [2] by explicit calculations).

THEOREM 4. *The boundary value problem* (3.5)–(3.6′) *can be solved uniquely for each $f_+\in H_+$.*

More precisely, the conditions (3.6) and (3.6′) are equivalent with $\psi(0)=Pf_+$. With the latter condition at $x=0$, the original boundary value problem becomes a regular initial value problem in the space H_p.

The function $\Lambda(\lambda)$ was defined as the W–A determinant of the pair of operators T, $A^{-1}T$. Applying Theorems 2 and 3 of § 2, we obtain

THEOREM 5. *We have*

$$\Lambda(\lambda)=\det(E+R_+(\lambda))\det(E+R_-(\lambda)), \qquad \lambda\notin I, \tag{3.14}$$

where

$$
\begin{aligned}
E+R_+(\lambda) &= (P_+A^{-1}TP-\lambda E)(P_+T-\lambda E)^{-1} \\
&= E+P_+B_0TP(P_+T-\lambda E)^{-1}, \\
E+R_-(\lambda) &= (P_-A^{-1}TQ-\lambda E)(P_-T-\lambda E)^{-1} = E+P_-B_0TQ(P_-T-\lambda E)^{-1}.
\end{aligned} \tag{3.15}
$$

Also,

$$(E+R_+(\lambda))(E+R_-(\lambda)) = (VA^{-1}TV^{-1}-\lambda E)(T-\lambda E)^{-1}.$$

The determinants in the right-hand member of (3.14) are W–A determinants of the pairs of operators P_+T, $P_+A^{-1}TP$ and P_-T, $P_-A^{-1}TQ$, respectively. The spectra of P_+T and P_-T are given by the sets $[0, 1]$ and $[-1, 0]$, respectively. Hence, the functions $Y_+(\lambda) = \det(E+R_+(\lambda))$, $Y_-(\lambda) = \det(E+R_-(\lambda))$ are holomorphic outside the intervals $[0, 1]$, $[-1, 0]$, respectively. For reasons of symmetry, we have $Y_-(\lambda) = Y_+(-\lambda)$, so that

$$\Lambda(\lambda) = Y_+(\lambda)Y_+(-\lambda), \qquad \lambda \notin I.$$

We remark that the operator $R_+(\lambda)$ is one-dimensional with range span (P_+e), the linear span of P_+e. Adopting the notation $e_+ = P_+e$ (i.e., e_+ is the characteristic function of the interval I_+), we see that $\det(E+R_+(\lambda))$ equals the determinant of the restriction of $E+R_+(\lambda)$ to the one-dimensional space span (e_+). In other words, $Y_+(\lambda)$ is the eigenvalue of the operator $E+R_+(\lambda)$ which corresponds to the eigenvector e_+:

$$(E+R_+(\lambda))e_+ = Y_+(\lambda)e_+, \qquad \lambda \notin I_+. \tag{3.16}$$

Moreover, $Y_+(\lambda)$ is the only eigenvalue of $E+R_+(\lambda)$ which may be different from 1.

The actual factorization of the dispersion function $\Lambda(\lambda)$ into a product of the form

$$\Lambda(\lambda) = X(\lambda)X(-\lambda) \tag{3.17}$$

plays a decisive role in the explicit determination of the projection operator P (see the end of this section). In order to guarantee the uniqueness of the latter factorization, certain conditions have to be imposed on $X(\lambda)$. In [2, 7.3] such conditions were formulated in the following way: $X(\lambda)$ has to be a function defined for $\lambda \notin (0, 1]$ with the properties (a) $X(\lambda)$ is holomorphic in the open half-space $\operatorname{Re}\lambda < 0$, (b) $X(\lambda)$ is continuous and without zeros in the closed half-space $\operatorname{Re}\lambda \leqq 0$, (c) $\lim_{\lambda\to\infty} X(\lambda) = 1$. We remark that the condition $X(0) \neq 0$ contained in (b) is superfluous since it is a consequence of $\lim_{\lambda\to 0}\Lambda(\lambda) = 1/(1-c) \neq 0$ (cf. (3.9)). In Lemma 2 we give three properties of the function $Y_+(\lambda)$ which imply that $Y_+(\lambda)$, extended with its limit value at $\lambda = 0$, satisfies the conditions (a), (b) and (c). Thus we conclude that

$$X(\lambda) = \det(E+R_+(\lambda)),$$

so that the actual factorization of the dispersion function $\Lambda(\lambda)$ may be looked upon as a means to calculate the W–A determinant of the pair P_+T, $P_+A^{-1}TP$.

LEMMA 2. *The function* $Y_+(\lambda) = \det(E + R_+(\lambda))$ *has the properties*:
(i) $Y_+(\lambda)$ *is holomorphic outside* $[0, 1]$ *and has* $\lambda = \nu_0$ *as its only zero,*
(ii) $\lim_{\lambda \to \infty} Y_+(\lambda) = 1$,
(iii) $Y_+(\lambda)$ *has a limit for* $\lambda \to 0$, $\frac{1}{2}\pi \leqq \arg \lambda \leqq \frac{3}{2}\pi$.

Proof. The operator $A^{-1}T|H_p$ has spectrum $N_+ = [0, 1] \cup \{\nu_0\}$. Since P_+ is an isomorphism from H_p onto H_+ with P as its partial inverse, it follows that $P_+A^{-1}TP$ also has N_+ as its spectrum, with ν_0 an eigenvalue. From the first equality in (3.15) we deduce that for $\lambda \notin I$ the operator $E + R_+(\lambda)$ has an eigenvalue 0 if and only if $\lambda = \nu_0$. According to the remark following (3.16) this means that $Y_+(\lambda)$ has ν_0 as its only zero.

From (3.15) and (3.16) it will be clear that the map $\lambda \to Y_+(\lambda)$ is regular at infinity with $\lim_{\lambda \to \infty} Y_+(\lambda) = 1$.

Using the projection Q along H_+ onto H_m, we can decompose the vector e as $e = Qe + p$ with $Qe \in H_m$, $p \in H_+$. Since H_p is orthogonal to H_m relative to the inner product $(\cdot, \cdot)_A$, we have $(g, e)_A = (g, p)_A$ for any $g \in H_p$. Then

$$(Tg, e) = (A^{-1}Tg, e)_A = (A^{-1}Tg, p)_A = (Tg, p), \qquad g \in H_p,$$

and

$$\begin{aligned} P_+A^{-1}TPf &= P_+Tf + P_+B_0TPf = P_+TF + \gamma(TPf, e)e_+ \\ &= P_+Tf + \gamma(TPf, p)e_+ = P_+Tf + \gamma(P_+Tf, p)e_+ \end{aligned}$$

since $P_+T = TP_+$, $P_+P = P_+$. Using (3.15) and (3.16), we obtain the following expression for $Y_+(\lambda)$:

$$Y_+(\lambda) = 1 + \gamma \int_0^1 \frac{\mu \overline{p(\mu)}}{\mu - \lambda}\, d\mu, \quad \text{where } p = (E - Q)e \in H_+ = L^2(I_+). \tag{3.18}$$

For $\frac{1}{2}\pi \leqq \arg \lambda \leqq \frac{3}{2}\pi$ and $\mu \in I_+$, we have $|\mu| < |\mu - \lambda|$. Then it follows from Lebesgue's theorem on the dominated convergence of integrable functions that the right-hand member of (3.18) converges to $1 + \gamma \int_0^1 \overline{p(\mu)}\, d\mu$ for $\lambda \to 0$, $\arg \lambda$ in the indicated interval.

Finally, we derive two methods to determine the projection P explicitly. For a third method we refer to [3].

In order to derive the first method, take an arbitrary $f \in H$. Then $Pf \in H_p$ and $FPf \in L^2(N_+, \sigma)$. Instead of Pf, we determine $g = FPf$. Since $P_+ = P_+P$, we can write

$$P_+f = P_+F^{-1}g, \qquad g \in L^2(N_+, \sigma). \tag{3.19}$$

If g happens to be Hölder continuous on N_+, then we may use (3.10) to write (3.19) in the explicit form

$$f(\mu) = \gamma \int_N \frac{g(\nu) - g(\mu)}{\nu - \mu}\, \nu\, d\sigma(\nu) + g(\mu), \qquad \mu \in I_+. \tag{3.20}$$

Thus we have derived a singular integral equation for g, which can be solved for an arbitrary function f that is Hölder continuous on I_+. In fact, the half-range completeness theorem of Case deals with a singular integral equation which is equivalent to (3.20). In the proof of that theorem, a solution of the equation is

constructed with the aid of (3.17). This solution is directly related to the solution g of (3.20) (cf. [1, 4.8], [2, 7.4]).

The second method is as follows. Using (3.15) we derive from (3.16)

$$(P_+A^{-1}TP-\lambda E)^{-1}e_+=\frac{1}{X(\lambda)}(P_+T-\lambda E)^{-1}e_+, \qquad \lambda\notin N_+,$$

where we have replaced $Y_+(\lambda)$ by $X(\lambda)$. For $\hat{\phi}$ some polynomial we take the Dunford–Taylor integral representation of $\hat{\phi}(P_+A^{-1}TP)$ in order to obtain

$$\hat{\phi}(P_+A^{-1}TP)e_+=\frac{1}{2\pi i}\int_\Gamma \hat{\phi}(\lambda)(\lambda E-P_+A^{-1}TP)^{-1}e_+\,d\lambda$$

$$=\frac{1}{2\pi i}\int_\Gamma \hat{\phi}(\lambda)(\lambda E-P_+T)^{-1}e_+\frac{d\lambda}{X(\lambda)},$$

where Γ is a simple closed contour in the complex λ-plane enclosing the spectrum N_+ of $P_+A^{-1}TP$. The integrals exist as limits of Riemann sums in H_+. Since e_+ is a continuous function on I_+, and P_+T induces a bounded linear operator in the Banach space $C(I_+)$ of continuous functions with the supremum norm, the integral in the right-hand member exists also in $C(I_+)$. This means that the left-hand member represents a continuous function ϕ, say, and that we may take values at an arbitrary point $\mu\in I_+$. Thus we obtain

$$\phi(\mu)=\frac{1}{2\pi i}\int_\Gamma \frac{\hat{\phi}(\lambda)}{\lambda-\mu}\frac{d\lambda}{X(\lambda)}, \qquad \mu\in I_+. \tag{3.21}$$

The integrand in (3.21) is holomorphic outside Γ with the exception of a pole of order $n-1$ at $\lambda=\infty$ if n is the degree of $\hat{\phi}$. An expansion of the integrand in powers of $1/\lambda$ yields as a result that ϕ is a polynomial in μ of the same degree n as $\hat{\phi}$. By analytic continuation, the formula holds for all μ inside Γ. The formula can be inverted so as to yield

$$\hat{\phi}(\nu)=\frac{1}{2\pi i}\int_{\Gamma'}\frac{\phi(\lambda)}{\lambda-\nu}X(\lambda)\,d\lambda, \qquad \nu \text{ inside } \Gamma', \tag{3.22}$$

where Γ' is any simple closed contour enclosing N_+. (Choose Γ' inside Γ and substitute $\phi(\lambda)$ from (3.21) in the right-hand member of (3.22)).

Formula (3.22) is introduced in [2, 7.3] without prior justification in order to prove the existence of the projection P by actual construction. Here we indicate briefly how the formulas (3.21) and (3.22) lead to an expression for P. For the details we refer to the cited thesis [2].

The latter two formulas are taken for $\mu\in I_+$, $\nu\in N_+$ and the contours are contracted to loops around the sets N_+, I_+, respectively. (Remark that $\lambda=\nu_0$ is a

regular point of the integrand in (3.22).) As a result, we obtain with the aid of (3.17) considering $\hat{\phi}$, ϕ as half-range functions (see [2])

$$(3.23)\quad \phi(\mu)=\gamma\int_N \{\hat{\phi}(\nu)X(-\nu)-\hat{\phi}(\mu)X(-\mu)\}\frac{\nu\,d\sigma(\nu)}{\nu-\mu}+\hat{\phi}(\mu)X(-\mu), \qquad \mu\in I_+,$$

$$(3.24)\quad \hat{\phi}(\nu)=\begin{cases} -\gamma\int_I \left\{\dfrac{\phi(\mu)}{X(-\mu)}-\dfrac{\phi(\nu)}{X(-\nu)}\right\}\dfrac{\mu\,d\mu}{\mu-\nu}+\dfrac{\phi(\nu)}{X(-\nu)}, & 0\leqq\nu\leqq+1,\\ -\gamma\int_I \dfrac{\phi(\mu)}{X(-\mu)}\dfrac{\mu\,d\mu}{\mu-\nu}, & \nu=\nu_0. \end{cases}$$

Comparing (3.23) with (3.20), we see that if we take f to be ϕ in the latter equation, then $g(\nu)=X(-\nu)\hat{\phi}(\nu)$. In other words, we have the result that for a given polynomial ϕ on I_+,

$$(FP\phi)(\nu)=X(-\nu)\hat{\phi}(\nu), \qquad \nu\in N_+,$$

where $\hat{\phi}(\nu)$ is given by (3.24).

By continuous extension to Banach spaces of Hölder continuous functions the formulas (3.23), (3.24) can be shown to hold also for ϕ, $\hat{\phi}$ that are Hölder continuous on I_+, N_+, respectively.

The results can be extended even further to $\phi\in H_+$, $\hat{\phi}\in L^2(N_+,\sigma)$ so as to yield ultimately

$$Pf=F^{-1}[\tilde{q}]F[p]f, \qquad f\in H,$$

where $[p]$, $[\tilde{q}]$ are bounded linear multiplication operators in H, $L^2(N,\sigma)$, respectively, which are defined by

$$([p]f)(\mu)=\begin{cases} f(\mu)/X(-\mu), & \mu\in I_+,\\ 0, & \mu\in I_-, \end{cases}$$

and

$$([\tilde{q}]g)(\nu)=\begin{cases} g(\nu)X(-\nu), & \nu\in N_+,\\ 0, & \nu\in N_-, \end{cases}$$

for $f\in H$, $g\in L^2(N,\sigma)$.

REFERENCES

[1] K. M. CASE AND P. F. ZWEIFEL, *Linear Transport Theory*, Addison-Wesley, Reading, Mass., 1967.

[2] R. J. HANGELBROEK, *A functional analytic approach to the linear Transport Equation*, Thesis, University of Groningen, Groningen, the Netherlands, 1973.

[3] E. W. LARSEN AND G. J. HABETLER, *A functional analytic derivation of Case's full and half-range formulas*, Comm. Pure Appl. Math., 26 (1973), pp. 525–537.

[4] I. C. GOHBERG AND M. G. KREIN, *Introduction to the Theory of Linear Nonselfadjoint Operators*, Transl. Math. Monographs, vol. 18, American Mathematical Society, Providence, R.I., 1969.

SIAM J. MATH. ANAL.
Vol. 8, No. 3, May 1977

NONLINEAR NETWORKS WITH CURRENT SOURCES AND TELLEGEN'S THEOREM*

VACLAV DOLEZAL†

Abstract. In this paper conditions are given guaranteeing the existence and uniqueness of a current distribution in a Hilbert or algebraic network driven by both EMF and current sources. Moreover, a generalization and converse of Tellegen's theorem for these networks is discussed.

Introduction. The purpose of the present paper is two-fold; first, to discuss solvability of general nonlinear networks driven by both EMF and independent current sources, second, to generalize Tellegen's theorem and give a converse.

To be more specific about our first goal, we will discuss Hilbert networks and thus extend the theory presented in papers [1]–[3]. The present solvability conditions are similar to those in [3], but more complicated in the case of a nonlinear network.

In addition to this, we will consider algebraic networks, i.e., finite nonlinear networks, whose variables (voltages, currents) belong to some linear space, which does not have any topological structure. Such networks are encountered if, for example, the variables are continuous or locally integrable functions on $[0, \infty)$. The corresponding results are analogous to those for Hilbert networks.

As for the second goal, we will deal with both Hilbert and algebraic networks and derive the respective extension of Tellegen's theorem.

It is a fact that both Hilbert and algebraic networks are special cases of an abstract network. Consequently, we will study the abstract network first.

For purposes apparent later, we will introduce this concept in a slightly more general way than was done in [3] and [1]. Also, the "quasi-inverse" will have a broader meaning here.

1. Abstract networks. First, let us define several concepts we will need.

Let X, Y be nonempty sets and let $\mathfrak{S}(Y)$ denote the collection of all nonempty subsets of Y; a mapping $A: X \to \mathfrak{S}(Y)$ will be called a set mapping from X to Y.

If $D \subset X, D \neq \emptyset$, we denote $(AD)^0 = \bigcup_{x \in D} Ax$. Moreover, if A is such that Ax is a singleton for each $x \in X$, then A will be called an operator.

Let $A: X \to \mathfrak{S}(Y)$ and let $D \subset X, D \neq \emptyset$; the set mapping $A^-: (AD)^0 \to \mathfrak{S}(D)$, defined by $A^- y = \{x : x \in D, y \in Ax\}$, will be called the quasi-inverse of A on D.

Clearly, if both A and A^- are operators, then A^- coincides with the ordinary inverse A^{-1}.

Next, let $A: X \to \mathfrak{S}(Y), D \subset X, D \neq \emptyset$; A will be called simple on D, if for $x_1, x_2 \in D, x_1 \neq x_2 \Rightarrow (Ax_1) \cap (Ax_2) = \emptyset$.

It is easy to see that:

$$A \text{ is simple on } D \Leftrightarrow A^-: (AD)^0 \to \mathfrak{S}(D) \text{ is an operator.}$$

* Received by the editors May 5, 1975, and in final revised form February 25, 1976.

† Department of Applied Mathematics and Statistics, State University of New York at Stony Brook, Stony Brook, New York 11794: on sabbatical leave at the Lehrkanzel für Systemtheorie, Johannes Kepler University in Linz, Austria. This research was supported by the National Science Foundation under Grant PO 33568-X00.

Finally, given $A: X \to \mathfrak{S}(Y)$ and an operator $B: Y \to Z$, we define the set mapping $BA: X \to \mathfrak{S}(Z)$ by $(BA)x = B(Ax) \subset Z$ for each $x \in X$; the definition of AC is analogous.

Now, let $\mathscr{L}$ be a nontrivial linear space over the field of complex or real numbers, and let M, N be linear subspaces of $\mathscr{L}$ such that $\mathscr{L} = M \oplus N$; if $D \subset \mathscr{L}$, $D \neq \emptyset$ and $Z: D \to \mathfrak{S}(\mathscr{L})$, then the ordered triplet $\mathscr{L} = [Z, N, M]$ will be called an abstract network over $\mathscr{L}$.

DEFINITION 1.1. Let $\mathscr{N} = [Z, N, M]$ be an abstract network over $\mathscr{L}$, and let $(e, k) \in \mathscr{L} \times \mathscr{L}$; an element $i \in \mathscr{L}$ will be called *a solution of $\mathscr{N}$ corresponding to the pair* (e, k), if

K_1: there exists $v \in Zi$ such that $v - e \in M$,

K_2: $i \in (k + N) \cap D$.

Observe that if $i \in \mathscr{L}$ is a solution of $\mathscr{N}$ corresponding to (e, k), then i is also a solution corresponding to $(e, k + k')$ for any $k' \in N$.

Remark. Referring to [3] and [1], let us point out the following important fact: If $\mathscr{N} = (Z, a)$ is an abstract network over $\mathscr{H}$ as defined in [3], we let $\mathscr{N}$ correspond to $[Z, N_a, N_a^{\perp}]$, where $N_a = \{x: x \in \mathscr{H}, ax = 0\}$. Then, by the above conditions K_1, K_2 and the definition in [3], an element $i \in \mathscr{H}$ is a solution of $\mathscr{N}$ corresponding to an $e \in \mathscr{H}$, exactly if i is a solution of $[Z, N_a, N_a^{\perp}]$ corresponding to the pair $(e, 0)$.

For an abstract network $\mathscr{N}$ we shall define the following sets:

If $\tilde{D} \subset D$, $\tilde{D} \neq \emptyset$, and if $k \in \mathscr{L}$, let

$$\tilde{D}_k = (k + N) \cap \tilde{D}; \tag{1.1}$$

if $\tilde{D}_k \neq \emptyset$, let

$$Q(\tilde{D}_k) = M + (Z\tilde{D}_k)^0. \tag{1.2}$$

Let P be the projection from $\mathscr{L}$ onto N along M; then it is easy to see that

$$Q(\tilde{D}_k) = P^{-1}[(PZ)\tilde{D}_k]^0. \tag{1.3}$$

Furthermore, let

$$K(\tilde{D}) = \{k: k \in \mathscr{L}, \tilde{D}_k \neq \emptyset\}, \tag{1.4}$$

and

$$R(\tilde{D}) = \{(x, y): y \in K(\tilde{D}), x \in Q(\tilde{D}_y)\}. \tag{1.5}$$

Because $\tilde{D} \neq \emptyset$, it is clear that $K(\tilde{D}) \neq \emptyset$ and consequently, $R(\tilde{D}) \neq \emptyset$.

Now the following is true:

THEOREM 1.1. *Let $\mathscr{N} = [Z, N, M]$ be an abstract network over $\mathscr{L}$, let $\tilde{D} \subset D$, $\tilde{D} \neq \emptyset$ and let $(e, k) \in \mathscr{L} \times \mathscr{L}$; then $\mathscr{N}$ possesses a solution $i \in \tilde{D}$ corresponding to (e, k), iff $(e, k) \in R(\tilde{D})$. In this case, the set I of all solutions in $\tilde{D}$ corresponding to (e, k) is given by*

$$I = (PZ)_k^{-} Pe, \tag{1.6}$$

where $(PZ)_k^{-}: [(PZ)\tilde{D}_k]^0 \to \mathfrak{S}(\tilde{D}_k)$ is the quasi-inverse of the set mapping $PZ: \tilde{D}_k \to \mathfrak{S}([(PZ)\tilde{D}_k]^0)$.

Moreover, $\mathcal{N}$ possesses a unique solution in $\tilde{D}$ for any $(e, k) \in R(\tilde{D})$, iff, for every $k^* \in K(\tilde{D})$, the set mapping PZ is simple on $\tilde{D}_{k^*}$.

The proof of this theorem follows almost the same pattern as the proofs of Theorems 1.1 and 1.2 in [3]; therefore, we omit the details. Instead, let us make a few comments.

(a) To motivate Theorem 1.1, we shall see in §§ 2 and 3 that this theorem constitutes the mathematical background for proving an existence and uniqueness result on Hilbert and algebraic networks.

(b) The main reason why we discuss the existence and uniqueness of a solution in some subset $\tilde{D}$ of D, is that we will consider this situation in the subsequent Theorem 1.2.

(c) In the above conditions we allowed k in the pair (e, k) to be any element of $K(\tilde{D})$, and not necessarily an element of some specified subset $K' \subset K(\tilde{D})$ only. This, however, is no loss of generality since the latter more general situation is obtained if $\tilde{D}$ is suitably restricted. Indeed, if $K' \subset K(\tilde{D})$, $K' \neq \emptyset$, we can put

$$\tilde{\tilde{D}} = \bigcup_{x \in K'} [(x+N) \cap \tilde{D}]; \tag{1.7}$$

then it is easy to see that $K(\tilde{\tilde{D}}) = K'$.

Also, observe this fact: Assume that $N \cap \tilde{D} \neq \emptyset$. Then $0 \in K(\tilde{D})$, and putting $K' = \{0\}$, we get by (1.7), $\tilde{\tilde{D}} = N \cap \tilde{D}$; thus, by (1.5), $R(\tilde{\tilde{D}}) = \{(x, y): y \in K', x \in Q(\tilde{\tilde{D}}_y)\} = \{(x, 0): x \in Q(\tilde{\tilde{D}}_0)\}$. Since $\tilde{\tilde{D}}_0 = N \cap \tilde{\tilde{D}} = N \cap \tilde{D}$, we have by (1.2), $R(\tilde{\tilde{D}}) = \{(x, 0): x \in M + (Z(N \cap \tilde{D}))^0\}$. Now taking into account Definition 1.1 and the above remark, we see readily that the present Theorem 1.1 yields Theorems 1.1 and 1.2 in [3].

It is worth noting that the uniqueness condition simplifies if the abstract network is linear; indeed, we have the following result.

THEOREM 1.2. *Let $\mathcal{N} = [Z, N, M]$ be an abstract network over $\mathcal{L}$, let D be a linear subspace of $\mathcal{L}$ and let $Z: D \to \mathcal{L}$ be a linear operator. Moreover, let $D' \subset D$ be a linear subspace, let $d \in N \cap D$ be a fixed element and let $\tilde{D} = d + D'$. Then for each $(e, k) \in R(\tilde{D})$ there exists in $\tilde{D}$ a unique solution i of $\mathcal{N}$ corresponding to (e, k), iff the operator PZ is 1-to-1 on $N \cap D'$. In this case*

$$i = i^* + (PZ)_0^{-1} P(e - Zi^*), \tag{1.8}$$

where $(PZ)_0^{-1}$ is the inverse of $PZ: N \cap D' \to (PZ)(N \cap D')$, and i^* is any chosen element in $\tilde{D}_k = (k+N) \cap (d+D')$.

The proof is an elementary consequence of Theorem 1.1 and is omitted. The assumption that D is a linear flat in $\mathcal{L}$ is motivated by the rather common case of a network $\mathcal{N}$ such that $\mathcal{L}$ is some function space, $\mathcal{N}$ contains differentiators and some initial condition is prescribed for the current distribution in $\mathcal{N}$.

2. Hilbert networks. Let us now apply the above results to Hilbert networks. In order to facilitate reading the paper, we will first review all previously introduced concepts which are needed in the sequel.

Let G be a locally finite oriented graph [1] which has the set of branches $\{b_1, b_2, \cdots\}$ with cardinal $c_2 \leqq \aleph_0$, the set of vertices $\{v_1, v_2, \cdots\}$ with cardinal $c_1 \leqq \aleph_0$, and let d be the incidence matrix of G (having type $c_2 \times c_1$). Let

$a=[a_{ik}]=K\cdot d^T$, where $K=\operatorname{diag}(k_1, k_2, k_3, \cdots)$ of type $c_1\times c_1$ is chosen so that the number $k_j\neq 0$ for all j's and $\sum_{ik}|a_{ik}|^2<\infty$.

Furthermore, let H be a fixed separable Hilbert space; if $c\leqq\aleph_0$, we put

$$H^c=\{x: x=[x_k] \text{ is a } c\text{-vector}, x_k\in H, \sum_k \|x_k\|^2<\infty\}$$

and

$$\langle x, y\rangle_c=\sum_k \langle x_k, y_k\rangle$$

for all $x, y\in H^c$. Then H^c is again a separable Hilbert space.

Let the operator $\hat{a}: H^{c_2}\to H^{c_1}$ be defined by $\hat{a}x=a\cdot x$. Then $\hat{a}$ is a linear bounded operator on H^{c_2}, its null-space $N_{\hat{a}}$ is closed in H^{c_2} and does not depend on the choice of the matrix K [1].

Next, let X be a $c_2\times c_0$ matrix whose columns constitute an orthonormal basis in the solution space of the equation $a\cdot\xi=0$, $\xi\in R^{c_2}$ (thus, entries of X are numbers), and let $\hat{X}: H^{c_0}\to H^{c_2}$ be defined by $\hat{X}z=X\cdot z$. As shown in [1], $\hat{X}$ is a norm-preserving isomorphism between H^{c_0} and $N_{\hat{a}}\subset H^{c_2}$.

Now, let $D\subset H^{c_2}$, $D\neq\emptyset$, and let $\hat{Z}: D\to\mathfrak{S}(H^{c_2})$ be a set mapping; then the ordered pair $\hat{\mathcal{N}}=(\hat{Z}, G)$ will be called a Hilbert network [3].

Finally, denote

$$d^TH^{c_2}=\{d^T\cdot x: x\in H^{c_2}\}; \tag{2.1}$$

clearly, $d^TH^{c_2}$ is a linear subspace of the space consisting of all c_1-vectors whose elements are in H.

DEFINITION 2.1. Let $\hat{\mathcal{N}}=(\hat{Z}, G)$ be a Hilbert network and let $(e, j)\in H^{c_2}\times d^TH^{c_2}$; an element $i\in H^{c_2}$ will be called *a solution of $\hat{\mathcal{N}}$ corresponding to the pair* (e, j), if i is a solution of the abstract network $\mathcal{N}=[\hat{Z}, N_{\hat{a}}, N_a^{\perp}]$ over H^{c_2} corresponding to $(e, k)\in H^{c_2}\times H^{c_2}$, where k is any element in H^{c_2} satisfying the equation $d^T\cdot k=j$.

In other words,

(i) there exists $v\in\hat{Z}i$ such that $v-e\in N_{\hat{a}}^{\perp}$,

(ii) $i\in(k+N_{\hat{a}})\cap D$.

Observe that this definition is meaningful, i.e., i does not depend on the choice of k.

DEFINITION 2.2. Let $\hat{\mathcal{N}}$ be a Hilbert network and let $(e, j)\in H^{c_2}\times d^TH^{c_2}$; an element $i\in H^{c_2}$ will be called *a classical solution of $\hat{\mathcal{N}}$ corresponding to the pair* (e, j) if

$$\begin{aligned} &K_1^+: \quad \text{there exists a } v\in\hat{Z}i \text{ such that} \\ &\qquad\qquad \bar{\gamma}^T\cdot(v-e)=0 \\ &\qquad \text{for all } \gamma\in R^{c_2} \text{ satisfying the equation } d^T\cdot\gamma=0, \\ &K_2^+: \quad i\in D \text{ and } d^T\cdot i=j. \end{aligned} \tag{2.2}$$

A comment on Definition 2.2 is in order. Let $\hat{\mathcal{N}}=(\hat{Z}, G)$ be a Hilbert network; then clearly G describes the structure of $\hat{\mathcal{N}}$ and $\hat{Z}$ the behavior of its elements. In the pair (e, j) let the c_2-vector e be interpreted as a vector of EMF's in

branches $b_1, b_2, \cdots$, and the c_1-vector j as a vector of currents forced to nodes $v_1, v_2, \cdots$ by independent current sources. Moreover, let the c_2-vector i be interpreted as a vector of currents $i_1, i_2, \cdots$ flowing through branches $b_1, b_2, \cdots$ and the c_2-vector $v \in \hat{Z}i$ as a vector of voltage drops caused by currents $i_1, i_2, \cdots$. Then condition K_2^+ expresses the Kirchhoff's node law, and K_1^+ the loop law. Thus, K_1^+ and K_2^+ constitute a complete description of physical phenomena in the network.

Using the same argument as in [1], we readily conclude that the following is true.

THEOREM 2.1. *Let $(e, j) \in H^{c_2} \times d^T H^{c_2}$, and let $i \in H^{c_2}$; then i is a solution of $\hat{\mathcal{N}}$ corresponding to (e, j), iff i is a classical solution of $\hat{\mathcal{N}}$ corresponding to (e, j).*

Observe the significance of this theorem. The Definition 2.2, which is a formulation of Kirchhoff's laws, employs two different spaces R^{c_2} and H^{c_2}. On the other hand, the Definition 2.1 uses a single space H^{c_2}. Thus, due to Theorem 2.1, we can work only with a simpler structure of an abstract network and remain in the framework of the space H^{c_2}.

Let us now present a (rather complicated) analog of Theorem 2.1 in [3]. To this end, some new notation is needed.

Let $\tilde{D} \subset D \subset H^{c_2}$, $\tilde{D} \neq \emptyset$; if $k \in H^{c_2}$, we will define the set $\hat{D}_k$ again by (1.1), and $Q(\hat{D}_k)$ by (1.2). (Here, of course, $N = N_{\hat{a}}$ and $M = N_{\hat{a}}^{\perp}$.) Also, let $K(D)$ and $R(D)$ be defined by (1.4) and (1.5), respectively. In addition to this, let

$$\mathcal{S}(\tilde{D}) = \{(x, d^T \cdot y) : y \in K(\tilde{D}), x \in Q(\tilde{D}_y)\}. \tag{2.3}$$

If $\xi \in H^{c_2}$, let the translation operator $T_\xi : H^{c_2} \to H^{c_2}$ be defined by

$$T_\xi z = z + \xi. \tag{2.4}$$

Finally, for every $x \in K(\tilde{D})$, choose a fixed $\xi_x \in \tilde{D}_x$ and define a subset $F_x \subset H^{c_0}$ by the relation

$$\hat{X} F_x = T_{-\xi_x} \tilde{D}_x. \tag{2.5}$$

This definition is clearly meaningful. Indeed, if $u \in T_{-\xi_x} \tilde{D}_x$, then $u = -\xi_x + v$, $v \in x + N_{\hat{a}}$, $v \in \tilde{D}$, so that $u \in -\xi_x + x + N_{\hat{a}}$. However, since $\xi_x \in x + N_{\hat{a}}$, i.e., $-\xi_x \in -x + N_{\hat{a}}$, we have $u \in N_{\hat{a}}$. Hence, $T_{-\xi_x} \tilde{D}_x \subset N_{\hat{a}}$. On the other hand, as mentioned above, $\hat{X}$ is a 1-to-1 correspondence between H^{c_0} and $N_{\hat{a}}$; consequently, F_x is uniquely defined by (2.5).

Now we have

THEOREM 2.2. *Let $\hat{\mathcal{N}} = (\hat{Z}, G)$ be a Hilbert network, let $\tilde{D} \subset D$, $\tilde{D} \neq \emptyset$ and let $(e, j) \in H^{c_2} \times d^T H^{c_2}$; then $\hat{\mathcal{N}}$ possesses a solution $i \in \tilde{D}$ corresponding to (e, j), iff $(e, j) \in \mathcal{S}(\tilde{D})$. In this case, the set I of all solutions in $\tilde{D}$ corresponding to (e, j) is given by*

$$I = T_{\xi_k} \hat{X} W_{\xi_k}^{-} \hat{X}^* e, \tag{2.6}$$

where k is any element in H^{c_2} satisfying the equation $d^T \cdot k = j$, ξ_k is in $\tilde{D}_k$, and $W_{\xi_k}^{-}$ denotes the quasi-inverse of the set mapping $W_{\xi_k} : F_k \to \mathfrak{S}([W_{\xi_k} F_k]^0)$ with W_{ξ_k} being defined by

$$W_{\xi_k} = \hat{X}^* \hat{Z} T_{\xi_k} \hat{X}. \tag{2.7}$$

Moreover, $\hat{\mathcal{N}}$ possesses a unique solution in $\tilde{D}$ for any $(e, j) \in \mathcal{S}(\tilde{D})$, iff, for every $x \in K(\tilde{D})$ and some $\xi_x \in \tilde{D}_x$, the mapping $W_{\xi_x} = \hat{X}^* \hat{Z} T_{\xi_x} \hat{X}$ is simple on F_x.

The proof of this theorem follows readily from Theorem 1.1 by applying similar arguments as in the proof of Theorem 2.1 in [3]. The details are omitted.

The comments (b) and (c) that we made on Theorem 1.1 apply also to Theorem 2.2. Thus, in particular, Theorem 2.1 in [3] appears as a special case of the present Theorem 2.2. To see this, it suffices to realize that saying "$i \in H^{c_2}$ is a solution corresponding to $e \in H^{c_2}$" defined in [3] (no current present) is equivalent to "i is a solution corresponding to the pair $(e, 0)$" in the present context.

Also, observe the following fact: If $(e, j) \in \mathcal{S}(\tilde{D})$, then clearly $j \in d^T H^{c_2}$; if $\hat{\mathcal{N}}$ is a finite network, i.e., $c_2 < \aleph_0$ then as is known,

$$j = [j_1, j_2, \cdots, j_{c_1}]^T \in d^T H^{c_2} \Leftrightarrow \sum_{m=1}^{c_1} j_m = 0.$$

The physical interpretation of this fact is plausible.

Applying Theorem 1.2 to a Hilbert network, we obtain the following result:

THEOREM 2.3. *Let $\hat{\mathcal{N}}$ be a Hilbert network, let D be a linear subspace of H^{c_2} and let $\hat{Z}: D \to H^{c_2}$ be a linear operator. Moreover, let $D' \subset D$ be a linear subspace, let $d \in N_{\hat{a}} \cap D$ be a fixed element and let $\tilde{D} = d + D'$. Then for each $(e, j) \in \mathcal{S}(\tilde{D})$ there exists in $\tilde{D}$ a unique solution i of $\hat{\mathcal{N}}$ corresponding to (e, j), iff the operator $W = \hat{X}^* \hat{Z} \hat{X}$ is 1-to-1 on F', where the linear subspace $F' \subset H^{c_0}$ is defined by*

$$\hat{X} F' = N_{\hat{a}} \cap D'. \tag{2.8}$$

In this case,

$$i = i^* + \hat{X} W^{-1} \hat{X}^* (e - \hat{Z} i^*), \tag{2.9}$$

where W^{-1} denotes the inverse of $W: F' \to WF'$, i^* is any chosen element in $\tilde{D}_k = (k + N_{\hat{a}}) \cap (d + D')$ and $k \in H^{c_2}$ is any solution of $d^T \cdot k = j$.

(The proof is obvious.)

Comparing this result with Theorem 2.1 in [3], we see that if a linear network $\hat{\mathcal{N}}$ is regular on D' (i.e., each solution in D' corresponding to an excitation by EMF's only is determined uniquely [3]), then each current distribution in $\hat{\mathcal{N}}$ forced by both EMF's and other current sources is unique. Such a statement, however, cannot be made about nonlinear networks.

Moreover, (2.9) shows that, in the case of a linear network, presence of outer current sources described by a vector j can be replaced by additional EMF sources described by the vector $-\hat{Z} i^*$. As manifested by Theorem 2.2, such a conclusion is not true for a nonlinear network.

Finally, let us mention the following useful fact: Sufficient conditions for uniqueness of a solution can easily be derived on the basis of conditions given in Theorem 1.4 in [3]. For example, if any of conditions (1.8), (1.10) or (1.12) is satisfied for all $x_1, x_2 \in D \subset H^{c_2}$, then we have uniqueness of a solution in D for any $(e, j) \in \mathcal{S}(D)$. This follows immediately from Theorem 1.1 and the fact that $\tilde{D}_k \subset D$ for each $k \in K(D)$. Since these conclusions are straightforward, we omit the details.

3. Algebraic networks. In many engineering problems we encounter networks whose variables (voltages, currents) are continuous functions on $[0, \infty)$, or locally bounded measurable functions, or locally integrable functions on $[0, \infty)$. Such variables cannot be imbedded into a Hilbert space, and consequently, we are unable to apply the model of a Hilbert network. Fortunately, many results valid for Hilbert networks can be extended to the case that the underlying space is just a linear space not necessarily equipped with any topological structure. For this generalization, however, we have to pay a price: We have to assume that the networks under consideration are finite since the concept of convergence is missing in this setting.

In order to study these (algebraic) networks more closely, we will again use the concept and properties of the abstract network. To this end, let us carry out some auxiliary considerations.

Let G be a finite oriented graph having c_2 branches and c_1 vertices, and let d be the $(c_2 \times c_1)$ incidence matrix of G. Define the operator $\tilde{a}: R^{c_2} \to R^{c_1}$ by $\tilde{a}\xi = d^T \cdot \xi$, and let $N_{\tilde{a}} = \{\xi: \xi \in R^{c_2}, \tilde{a}\xi = 0\}$.

We will assume in the sequel that $N_{\tilde{a}} \neq \{0\}$, which amounts to requiring that G contains at least one loop.

Next, since every Euclidean space R^c is a Hilbert space with inner product $\langle \xi_1, \xi_2 \rangle = \bar{\xi}_1^T \cdot \xi_2$, the set $N_{\tilde{a}}$ is a closed linear subspace of R^{c_2}. Choose some fixed orthonormal basis $\{\xi^1, \xi^2, \cdots, \xi^{c_0}\}$ in $N_{\tilde{a}}$, and let X be the $c_2 \times c_0$ matrix having the c_2-vectors ξ^i as columns. Define operators $\tilde{X}$ and $\tilde{X}^*$ by

$$
\begin{aligned}
&\tilde{X}: R^{c_0} \to R^{c_2}, \qquad \tilde{X}\xi = X \cdot \xi, \\
&\tilde{X}^*: R^{c_2} \to R^{c_0}, \qquad \tilde{X}^*\eta = \bar{X}^T \cdot \eta,
\end{aligned} \tag{3.1}
$$

and let $N_{\tilde{X}^*} = \{\eta: \eta \in R^{c_2}, \tilde{X}^*\eta = 0\}$.

Using elementary arguments we see easily that:

(a) $N_{\tilde{a}} = \tilde{X}R^{c_0}$ and $\tilde{X}$ is 1-to-1.

(b) $\tilde{a}\tilde{X} = 0$ on R^{c_0}.

(c) $\tilde{X}^*\tilde{X} = I$ on R^{c_0}.

(d) $R^{c_2} = N_{\tilde{X}^*} \oplus N_{\tilde{a}}$.

Now let L be a fixed nontrivial linear space over the field of complex or real numbers; if $c \geqq 1$ is an integer, L^c will denote the Cartesian product $L \times L \times \cdots \times L$ having c factors. Clearly, L^c is a real or complex linear space with "elementwise" operations, depending on whether L is real or complex. For convenience, we will interpret the elements in L^c as c-vectors.

Define the operators $\hat{a}$, $\hat{X}$ and $\hat{X}^*$ by

$$
\begin{aligned}
&\hat{a}: L^{c_2} \to L^{c_1}, \qquad \hat{a}x = d^T \cdot x, \\
&\hat{X}: L^{c_0} \to L^{c_2}, \qquad \hat{X}y = X \cdot y, \\
&\hat{X}^*: L^{c_2} \to L^{c_0}, \qquad \hat{X}^*z = \bar{X}^T \cdot z,
\end{aligned} \tag{3.2}
$$

and let $N_{\hat{a}} = \{x: x \in L^{c_2}, \hat{a}x = 0\}$, $N_{\hat{X}^*} = \{z: z \in L^{c_2}, \hat{X}^*z = 0\}$. In definitions (3.2) we understand that if L is a real space, then X is a real matrix (i.e., in (3.1) the Euclidean spaces are real); in the opposite case X can be complex.

Using the above relations (a)–(d), we can prove the following proposition:

LEMMA 3.1.

(i) $N_{\hat{a}} = \hat{X}L^{c_0}$ *and* $\hat{X}$ *is* 1-*to*-1.

(ii) $\hat{a}\hat{X} = 0$ *on* L^{c_0}.

(iii) $\hat{X}^*\hat{X} = I$ *on* L^{c_0}.

(iv) $L^{c_2} = N_{\hat{X}^*} \oplus N_{\hat{a}}$.

(v) *If* P *is the projection from* L^{c_2} *onto* $N_{\hat{a}}$ *along* $N_{\hat{X}^*}$, *then* $P = \hat{X}\hat{X}^*$.

(vi) *Let* $u \in L^{c_2}$; *then* $\bar{\gamma}^T \cdot u = 0$ *for all* $\gamma \in N_{\hat{a}} \Leftrightarrow u \in N_{\hat{X}^*}$.

Proof. First note the fact that all operators involved are defined via finite matrices whose entries are numbers, and consequently, the product of such matrices obeys the associative law. Having this in mind, (ii) and (iii) are trivial.

The inclusion $\hat{X}L^{c_0} \subset N_{\hat{a}}$ is obvious. Conversely, let $x \in N_{\hat{a}}$, i.e., $d^T \cdot x = 0$. Choosing some basis $\mathscr{L}$ in L we can find elements $l_1, l_2, \cdots, l_m \in \mathscr{L}$ and $c_2 \times m$ matrix R with constant elements such that $x = R \cdot l$, where $l = [l_1, l_2, \cdots, l_m]^T$. Thus, $d^T \cdot (R \cdot l) = (d^T \cdot R) \cdot l = 0 \Rightarrow d^T \cdot R = 0$. Hence, by proposition (a), there exists a $c_0 \times m$ matrix E such that $R = X \cdot E$, so that $x = (X \cdot E) \cdot l = X \cdot (E \cdot l)$; since $E \cdot l \in L^{c_0}$, we have $x \in \hat{X}L^{c_0}$ and the inclusion $N_{\hat{a}} \subset \hat{X}L^{c_0}$ is proven. An analogous argument shows that $\hat{X}$ is 1-to-1.

The relation (iv) follows readily from (d) by employing the basis $\mathscr{L}$ in L.

As for (v), let $Q = \hat{X}\hat{X}^*: L^{c_2} \to L^{c_2}$. From (iii) we infer that $Q^2 = Q$, i.e., Q is a projection. Moreover, (ii) implies that $QL^{c_2} \subset N_{\hat{a}}$. Conversely, let $x \in N_{\hat{a}}$; then by (i), $x = \hat{X}z$ for some $z \in L^{c_0}$, and since $\hat{X}^*$ maps L^{c_2} onto L^{c_0} by (iii), we have $z = \hat{X}^*w$ for some $w \in L^{c_2}$. Hence, $x = \hat{X}\hat{X}^*w = Qw$, i.e., $N_{\hat{a}} \subset QL^{c_2}$. Consequently, by (iv), Q is the projection onto $N_{\hat{a}}$ and $N_{\hat{X}^*}$.

Finally, (vi) is an elementary consequence of (a) and (iv).

We are ready to define the algebraic network.

Let G be a finite oriented graph, let $D \subset L^{c_2}$, $D \neq \emptyset$, and let $\hat{Z}: D \to \mathfrak{S}(L^{c_2})$ be a set mapping; then the ordered pair $\hat{\mathscr{N}} = (\hat{Z}, G)$ will be called an algebraic network.

Denote

$$d^T L^{c_2} = \{d^T \cdot x : x \in L^{c_2}\} \subset L^{c_1}. \tag{3.3}$$

DEFINITION 3.1. Let $\hat{\mathscr{N}} = (\hat{Z}, G)$ be an algebraic network, and let $(e, j) \in L^{c_2} \times d^T L^{c_2}$; an element $i \in L^{c_2}$ will be called *a solution of* $\hat{\mathscr{N}}$ *corresponding to* (e, j), if

$$\begin{aligned} K_1^+: &\quad \text{there exists a } v \in \hat{Z}i \text{ such that} \\ &\qquad \bar{\gamma}^T \cdot (v - e) = 0 \\ &\qquad \text{for all } \gamma \in R^{c_2} \text{ satisfying the equation } d^T \cdot \gamma = 0, \\ K_2^+: &\quad d^T \cdot i = j \quad \text{and} \quad i \in D. \end{aligned} \tag{3.4}$$

Clearly, conditions K_1^+, K_2^+ are formulations of Kirchhoff's laws for our algebraic network.

As pointed out above, an algebraic network is a special case of an abstract network. Indeed, we have the following:

THEOREM 3.1. *Let* $\hat{\mathscr{N}} = (\hat{Z}, G)$ *be an algebraic network with* $N_{\hat{a}} \neq \{0\}$, *and let* $(e, j) \in L^{c_2} \times d^T L^{c_2}$; *then an element* $i \in L^{c_2}$ *is a solution of* $\hat{\mathscr{N}}$ *corresponding to* (e, j),

iff i is a solution of the abstract network $\mathcal{N} = [\hat{Z}, N_{\hat{a}}, N_{\hat{X}^*}]$ *over* L^{c_2} *corresponding to* (e, k), *where k is any element in* L^{c_2} *satisfying the equation* $d^T \cdot k = j$.

Proof. First, $\mathcal{N}$ is truly an abstract network over L^{c_2}, since $L^{c_2} = N_{\hat{a}} \oplus N_{\hat{X}^*}$ by Lemma 3.1(iv). Moreover, it is clear that $K_2^+ \Leftrightarrow i \in (k + N_{\hat{a}}) \cap D \Leftrightarrow K_2$ in Definition 1.1 for $[\hat{Z}, N_{\hat{a}}, N_{\hat{X}^*}]$. Also, letting $u = v - e$, we have by Lemma 3.1(vi), $K_1^+ \Leftrightarrow u \in N_{\hat{X}^*} \Leftrightarrow K_1$ in Definition 1.1 for $[\hat{Z}, N_{\hat{a}}, N_{\hat{X}^*}]$. Hence, the proof.

Using the same arguments as in § 2, we can easily derive necessary and sufficient conditions for the existence and/or uniqueness of a solution of an algebraic network. To this end, define the following concepts:

Let $\tilde{D} \subset D$, $\tilde{D} \neq \varnothing$ be a given set; for each $k \in L^{c_2}$, let

$$\tilde{D}_k = (k + N_{\hat{a}}) \cap \tilde{D}, \tag{3.5}$$

and

$$Q(\tilde{D}_k) = N_{\hat{X}^*} + (\hat{Z}\tilde{D}_k)^0 \tag{3.6}$$

provided $\tilde{D}_k \neq \varnothing$. Furthermore, let

$$K(\tilde{D}) = \{k : k \in L^{c_2}, \tilde{D}_k \neq \varnothing\}, \tag{3.7}$$

$$\mathcal{S}(\tilde{D}) = \{(x, d^T \cdot y) : y \in K(\tilde{D}), x \in Q(\tilde{D}_y)\}. \tag{3.8}$$

If $\xi \in L^{c_2}$, let $T_\xi : L^{c_2} \to L^{c_2}$ be defined by $T_\xi z = z + \xi$.

Finally, if $\xi_x \in \tilde{D}_x$, define the set $F_x \subset L^{c_0}$ by

$$\hat{X}F_x = T_{-\xi_x}\tilde{D}_x. \tag{3.9}$$

As in § 2 it follows by Lemma 3.1 that F_x is uniquely defined. Now, the following assertion is true.

THEOREM 3.2. *Theorem 2.2 remains true, if the term "Hilbert network" is replaced by "algebraic network", H is replaced by L, and the symbols have the meaning defined by* (3.2)–(3.9).

In particular, if $N_{\hat{a}} \cap D \neq \varnothing$ *and* $e \in L^{c_2}$, *then* $\hat{\mathcal{N}}$ *possesses a solution corresponding to* $(e, 0)$, *iff*

$$e \in Q(D) = N_{\hat{X}^*} + [\hat{Z}(N_{\hat{a}} \cap D)]^0. \tag{3.10}$$

In this case, the set I of all solutions of $\hat{\mathcal{N}}$ *corresponding to* $(e, 0)$ *is given by*

$$I = \hat{X}W^-\hat{X}^*e, \tag{3.11}$$

where W^- *denotes the quasi-inverse of the set mapping* $W = \hat{X}^*\hat{Z}\hat{X} : F \to \mathfrak{S}((WF)^0)$ *and* $F \subset L^{c_0}$ *is defined by* $\hat{X}F = N_{\hat{a}} \cap D$. *Moreover, I is a singleton for every* $e \in Q(D)$, *iff W is simple on F.*

Remark. In formulas (3.10), (3.11) the operators $\hat{X}$ and $\hat{X}^*$ can be replaced by operators $\hat{Y}$ and $\hat{Y}^*$ generated by matrices Y and $\bar{Y}^T$, respectively, where the columns of Y constitute a (not necessarily orthonormal) basis in $N_{\hat{a}}$. Then, of course, F has to be replaced by a set F' defined by $\hat{Y}F' = N_{\hat{a}} \cap D$. To verify this fact it suffices to realize that $X = Y \cdot S$ with S being a nonsingular $c_0 \times c_0$ matrix. Then $\hat{X} = \hat{Y}\hat{S}$ and $\hat{S}$ being a bijection between L^{c_2} and itself, and similarly for $\hat{X}^*$.

Note that, with changes described above, Theorem 2.3 also holds for algebraic networks; however, since this matter is straightforward, we omit the details.

4. Tellegen's theorem. Let us now consider a generalization and converse of the classical Tellegen theorem [4]. First, we are going to discuss the case of a Hilbert network.

Having Kirchhoff's laws in mind and thus accepting the classical solution as the "true solution concept", we see readily that the sought generalization and converse is already furnished by Theorem 2.1 and Definition 2.1. Actually, realizing that $u \in N_{\hat{a}}^{\perp} \Leftrightarrow \langle c, u\rangle_{c_2} = 0$ for all $c \in N_{\hat{a}}$, we can rephrase our results as follows:

THEOREM 4.1. *An element $i \in H^{c_2}$ is a (classical) solution of a Hilbert network $(\hat{Z}, G)$ corresponding to $(e, j) \in H^{c_2} \times d^T H^{c_2}$, iff*

$$K_1^*: \quad \text{there exists a } v \in \hat{Z}i \text{ such that}$$

$$\langle c, v - e\rangle_{c_2} = 0 \tag{4.1}$$

for every vector $c \in H^{c_2}$ satisfying the equation $d^T \cdot c = 0$,

$$K_2^*: \quad i \in D \quad \text{and} \quad d^T \cdot i = j.$$

Obviously, condition K_1^* is a type of relation we encounter in the classical Tellegen theorem. Also, let us stress the sufficiency part of our assertion, i.e., if $i \in H^{c_2}$ meets K_1^* amd K_2^*, it must be a solution of the network. Further comments on this result are not necessary, since they have been made in § 2.

Let us now establish a generalization of the Tellegen theorem for algebraic networks. To do this, we define some additional concepts.

Let L' be a fixed linear space over the same system of scalars as L has, and let (L, L') denote the linear space of all linear operators from L into L'; moreover, let K be a fixed linear subspace of (L, L'). The subspace K will be called complete if $x \in L$, $Ax = 0$ for all $A \in K$ implies that $x = 0$.

Observe that if K contains at least one 1-to-1 operator, then K is complete.

THEOREM 4.2. *Let $\hat{\mathcal{N}} = (\hat{Z}, G)$ be an algebraic network, and let $(e, j) \in L^{c_2} \times d^T L^{c_2}$.*

(i) *Let $i \in L^{c_2}$ be a solution of $\hat{\mathcal{N}}$ corresponding to (e, j); then there exists $v \in \hat{Z}i$ such that*

$$\Gamma^T \cdot (v - e) = 0 \tag{4.2}$$

for every $\Gamma \in K^{c_2}$ satisfying the equation $d^T \cdot \Gamma = 0$.

(ii) *Let K be complete, let $i \in D$ and $d^T \cdot i = j$; if there exists $v \in \hat{Z}i$ such that (4.2) holds for every $\Gamma \in K^{c_2}$ satisfying the equation $d^T \cdot \Gamma = 0$, then i is a solution of $\hat{\mathcal{N}}$ corresponding to (e, j).*

Proof. Denote $\mathfrak{N} = \{\Gamma: \Gamma \in K^{c_2}, d^T \cdot \Gamma = 0\}$; then we have

$$\Gamma \in \mathfrak{N} \Leftrightarrow \Gamma = X \cdot \omega \quad \text{with } \omega \in K^{c_0}. \tag{4.3}$$

Indeed, recalling the definitions (3.2) of $\hat{a}$ and $\hat{X}$, we see readily that the equivalence (4.3) is exactly the proposition (i) in Lemma 3.1, where L is replaced by K.

Next, observe the following fact: since d is a real matrix, we can select an orthonormal basis $\{\xi^1, \xi^2, \cdots, \xi^{c_0}\}$ in $N_{\hat{a}}$ so that each vector ξ^i is real, independent of whether R^{c_2} is a real or complex space. Consequently, we can assume that X is a real matrix.

(i) Let $i \in L^{c_2}$ be a solution of $\hat{\mathcal{N}}$ corresponding to (e, j). Then, by K_1^+, there exists $v \in \hat{Z}i$ such that $\bar{\gamma}^T \cdot u = 0$ for any $\gamma \in N_{\hat{a}}$, where $u = v - e$. Thus, by (vi) in Lemma 3.1, $\hat{X}^* u = \bar{X}^T \cdot u = X^T \cdot u = 0$. Now, if $\Gamma \in \mathfrak{N}$, we have by (4.3), $\Gamma = X \cdot \omega$ for some $\omega \in K^{c_0}$. Hence, $0 = \omega^T \cdot (X^T \cdot u) = (\omega^T \cdot X^T) \cdot u = \Gamma^T \cdot u$ and (i) is proved.

For (ii), if the hypothesis is satisfied, then i fulfills K_2^+. Again letting $u = v - e$, we have by (4.3) for any $\omega \in K^{c_0}$,

$$0 = (X \cdot \omega)^T \cdot u = \omega^T \cdot (X^T \cdot u). \tag{4.4}$$

However, if $X^T \cdot u$ has components $(X^T \cdot u)_j$, $j = 1, 2, \cdots, c_0$, then (4.4) implies that $\nu(X^T \cdot u)_j = 0$ for each $\nu \in K$ and every j; thus, by completeness of K, $(X^T \cdot u)_j = 0$ for $j = 1, 2, \cdots, c_0$. Consequently, $X^T \cdot u = 0 \Rightarrow u \in N_{\hat{X}^*} \Rightarrow K_1^+$ holds by (vi) in Lemma 3.1. Hence i is a solution of $\mathcal{N}$ corresponding to (e, j) and (ii) is proved.

Observe that if $\hat{\mathcal{N}} = (\hat{Z}, G)$ is an algebraic network (consequently, finite) whose underlying space L happens to be a Hilbert space, then Theorem 4.1, which is valid for Hilbert (not necessarily finite) networks, follows for $\hat{\mathcal{N}}$ from Theorem 4.2. To verify this, let $L' = R^1$ and let $K = \{A_\alpha : \alpha \in L\}$, where each operator $A_\alpha : L \to R^1$ is defined by $A_\alpha x = \langle \alpha, x \rangle$. Clearly, K is complete, since $\langle \alpha, x \rangle = 0$ for all $\alpha \in L$ implies that $x = 0$. Moreover, it is easy to see that $\Gamma = [A_{\alpha_1}, A_{\alpha_2}, \cdots, A_{\alpha_{c_2}}]^T \in K^{c_2}$ and $d^T \cdot \Gamma = 0 \Leftrightarrow \alpha = [\alpha_1, \alpha_2, \cdots, \alpha_{c_2}]^T \in L^{c_2}$ and $d^T \cdot \alpha = 0$.

Thus, with $v - e = u = [u_1, u_2, \cdots, u_{c_2}]^T$, (4.2) reads $0 = \Gamma^T \cdot u = \sum_{j=1}^{c_2} A_{\alpha_j} u_j = \sum_{j=1}^{c_2} \langle \alpha_j, u_j \rangle = \langle \alpha, u \rangle_{c_2}$; this, however, is exactly equation (4.1). Hence, our claim.

Let us now consider two simple examples which illustrate the applications of Theorem 4.2.

Example 1. Let G be a finite oriented graph having c_2 branches, and let d be its incidence matrix. Let L be the space of all locally integrable functions on $[0, \infty)$. Given $i_0 \in R^{c_2}$ with $d^T \cdot i_0 = 0$, let $D \subset L^{c_2}$ be defined by $D = \{x : x = [x_k]$ is a c_2-vector, x_k absolutely continuous on $[0, \infty)$ for $k = 1, 2, \cdots, c_2, x(0) = i_0\}$.

Furthermore, denote $T = R^{c_2} \times [0, \infty)$, and assume that the functions $l, r, s : T \to R^{c_2}$ have the following properties:

(i) l is differentiable at every point of T,

(ii) $r(x(t), t), s(x(t), t) \in L^{c_2}$ for every $x \in L^{c_2}$.

Now, let $\hat{Z} : D \to L^{c_2}$ be defined by

$$(\hat{Z}x)(t) = \{l(x(t), t)\}' + r(x(t), t) + s\left(\int_0^t x(\tau)\, d\tau, t\right); \tag{4.5}$$

then the algebraic network $\hat{\mathcal{N}} = (\hat{Z}, G)$ will be called an L, R, C-network.

It is easy to see that a nonlinear, time-varying L, R, C-network is truly under consideration. Indeed, if $\xi \in R^{c_2}$ and $d^T \cdot \xi = 0$, we can interpret $l(\xi, t)$ as a vector of magnetic fluxes at time t generated by the (direct) current vector ξ. Similarly, $r(\xi, t)$ can be interpreted as a vector of voltage drops on resistors at time t. Finally, if $\eta \in R^{c_2}$ has a meaning of charges at capacitors, $s(\eta, t)$ is a vector of capacitive voltage drops at t. Hence, if the vector function $x(t)$ has a meaning of currents, then the right-hand side of (4.5) is a vector of total voltage drops in branches of our network.

Note also that if there are no mutual couplings between branches, then l, r, s have, of course, the "diagonal form", i.e., $l(\xi, t) = [l_1(\xi_1, t), l_2(\xi_2, t), \cdots, l_{c_2}(\xi_{c_2}, t)]^T$, and similarly for r and s.

To avoid ambiguity, a solution $i \in D$ of $\hat{\mathcal{N}}$ corresponding to $(e, 0)$ for some $e \in L^{c_2}$ will be called an elementary solution.

It is well known [5] that the requirement $i \in D$, which is enforced by presence of differentiators in the network, imposes severe limitations on the existence of an elementary solution. (Note that this is so even if $\hat{Z}$ is linear.)

However, this inconvenience can easily be surmounted by defining the concept of a "generalized solution". Indeed, let $\hat{\hat{Z}}: L^{c_2} \to L^{c_2}$ be defined by

$$(4.6)\quad (\hat{\hat{Z}}x)(t) = l(x(t), t) - l(i_0, 0) + \int_0^t r(x(\tau), \tau)\, d\tau + \int_0^t s\left(\int_0^t x(\sigma)\, d\sigma, \tau\right) d\tau.$$

(Note that now $\hat{\hat{Z}}$ is defined on the entire space L^{c_2}!)

Let $e \in L^{c_2}$; an element $i \in L^{c_2}$ will be called a generalized solution of $\hat{\mathcal{N}}$ corresponding to $(e, 0)$, if i is a solution of $\hat{\hat{\mathcal{N}}} = (\hat{\hat{Z}}, G)$ corresponding to $(\int_0^t e(\tau)\, d\tau, 0)$.

It is clear that the class of generalized solutions is much bigger than the class of elementary solutions. On the other hand, the relation between the two solution concepts is very simple as is documented by the following:

PROPOSITION 1. *Let $e \in L^{c_2}$; then i is an elementary solution of $\hat{\mathcal{N}}$ corresponding to $(e, 0) \Leftrightarrow i$ is a generalized solution of $\hat{\mathcal{N}}$ corresponding to $(e, 0)$ and $i \in D$.*

Although this fact follows directly from (4.5), (4.6) and K_1^+, let us prove it by using Theorem 4.2; to this end, put $L' = L$, and let $K = \{\alpha J: \alpha \in R^1\}$, where $(Jx)(t) = \int_0^t x(\tau)\, d\tau$. Clearly, K is complete, since J is 1-to-1 on L. Thus, by Theorem 4.2 and the above definitions, i is an elementary solution of $\hat{\mathcal{N}}$ corresponding to $(e, 0) \Leftrightarrow \Gamma^T \cdot (\hat{Z}i - e) = 0$ for all $\Gamma \in K^{c_2}$ with $d^T \cdot \Gamma = 0$, and $i \in D$, $d^T \cdot i = 0$. However, by (4.5), (4.6), the relation $\Gamma^T \cdot (\hat{Z}i - e) = 0$ is nothing else than K_1^+ for $\hat{\mathcal{N}}$ with e replaced by Je. Since $D \subset L^{c_2}$, the right-hand side of our equivalence reads: i is a generalized solution of $\hat{\mathcal{N}}$ corresponding to $(e, 0)$ and $i \in D$. This verifies our claim.

Example 2. Let $\hat{\mathcal{N}} = (\hat{Z}, G)$ be the same L, R, C-network as in Example 1, but now let L be the space of all measurable, locally bounded functions on $[0, \infty)$. Also, let the definition of D be modified accordingly, i.e., let us add the requirement $x' \in L^{c_2}$. We are going to show that the following is true.

PROPOSITION 2. *Let $e \in L^{c_2}$; then i is an elementary solution of $\hat{\mathcal{N}}$ corresponding to $(e, 0)$, iff $i \in D$, $d^T \cdot i = 0$ and*

$$(4.7)\qquad \int_0^t c^T(\tau) \cdot [(\hat{Z}i)(\tau) - e(\tau)]\, d\tau = 0$$

for all $c \in L^{c_2}$ with $d^T \cdot c = 0$.

Note that by setting $c = i$ in (4.7), we get $\int_0^t i^T \cdot \hat{Z}i\, d\tau = \int_0^t i^T \cdot e\, d\tau$, i.e., an energy relation analogous to (4.1) holding for a Hilbert network.

To prove our proposition, let $L' = L$ and put $K = \{J_a: a \in L\}$, where $J_a: L \to L$ is defined by $(J_a x)(t) = \int_0^t a(\tau)x(\tau)\, d\tau$. Clearly, K is complete, since J_a with $a = \text{const.} \neq 0$ is 1-to-1. Moreover, it is easy to see that (4.2) is precisely the relation (4.7). The rest follows from Theorem 4.2.

REFERENCES

[1] V. DOLEZAL, *Hilbert networks I*, SIAM J. Control, 12 (1974), pp. 755–778.

[2] V. DOLEZAL AND A. H. ZEMANIAN, *Hilbert networks II.—Some qualitative properties*, Ibid., 13 (1975), pp. 153–161.

[3] V. DOLEZAL, *Generalized Hilbert networks*, Ibid., 14 (1976), pp. 26–41.

[4] P. PENFIELD, R. SPENCE AND S. DUINKER, *Tellegen's Theorem and Electrical Networks*, MIT Press, Cambridge, Mass., 1970.

[5] V. DOLEZAL, *Dynamics of Linear Systems*, Academia/Noordhoff, Groningen, the Netherlands, 1967, p. 9–17.

SIAM J. MATH. ANAL.
Vol. 8, No. 3, May 1977

INTEGRAL REPRESENTATIONS AND INEQUALITIES FOR BESSEL FUNCTIONS*

A. McD. MERCER†

Abstract. In a previous note the author showed how various inequalities, typified by Grünbaum's inequality $1+J_0(a)\geqq J_0(b)+J_0(c)$ $(a^2=b^2+c^2)$, could easily be obtained from a certain integral representation for the Bessel functions. In the present note a method is described which provides more general representations of this kind, and examples of the resulting inequalities are given.

1. Introduction. In [1] we showed how short proofs of Grünbaum's inequality

$$1+J_0(a)\geqq J_0(b)+J_0(c), \qquad a^2=b^2+c^2, \tag{1.1}$$

and others of a similar nature could be constructed using the integral representation

$$\mathscr{J}_\nu(\lambda)=\frac{1}{S(1)}\int_{S(1)}\left\{\prod_{k=1}^{n}\cos\lambda_k x_k\right\}\sigma_1, \qquad n\geqq 2, \tag{1.2}$$

where $\sum_{k=1}^{n}\lambda_k^2=\lambda^2$.

Here, $\mathscr{J}_\nu(\lambda)=\Gamma(\nu+1)(2/\lambda)^\nu J_\nu(\lambda)$, $\nu=\frac{1}{2}n-1$, $S(1)$ denotes both the $(n-1)$-dimensional manifold $\|\underline{x}\|=1$ in Euclidian space E^n and its volume, while σ_1 is the volume element in $S(1)$. This formula is a consequence of a theorem proved in [2].

The technique which deduces (1.1), for example, from (1.2) would work equally well to provide inequalities for functions which had integral representations of a kind more general than (1.2). For example, there would be no essential difference if, instead of the integral in (1.2), we had

$$\frac{1}{S(1)}\int_{S(1)}\left\{\prod_{k=1}^{n}\cos\lambda_k x_k\right\}w\sigma_1$$

in which w denotes a nonnegative function defined in $S(1)$. Now the theorem proved in [2] does not provide representations of this more general kind and so it is the purpose of this note to present a method which does. We shall describe this method in the next section and the basic result is to be found in (2.7) below. In the third and final section we give some examples both of the representations and of the resulting inequalities.

2. Let r denote distance from the origin in E^n. Then the Helmholtz equation

$$\sum_{k=1}^{n}\frac{\partial^2\phi}{\partial x_k^2}+\lambda^2\phi=0 \tag{2.1}$$

can be written as

$$\frac{\partial^2\phi}{\partial r^2}+\frac{n-1}{r}\frac{\partial\phi}{\partial r}+\frac{1}{r^2}\Delta\phi+\lambda^2\phi=0. \tag{2.2}$$

* Received by the editors January 11, 1976.

† Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario, Canada.

Here Δ is the Beltrami operator or surface Laplacian in the manifold $S(1)$. If σ_1 denotes the element of volume in $S(1)$, then we have the identity of $(n-1)$-forms $(\Delta\phi)\sigma_1 = d(*\, d\phi)$, where d is the operator of exterior differentiation and $*$ is the Hodge operator in $S(1)$. For more details of these matters we refer the reader to [3]. It follows then that (2.2) can be written in the Pfaffian form

$$r^2\left\{\frac{\partial^2\phi}{\partial r^2}+\frac{n-1}{r}\frac{\partial\phi}{\partial r}\right\}\sigma_1+d(*\, d\phi)+r^2\lambda^2\phi\sigma_1=0. \tag{2.3}$$

We have written (2.2) in this form so as to be able to separate variables. If $\underline{v}$ denotes the $(n-1)$-tuple of hyperspherical polar coordinates in the manifold $S(1)$, then the variables we shall want to "separate" will be r and $\underline{v}$. Accordingly let $R(r)H(\underline{v})$ be a solution of (2.3). The usual technique leads to the separated equations

$$\frac{d^2R}{dr^2}+\frac{n-1}{r}\frac{dR}{dr}+\left(\lambda^2-\frac{a}{r^2}\right)R=0, \tag{2.4}$$

$$d(*\, dH)+aH\sigma_1=0. \tag{2.5}$$

This second equation is simply the Pfaffian form of Helmholtz' equation in the manifold $S(1)$. As is well known, it has solutions which are single-valued in $S(1)$ when $a=a_m=m(m+n-2)$ $(m=0,1,2,\cdots)$. These are the hyperspherical harmonics of degree m and we denote a typical one by $H_m(\underline{v})$. The general solution of (2.4) is then

$$r^m\{A\mathcal{J}_{m+\nu}(\lambda r)+B\mathcal{Y}_{m+\nu}(\lambda r)\},$$

where ν and $\mathcal{J}_\nu$ are as defined in § 1 and $\mathcal{Y}_\nu(\lambda)=\Gamma(\nu+1)(2/\lambda)^\nu Y_\nu(\lambda)$.

Let $\Phi(\underline{x})$ denote any solution of (2.1) in E^n. Then $\Phi(\underline{x})$ (or more accurately its transformed form) will satisfy (2.3). Accordingly we multiply (2.3) by $H_m(\underline{v})$, (2.5) by $\Phi(\underline{x})$, subtract and we get

$$r^2\left\{\frac{\partial^2\Phi}{\partial r^2}+\frac{n-1}{r}\frac{\partial\Phi}{\partial r}+\lambda^2\Phi\right\}H_m\sigma_1+\{H_m d(*d\Phi)-\Phi d(*dH_m)\} \tag{2.6}$$
$$-a_mH_m\Phi\sigma_1=0.$$

Now the second expression in brackets can be written as $d\{H_m(*\, d\Phi)-\Phi(*\, dH_m)\}$ which is to say that it is an exact $(n-1)$-form in the $(n-1)$-dimensional manifold $S(1)$. If we denote it by dW, then by Stokes' theorem we will have

$$\int_{S(1)} dW=\int_{\partial S(1)} W;$$

this is zero because the manifold $S(1)$ is closed in the sense that its boundary $\partial S(1)$ vanishes. Hence if we integrate throughout (2.6) over $S(1)$ we will get

$$\int_{S(1)} r^2\left\{\frac{\partial^2\Phi}{\partial r^2}+\frac{n-1}{r}\frac{\partial\Phi}{\partial r}+\lambda^2\Phi\right\}H_m\sigma_1-a_m\int_{S(1)} H_m\Phi\sigma_1=0.$$

If we suppose, for example, that all our functions are continuously twice differentiable, then the differential operator appearing here and the operation of

integration over the fixed manifold $S(1)$ (i.e., fixed with respect to r) can be interchanged. The result is that

$$\left\{\frac{\partial^2}{\partial r^2}+\frac{n-1}{r}\frac{\partial}{\partial r}+\left(\lambda^2-\frac{a_m}{r^2}\right)\right\}\int_{S(1)} H_m(\underline{v})\Phi(\underline{x})\sigma_1=0,$$

which is to say that the integral here is a solution of (2.4). We conclude then that if $\Phi(\underline{x})$ denotes *any* solution of (2.1), then for some choice of A and B we will have

$$(2.7)\qquad \frac{1}{S(1)}\int_{S(1)} \Phi(\underline{x})H_m(\underline{v})\sigma_1=r^m\{A\mathscr{J}_{m+\nu}(\lambda r)+B\mathscr{Y}_{m+\nu}(\lambda r)\},\qquad \|\underline{x}\|=r.$$

This is as far as we can proceed with these general considerations. In any particular case, the coefficients A and B, which will depend on the choice of $\Phi(\underline{x})$, will be determined by considering the behavior of each side of (2.7) as $r\to 0$.

3. Examples. The simplest case is that quoted in § 1 and arises on taking $n\geqq 2$,

$$\Phi(\underline{x})=\prod_{k=1}^{n}\cos\lambda_k x_k,\quad \text{where}\ \sum_{k=1}^{n}\lambda_k^2=\lambda^2,$$

$m=0$ and $H_0(\underline{v})=1$. It is a simple matter to see that $A=1$, $B=0$ and the result, after putting $r=1$, is (1.2). Nothing of interest is obtained by replacing any of the cosines by sines because in all these cases $A=B=0$.

Next consider the case $n=2$. We shall use the usual notation x, y instead of x_1, x_2 etc. If we take $\Phi(\underline{x})=\cos ax\cos by$, where $a^2+b^2=\lambda^2$ and $H_m(\underline{v})=\cos m\theta$ ($m=0,1,2,\cdots$), we get

$$\frac{1}{2\pi}\int_0^{2\pi}\cos ax\cos by\cos m\theta\,d\theta=r^mA\mathscr{J}_m(\lambda r),\qquad r^2=x^2+y^2,$$

because one sees at once that $B=0$. Writing $\cos ax\cos by$ as a sum of two cosines, and then using elementary analysis, we derive the value of A. After putting $r=1$, which we may do without loss of generality, the result reads

$$(3.1)\qquad \frac{1}{2\pi}\int_0^{2\pi}\cos ax\cos by\cos m\theta\,d\theta=\begin{cases}(-1)^{m/2}J_m(\lambda)\cos m\gamma, & m\ \text{even},\\ 0, & m\ \text{odd},\end{cases}$$

where $\cos\gamma=a/\lambda$, $\sin\gamma=b/\lambda$.

In the same way we can find the following formulas:

$$(3.2)\qquad \frac{1}{2\pi}\int_0^{2\pi}\sin ax\sin by\sin m\theta\,d\theta=\begin{cases}(-1)^{m/2}J_m(\lambda)\sin m\gamma, & m\ \text{even},\\ 0, & m\ \text{odd},\end{cases}$$

$$(3.3)\qquad \frac{1}{2\pi}\int_0^{2\pi}\cos ax\sin by\sin m\theta\,d\theta=\begin{cases}0, & m\ \text{even},\\ (-1)^{(m-1)/2}J_m(\lambda)\sin m\gamma, & m\ \text{odd},\end{cases}$$

$$(3.4)\qquad \frac{1}{2\pi}\int_0^{2\pi}\sin ax\cos by\cos m\theta\,d\theta=\begin{cases}0, & m\ \text{even},\\ (-1)^{(m-1)/2}J_m(\lambda)\cos m\gamma, & m\ \text{odd},\end{cases}$$

in each of which γ has the same meaning as above. By symmetry considerations the other four integrals of this kind are each seen to be zero for all m.

Naturally as n gets larger, the coefficients A and B become more tedious to calculate. Therefore we shall merely quote one example from E^3: If $n=3$, $\Phi(\underline{x})=\cos\lambda z$, $H_m(\underline{y})=P_m(\cos\theta)$, then we get, on setting $r=1$,

$$(3.5)\quad \frac{1}{S(1)}\int_{S(1)}\cos\lambda z P_m(\cos\theta)\sigma_1=\begin{cases}(-1)^{m/2}\sqrt{\dfrac{\pi}{2\lambda}}\,J_{m+(1/2)}(\lambda), & m \text{ even},\\[2ex] 0, & m \text{ odd}.\end{cases}$$

To illustrate the derivation of inequalities from results of this type, let us take (3.1) for example. If $w(\theta)$ denotes a linear combination of the functions $\cos m\theta (m=0,1,2,\cdots)$ such that $w(\theta)\geqq 0$ in $0\leqq\theta<2\pi$, let us write the result which follows from (3.1) as

$$(3.6)\quad \frac{1}{2\pi}\int_0^{2\pi}\cos ax\cos by\, w(\theta)\,d\theta=Q(a,b).$$

Then if $\varepsilon_1,\varepsilon_2=\pm1$, we have

$$(3.7)\quad \frac{1}{2\pi}\int_0^{2\pi}(1+\varepsilon_1\cos ax)(1+\varepsilon_2\cos by)w(\theta)\,d\theta\geqq 0.$$

Let $w(\theta)$ be normalized so that $\int_0^{2\pi}w(\theta)\,d\theta=2\pi$. Then multiplying out in (3.7) and using (3.6) we get

$$1+\varepsilon_1 Q(a,0)+\varepsilon_2 Q(0,b)+\varepsilon_1\varepsilon_2 Q(a,b)\geqq 0.$$

Making suitable choices for ε_1 and ε_2, we obtain at once

$$1+Q(a,b)\geqq|Q(a,0)+Q(0,b)|,$$
$$1-Q(a,b)\geqq|Q(a,0)-Q(0,b)|.$$

For example, if we take $w(\theta)=1+\eta\cos 2\theta$, where $-1\leqq\eta\leqq 1$, we obtain these two inequalities with

$$Q(a,b)\equiv J_0\left(\sqrt{a^2+b^2}\right)-\eta\frac{a^2-b^2}{a^2+b^2}J_2\left(\sqrt{a^2+b^2}\right).$$

The former of these can be regarded as a generalization of Grünbaum's inequality (1.1) because it reduces to this if we take $\eta=0$ and omit the modulus signs.

Again if we take $w(\theta)=1+\cos\theta$ and use (3.4) in this way, we obtain

$$1+\frac{a}{\sqrt{a^2+b^2}}J_1\left(\sqrt{a^2+b^2}\right)\geqq|J_1(a)+J_0(b)|,$$

and

$$1-\frac{a}{\sqrt{a^2+b^2}}J_1\left(\sqrt{a^2+b^2}\right)\geqq|J_1(a)-J_0(b)|.$$

Many interesting inequalities can be found in this fashion, but we shall not enlarge on the number of examples already given except to illustrate the following remark. In deducing examples from the representation (2.7), we have so far

assumed that the function Φ and the parameter λ are real, but there is, of course, no need for this. For example, the representation

$$J_0(\lambda)=\frac{1}{2\pi}\int_0^{2\pi}\cos ax\cos by\,d\theta$$

is valid for all (real or complex) values of a, b, λ satisfying $a^2+b^2=\lambda^2$. With α, β, γ real, let us take $a=i\alpha$, $b=i\beta$, $\lambda=i\gamma$. Thus

$$I_0(\gamma)=\frac{1}{2\pi}\int_0^{2\pi}\cosh\alpha x\cosh\beta y\,d\theta, \qquad \alpha^2+\beta^2=\gamma^2. \tag{3.8}$$

Observing that $(\cosh\alpha x+\varepsilon_1)(\cosh\beta y+\varepsilon_2)\geqq 0$ if $\varepsilon_1, \varepsilon_2=\pm 1$, we deduce from (3.8) that

$$I_0(\gamma)+1\geqq|I_0(\alpha)+I_0(\beta)|, \qquad \alpha^2+\beta^2=\gamma^2,$$
$$I_0(\gamma)-1\geqq|I_0(\alpha)-I_0(\beta)|, \qquad \alpha^2+\beta^2=\gamma^2.$$

Acknowledgment. The author wishes to thank R. G. Buschman for several helpful conversations during the preparation of this paper.

REFERENCES

[1] A. McD. Mercer, *Grünbaum's inequality for Bessel functions and its extensions*, this Journal, 6 (1975), pp. 1021–1023.

[2] ———, *On certain functional identities in E^N*, Canad. J. Math., 23 (1971), pp. 315–324.

[3] H. Flanders, *Differential forms with applications to the physical sciences*, Academic Press, New York, 1963.

SIAM J. MATH. ANAL.
Vol. 8, No. 3, May 1977

AN EIGENVALUE ESTIMATION METHOD OF WEINBERGER AND WEINSTEIN'S INTERMEDIATE PROBLEMS*

DAVID W. FOX AND JAMES T. STADTER†

Abstract. A method of H. F. Weinberger for calculating lower bounds to eigenvalues of semi-bounded self-adjoint operators is given a new formulation in an operator setting. This analysis makes clear the relation of this method to the Weinstein–Aronszajn intermediate problems and completes the study of the method initiated by Bazley and Fox. The operators that are developed here are useful in showing the monotone properties of the method in a direct and natural way.

Introduction. This article is about a method for calculating lower bounds to eigenvalues of semi-bounded self-adjoint operators. The method, due to H. F. Weinberger [6], [7], is given a new formulation in an operator setting that makes evident its relation to the Weinstein–Aronszajn intermediate problems [1], [5], [8], [9]. This represents a completion of an analysis of the Weinberger method started by Bazley and Fox in [2]. The operators that we develop here are useful in other ways. They make the demonstrations of the monotone properties of the method much more direct; the clarifications that they offer will, we hope, encourage the application of the method.

In § 1 we develop the operators from scratch. The subspaces that enter into the construction are set up carefully, then as a first step a base operator in the sense of intermediate operators is constructed. After that the difference between the base operator and the given operator is used to construct intermediate operators, and the resolution of the spectral problems of the intermediate operators is worked out. This section is completed by showing how the method can be used when only enough information is given to construct operators on a subspace.

Section 2 shows that although the base operators constructed in the method have considerable freedom, this freedom is without consequence for the intermediate operators. The resulting operators depend only on the subspaces used and not on the intervening construction. This demonstration casts the operators in a form from which it is easy to see that the operators have the quadratic form on which Weinberger's method is based. The section concludes by showing the relationship of the matrix problems for our operators on their reducing spaces with the matrix eigenvalue problems given by Weinberger in his formulation.

Section 3 gives new demonstrations of the monotone properties of the method based on the operator formulation.

1. Development of the method. The setting of the method is standard. We suppose that A is a self-adjoint operator in a separable complex Hilbert space $\mathfrak{H}$ with its inner product designated by $(\cdot, \cdot\,\cdot)$. The domain of A is denoted $\mathfrak{D}$, and A is assumed to be bounded below and to have the lowest part of its spectrum made up of eigenvalues λ_ν of finite multiplicity lying below the first limit point λ_* in its spectrum. We use the customary ordering of the eigenvalues accounting for

* Received by the editors December 23, 1975, and in revised form May 27, 1976.

† 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 N00017-72-4401.

multiplicity:

$$\lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_*.$$

The object of the method is to determine lower bounds for these lowest eigenvalues of A, i.e., to find real numbers μ_ν that can be computed precisely and that satisfy $\mu_\nu \leqq \lambda_\nu$ for $\nu = 1, 2, \cdots$.

The minimum information that seems to be necessary to develop a general procedure is that there be known a finite-dimensional space $\mathfrak{P}$ and a number ρ such that

$$(Au, u) \geqq \rho(u, u) \quad \forall u \in \mathfrak{D} \cap \mathfrak{P}^\perp. \tag{1.1}$$

Clearly ρ and $\mathfrak{P}$ must be derived from some special a priori knowledge about A. By itself (1.1) implies only that $\lambda_{n+1} \geqq \rho$, where $n = \dim \mathfrak{P}$; however by using the values of A on additional vectors further information is available. It is this information along with that given by (1.1) that is organized into a lower bound method.

1.1. Setting up the subspaces. We suppose that $\mathfrak{P}$ and ρ are fixed and choose an m-dimensional subspace $\mathfrak{Q}$ of $\mathfrak{D}$. Other than being in $\mathfrak{D}$, $\mathfrak{Q}$ can be completely arbitrary. An important property of $\mathfrak{Q}$ is its rank with $\mathfrak{P}$, defined by

$$r = \operatorname{rank}(\mathfrak{P}, \mathfrak{Q}) = \operatorname{rank}\{(p_i, q_j)\},$$

where $\{p_i\}$ and $\{q_j\}$ are any bases for $\mathfrak{P}$ and $\mathfrak{Q}$, respectively. Clearly $r \leqq m, n$. We make the following decompositions of the suspaces $\mathfrak{Q}$ and $\mathfrak{P}$:

$$\mathfrak{Q} = \mathfrak{Q}_1 \vee \mathfrak{Q}_2, \quad \text{where } \mathfrak{Q}_2 = \mathfrak{Q} \cap \mathfrak{P}^\perp, \tag{1.2}$$

and

$$\mathfrak{P} = \mathfrak{P}_1 \oplus \mathfrak{P}_2, \quad \text{where } \mathfrak{P}_2 = \mathfrak{P} \cap \mathfrak{Q}^\perp. \tag{1.3}$$

The symbols $\vee$ and $\oplus$ denote the linear span and orthogonal sum, respectively. The decomposition (1.2) is not unique, and it is to be understood in the following way: the second of (1.2) defines $\mathfrak{Q}_2$ uniquely; then the manifold $\mathfrak{Q}_1$ can be taken to be *any* fixed r-dimensional subspace of $\mathfrak{Q}$ that, together with $\mathfrak{Q}_2$, spans $\mathfrak{Q}$. The partition of $\mathfrak{P}$ given by (1.3) is unique.

If $\mathfrak{P}_2$ is not void, i.e., if $r = \operatorname{rank}(\mathfrak{P}, \mathfrak{Q}) < \dim \mathfrak{P} = n$, then the operator construction that we are about to outline will take place in $\mathfrak{P}_2^\perp$, a subspace of $\mathfrak{H}$ of deficiency $n - r$, and the method gives lower bounds to eigenvalues of A starting with the $(n - r + 1)$st. To avoid that complication, we assume for the time being that $r = n$ so that $\mathfrak{P}_1 = \mathfrak{P}$. A little later this restriction will be lifted.

1.2. Construction of a base operator. As a first step we construct a *base operator* A_0 depending on the given $\mathfrak{P}$ and ρ and on the values of A on $\mathfrak{Q}_1$. For A_0 to be a base operator, it must be smaller than A, and it must be *explicitly resolvable*, i.e., its eigenvalues and eigenvectors have to be determinable to any desired precision.[1]

[1] Here the requirement that A_0 be explicitly resolvable does not enter in a practical way, since we shall not need to determine its eigenvalues or eigenvectors.

To carry out this construction we use the fact that the bounded operator Q_1 is completely determined by the requirements on the ranges of Q_1 and $(I-Q_1)$

$$\mathfrak{R}(Q_1)\subset\mathfrak{Q}_1, \qquad \mathfrak{R}(I-Q_1)\perp\mathfrak{P}_1. \tag{1.4}$$

These imply that Q_1 and its adjoint Q_1^* satisfy

$$\mathfrak{R}(Q_1)=\mathfrak{Q}_1, \quad \mathfrak{R}(Q_1^*)=\mathfrak{P}_1, \quad \text{and} \quad Q_1^2=Q_1.$$

From (1.4) and (1.1) it follows[2] that

$$([A-\rho][I-Q_1]u,[I-Q_1]u)\geqq 0 \quad \forall u\in\mathfrak{D}. \tag{1.5}$$

This means that the symmetric operator A_1 defined by

$$A_1=(I-Q_1)^*(A-\rho)(I-Q_1) \tag{1.6}$$

on $\mathfrak{D}$ is nonnegative. Now we define A_0 on $\mathfrak{D}$ by

$$A_0=A-A_1. \tag{1.7}$$

A short calculation gives

$$A_0=\rho I+(A-\rho)Q_1+Q_1^*(A-\rho)-Q_1^*(A-\rho)Q_1. \tag{1.8}$$

Since $\mathfrak{R}(Q_1)\subset\mathfrak{D}$, it follows that $Q_1^*(A-\rho)$, as well as $(A-\rho)Q_1$ and $Q_1^*(A-\rho)Q_1$, are bounded on $\mathfrak{D}$. Consequently we can extend A_0 to all of $\mathfrak{H}$ by continuity. It follows from (1.5) and (1.6) that A_0 is self-adjoint on $\mathfrak{H}$ and satisfies $A_0\leqq A$.

From (1.8) it is clear that $A_0-\rho I$ has its range in the finite-dimensional subspace $\mathfrak{M}_0$ given by

$$\mathfrak{M}_0=\mathfrak{P}_1\vee(A-\rho)\mathfrak{Q}_1, \tag{1.9}$$

and thus, since A_0 is symmetric, $\mathfrak{M}_0$ is a reducing space for A_0. On $\mathfrak{M}_0^\perp$ we have $A_0=\rho I$, since $\mathfrak{M}_0^\perp$ is contained in the null spaces of Q_1 and $Q_1^*(A-\rho)$. Thus A_0 is explicitly resolvable by diagonalizing a representing matrix for it on $\mathfrak{M}_0$.

It is appropriate here to recall that the manifold $\mathfrak{Q}_1$ is not completely fixed by (1.2) and that each distinct selection of $\mathfrak{Q}_1$ will in general result in a different base operator A_0. Nevertheless the intermediate operator we construct in the next step turns out to be independent of the choice of $\mathfrak{Q}_1$.

1.3. Construction of intermediate operators. With a base problem in hand, the remaining part $\mathfrak{Q}_2$ of $\mathfrak{Q}$ can be used to increase the base operator by the intermediate operator method.[3] Of course, for $\mathfrak{Q}_2$ to be nonvoid we suppose $r<m=\dim\mathfrak{Q}$.

The construction is suggested immediately when we write

$$A=A_0+A_1, \tag{1.10}$$

since A_1 is nonnegative and A_0 is explicitly resolvable. Let Q_2 be a projection on $\mathfrak{Q}_2$ orthogonal with respect to the quadratic form of A_1. From the intermediate

[2] Recall that for the moment $\mathfrak{P}=\mathfrak{P}_1$.

[3] See, for example, [3], [4], [5], or [9] for descriptions of this method.

operator construction[4] it follows that the operator[5] A_1Q_2 is bounded, symmetric, and of finite rank on $\mathfrak{D}$ and can be extended to $\mathfrak{H}$ by continuity. Further, it satisfies

$$0 \leqq A_1Q_2 \leqq A_1. \tag{1.11}$$

Let A_2 be the operator defined by

$$A_2 = A_0 + A_1Q_2. \tag{1.12}$$

The operator A_2 is bounded, and by (1.11) it is intermediate between A_0 and A, i.e.,

$$A_0 \leqq A_2 \leqq A, \tag{1.13}$$

and consequently the eigenvalues of these operators satisfy the parallel inequalities

$$\mu_\nu^0 \leqq \mu_\nu^2 \leqq \lambda_\nu, \qquad \nu = 1, 2, \cdots. \tag{1.14}$$

The operator A_2, as we shall see a little later, is exactly that of Weinberger when $\mathfrak{P} = \mathfrak{P}_1$.

Before we go further, it is appropriate to record some additional properties of Q_2. Since Q_2 is an orthogonal projection on $\mathfrak{Q}_2$ in the quadratic form generated by A_1, it satisfies $Q_2^2 = Q_2$ and $A_1Q_2 = Q_2^*A_1 = Q_2^*A_1Q_2$, where Q_2^* is the adjoint of Q_2 with respect to the inner product of $\mathfrak{H}$. An equivalent definition of Q_2 on $\mathfrak{D}$ is given by

$$\mathfrak{N}(Q_2) \subset \mathfrak{Q}_2, \qquad \mathfrak{N}(I - Q_2) \subset \{A_1\mathfrak{Q}_2\}^\perp. \tag{1.15}$$

Since $\mathfrak{Q}_2 \subset \mathfrak{P}^\perp \subset \mathfrak{N}(Q_1)$, where $\mathfrak{N}(Q_1)$ is the null space of Q_1, the right side of the second of (1.15) can be written $\{(I - Q_1)^*(A - \rho)\mathfrak{Q}_2\}^\perp$. It follows immediately from this that $\mathfrak{N}(Q_2^*) = (I - Q_1)^*(A - \rho)\mathfrak{Q}_2$. Note that since $Q_1\mathfrak{Q}_2 = 0$, the vanishing of A_1 on a vector u in $\mathfrak{Q}_2$ is equivalent to $(Au, u) = \rho(u, u)$, which is independent of $\mathfrak{Q}_1$.

1.4. Resolution of the spectral problem for A_2. The operator A_2 is also resolvable by the diagonalization of a symmetric representing matrix on a finite-dimensional reducing space. In fact, the subspace $\mathfrak{M}_2$ defined by

$$\mathfrak{M}_2 = \mathfrak{P}_1 \vee (A - \rho)\mathfrak{Q} \tag{1.16}$$

reduces A_2. First observe that A_2 can be written in the form

$$A_2 = \rho I + Q_1^*(A - \rho) + (A - \rho)Q_1 - Q_1^*(A - \rho)Q_1 + Q_2^*A_1. \tag{1.17}$$

From this it follows that $\mathfrak{N}(A_2 - \rho I) = \mathfrak{N}(Q_1^*) \vee (A - \rho)\mathfrak{Q}_1 \vee \mathfrak{N}(Q_2^*) \subset \mathfrak{M}_2$. Since A_2 is symmetric, $\mathfrak{M}_2$ is reducing for A_2. Further, on $\mathfrak{M}_2^\perp$ the projection Q_2 vanishes, and since $\mathfrak{M}_2^\perp \subset \mathfrak{M}_0^\perp$ we have $A_2 = A_0 = \rho I$ on $\mathfrak{M}_2^\perp$. Thus the eigenvalue problem for A_2 in $\mathfrak{H}$ can be resolved by determining the eigenvalues of a representing matrix on $\mathfrak{M}_2$.

[4] See, for example, [4, p. 434] and [9, p. 80].

[5] When A_1 is not positive definite on $\mathfrak{Q}_2$, the projection Q_2 is undetermined to the extent of an operator into the subspace of $\mathfrak{Q}_2$ on which A_1 vanishes. Since the construction uses Q_2 only in the expression A_1Q_2, the construction is nonetheless uniquely determined.

1.5. Intermediate operators on a subspace. Now, we take care of the more general circumstance in which $\mathfrak{P}=\mathfrak{P}_1\oplus\mathfrak{P}_2$ with $\mathfrak{P}_2\neq 0$. In this case (1.5) is replaced by (1.1) in the form

$$([A-\rho][I-Q_1]u,[I-Q_1]u)\geqq 0 \qquad \forall u\in\mathfrak{D}\cap\mathfrak{P}_2^{\perp}. \tag{1.18}$$

The natural way to deal with this situation is to turn our attention to the *part* $\hat{A}$ of A in $\mathfrak{P}_2^{\perp}$. Thus we apply the method to $\hat{A}$, the self-adjoint operator in $\hat{\mathfrak{H}}=\mathfrak{P}_2^{\perp}$ defined by

$$\hat{A}=(I-P_2)A(I-P_2) \tag{1.19}$$

on $\hat{\mathfrak{D}}=\mathfrak{D}\cap\mathfrak{P}_2^{\perp}$, where P_2 is the orthogonal projection in $\mathfrak{H}$ on $\mathfrak{P}_2$. The construction we have given so far is valid when A is replaced by $\hat{A}$, $\mathfrak{D}$ by $\hat{\mathfrak{D}}$, and the λ_ν by the eigenvalues $\hat{\lambda}_\nu$ of $\hat{A}$. This is summarized by

$$\hat{A}_1=(I-Q_1)^*(\hat{A}-\rho)(I-Q_1), \tag{1.20}$$

$$\hat{A}=\hat{A}_0+\hat{A}_1, \tag{1.21}$$

$$\hat{A}_2=\hat{A}_0+\hat{A}_1Q_2. \tag{1.22}$$

Observe that everything in (1.20)–(1.22) makes good sense. Indeed, $\mathfrak{R}(Q_1)=\mathfrak{Q}_1$ and $\mathfrak{R}(Q_1^*)=\mathfrak{P}_1$ are both in $\hat{\mathfrak{H}}$ since $\mathfrak{P}_2^{\perp}=\mathfrak{P}^{\perp}\vee\mathfrak{Q}$ and $\mathfrak{P}\ominus\mathfrak{P}_2=\mathfrak{P}_1$ from (1.3). Further, since $\mathfrak{Q}_2\subset\mathfrak{P}^{\perp}\subset\hat{\mathfrak{H}}$, the definition of $\mathfrak{Q}_2$ already given holds in $\hat{\mathfrak{H}}$ when A is replaced by $\hat{A}$. The reducing space $\hat{\mathfrak{M}}_2$ for $\hat{A}_2$ has the same form as that for A_2 with A replaced by $\hat{A}$, i.e.,

$$\hat{\mathfrak{M}}_2=\mathfrak{P}_1\vee(\hat{A}-\rho)\mathfrak{Q}. \tag{1.23}$$

Since $\hat{\mathfrak{H}}$ is a subspace of $\mathfrak{H}$ of deficiency $n-r$, the eigenvalues of $\hat{A}$ and A satisfy the inequalities[6]

$$\lambda_\nu\leqq\hat{\lambda}_\nu\leqq\lambda_{\nu+n-r}, \qquad \nu=1,2,\cdots. \tag{1.24}$$

From this it follows that the eigenvalues $\hat{\mu}_\nu$ of $\hat{A}_2$ give lower bounds to higher eigenvalues of A, i.e.,

$$\hat{\mu}_\nu\leqq\hat{\lambda}_\nu\leqq\lambda_{\nu+n-r}, \qquad \nu=1,2,\cdots. \tag{1.25}$$

2. Equivalence with Weinberger's construction. In this section we show that the operator $\hat{A}_2$ is independent of the selection of the subspace $\mathfrak{Q}_1$ in (1.2). In doing this we express $\hat{A}_2$ in a form that shows it to be the operator corresponding to the quadratic form of Weinberger's construction.

2.1. Independence of $\mathfrak{Q}_1$. Although the choice of the subspace $\mathfrak{Q}_1$ determines Q_1 and subsequently Q_2, the intermediate operator $\hat{A}_2$ does not depend at all on this choice. We show this by demonstrating that $\hat{A}_2$ depends on $Q=Q_1+Q_2$ only and that Q is, in fact, determined by $\mathfrak{Q}$ and is independent of the choice of $\mathfrak{Q}_1$.

We observe first that Q_1 and Q_2 annihilate each other, that is,

$$Q_1Q_2=0=Q_2Q_1. \tag{2.1}$$

[6] See, for example, [5, p. 71].

The first follows from $\mathfrak{R}(Q_2) \subset \mathfrak{P}_1^{\perp} \subset \mathfrak{R}(Q_1)$ and the second from $\mathfrak{R}(Q_1) \subset \{(I-Q_1)^*(\hat{A}-\rho)\mathfrak{Q}_2\}^{\perp} \subset \mathfrak{R}(Q_2)$. Now using (2.1) we rewrite $\hat{A}_2$ as follows:

$$\begin{aligned} \hat{A}_2 &= \hat{A}_0 + \hat{A}_1 Q_2 = \hat{A} - (I-Q_2)^* \hat{A}_1 (I-Q_2) \\ &= \hat{A} - (I-Q_1-Q_2)^*(\hat{A}-\rho)(I-Q_1-Q_2). \end{aligned} \tag{2.2}$$

Now put $Q = Q_1 + Q_2$ and write $\hat{A}_2$ in the form

$$\hat{A}_2 = \hat{A} - (I-Q)^*(\hat{A}-\rho)(I-Q), \tag{2.3}$$

where

$$\mathfrak{R}(Q) \subset \mathfrak{Q}, \qquad \mathfrak{R}(I-Q) \subset \{\mathfrak{P} \vee (\hat{A}-\rho)\mathfrak{Q}_2\}^{\perp} \quad \text{in } \hat{\mathfrak{H}}. \tag{2.4}$$

The first of (2.4) is immediate from the definition of Q. The second follows from

$$\mathfrak{R}(I-Q) = \mathfrak{R}[(I-Q_1)(I-Q_2)] \subset \mathfrak{R}(I-Q_1) \subset \mathfrak{P}_1^{\perp},$$

and

$$\mathfrak{R}(I-Q) = \mathfrak{R}[(I-Q_2)(I-Q_1)] \subset \mathfrak{R}(I-Q_2) \subset \{(I-Q_1)^*(\hat{A}-\rho)\mathfrak{Q}_2\}^{\perp},$$

so that

$$\mathfrak{R}(I-Q) \subset \mathfrak{P}_1^{\perp} \cap \{(\hat{A}-\rho)\mathfrak{Q}_2\}^{\perp} = \{\mathfrak{P} \vee (A-\rho)\mathfrak{Q}_2\}^{\perp} \quad \text{in } \hat{\mathfrak{H}}.$$

On the other hand (2.4) determines Q up to an operator with range in $\mathfrak{Q}_2 \cap \mathfrak{R}(\hat{A}-\rho)$. This lack of uniqueness does not cause any indeterminacy in the operator $(I-Q)^*(\hat{A}-\rho)(I-Q)$; in fact, it corresponds to the nonuniqueness in Q_2 discussed earlier.

2.2. The quadratic form of $\hat{A}_2$. From (2.3) all that is needed is a brief calculation to obtain the quadratic form of $\hat{A}_2$ in the expression which is the starting point for Weinberger's analysis. Indeed, from (2.3) we obtain

$$\hat{A}_2 = Q^*\hat{A}Q + Q^*\hat{A}(I-Q) + (I-Q)^*\hat{A}Q + \rho(I-Q)^*(I-Q) \quad \text{on } \hat{\mathfrak{H}},$$

and from this the quadratic form

$$\begin{aligned} (\hat{A}_2 u, u) = (AQu, Qu) + ([I-Q]u, AQu) + (AQu, [I-Q]u) \\ + \rho([I-Q]u, (I-Q)u) \end{aligned}$$

for $u \in \hat{\mathfrak{H}}$.

For each u in $\hat{\mathfrak{H}}$ the projection Q decomposes u according to

$$u = Qu + (I-Q)u = q + v,$$

with $q \in \mathfrak{Q}$, $v \perp \mathfrak{P}$, and $v \perp (A-\rho)\mathfrak{Q}_2$. In terms of q and v the quadratic form of $\hat{A}_2$ is

$$(Aq, q) + (v, Aq) + (Aq, v) + \rho(v, v) \quad \forall u \in \hat{\mathfrak{H}},$$

which is, in fact, the expression on which Weinberger's analysis is based.

2.3. Matrices for the bounds. Here we give a transformation of the matrix eigenvalue problems for $\hat{A}_2$ on $\hat{\mathfrak{M}}_2$ into the form obtained by Weinberger. This demonstrates in an independent way the equivalence of the methods.

Recall that $\hat{\mathfrak{M}}_2 = \mathfrak{P}_1 \vee (\hat{A}-\rho)\mathfrak{Q}$ is the reducing space for $\hat{A}_2$. The space $\hat{\mathfrak{M}}_2$ may not have the full dimension $r+m$ since $\mathfrak{Q}$ may have a subspace on which $\hat{A}-\rho$ vanishes, and in addition $\mathfrak{P}_1$ and $(\hat{A}-\rho)\mathfrak{Q}$ may not be linearly independent. Starting from a matrix representation for $\hat{A}_2$ given by a linearly independent basis for $\hat{\mathfrak{M}}_2$, we transform the system of linear equations into that given by Weinberger. Two steps are involved: the first is the transformation to a system of equations in terms of a natural (but possibly dependent) spanning set for $\hat{\mathfrak{M}}_2$; the second is the expansion of the system of equations to order $m+n$.

We introduce special bases for $\mathfrak{P}$ and $\mathfrak{Q}$ to simplify the calculations. Let $\{p_1, p_2, \cdots, p_n\}$ be a basis for $\mathfrak{P}$ such that the first r span $\mathfrak{P}_1$ and the last $n-r$ span $\mathfrak{P}_2$. Further, let $\{q_1, q_2, \cdots, q_m\}$ be a basis for $\mathfrak{Q}$ such that the first r span $\mathfrak{Q}_1$ and the last $m-r$ span $\mathfrak{Q}_2$. If $\hat{A}-\rho$ vanishes on a subspace of $\mathfrak{Q}_2$, we suppose that this subspace is spanned by the last $m-m'$ vectors in the basis for $\mathfrak{Q}$; then $\hat{A}-\rho$ will be positive definite on the span of $\{q_{r+1}, q_{r+2}, \cdots, q_{m'}\}$.

It follows from (1.4) that Q_1 has the explicit representation

$$Q_1 = \sum_{i,j=1}^{r} (\cdot, p_i) b_{ij}^1 q_j, \quad \text{where } \{b_{ij}^1\} = \{(q_k, p_l)\}^{-1}. \tag{2.5}$$

Similarly from (1.15) we have on $\hat{\mathfrak{S}}_2$

$$Q_2 = \sum_{i,j=1}^{m'-r} ([I-Q_1]\cdot, [A-\rho]q_{i+r}) b_{ij}^2 q_{j+r}, \tag{2.6}$$

$$\text{where } \{b_{ij}^2\} = \{([A-\rho]q_{i+r}, q_{j+r})\}^{-1}.$$

Suppose $\{t_1, t_2, \cdots, t_s\}$ is a basis for $\hat{\mathfrak{M}}_2$. In terms of this basis the eigenvalue problem for $\hat{A}_2$ on $\hat{\mathfrak{M}}_2$ is

$$\sum_{i=1}^{s} x_i^1 \{(\hat{A}_2 t_i, t_j) - \hat{\mu}(t_i, t_j)\} = 0, \qquad j = 1, 2, \cdots, s. \tag{2.7}$$

A natural spanning set for $\hat{\mathfrak{M}}_2$ is $\{v_1, v_2, \cdots, v_{m'+r}\}$ given by

$$v_i = \begin{cases} (\hat{A}-\rho)q_i, & i = 1, 2, \cdots, m', \\ p_{i-m'}, & i = m'+1, m'+2, \cdots, m'+r. \end{cases}$$

Let Γ be a nonsingular matrix that expresses the t_i in terms of the v_i and that states $m'+r-s$ independent relations of linear dependence among the v_i, i.e.,

$$\sum_{j=1}^{m'+r} \Gamma_{ij} v_j = \begin{cases} t_i, & i = 1, 2, \cdots, s, \\ 0, & i = s+1, s+2, \cdots, m'+r. \end{cases}$$

Now augment the equations (2.7) in the trivial way:

$$\{x^1, x^2\} \left(\begin{array}{c|c} \{(\hat{A}_2 t_i, t_j) - \hat{\mu}(t_i, t_j)\} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \hline 0 & \end{array} \right) = 0,$$

where x^2 is arbitrary. In terms of Γ these are equivalent to

$$x\Gamma\{(\hat{A}_2 v_i, v_j) - \hat{\mu}(v_i, v_j)\}\Gamma^* = 0, \tag{2.8}$$

where $x=\{x^1, x^2\}$ and Γ^* is the transpose conjugate of Γ. A tedious calculation using (2.5) and (2.6) allows us to write (2.8) in the form

$$x\Gamma[\hat{F}\hat{G}^{-1}\hat{F}-(\hat{\mu}-\rho)\hat{F}]\Gamma^*=0, \tag{2.9}$$

where $\hat{F}$ and $\hat{G}$ are the square matrices of order $m'+r$ given by

$$\hat{F}=(v_i, v_j)=\left(\begin{array}{c|c} ([\hat{A}-\rho]q_i, [A-\rho]q_j) & ([\hat{A}-\rho]q_i, p_{j-m'}) \\ \hline (p_{i-m'}, [\hat{A}-\rho]q_j) & (p_{i-m'}, p_{j-m'}) \end{array}\right) \tag{2.10}$$

and

$$\hat{G}=\left(\begin{array}{c|c} ([\hat{A}-\rho]q_i, q_j) & (q_i, p_{j-m'}) \\ \hline (p_{i-m'}, q_j) & 0 \end{array}\right). \tag{2.11}$$

Since from the definitions of $\hat{F}$ and Γ we have

$$\Gamma\hat{F}\Gamma^*=\left(\begin{array}{c|c} \{(t_i, t_j)\} & 0 \\ \hline 0 & 0 \end{array}\right),$$

(2.9) can be put in the form

$$\{y^1, 0\}\Gamma^{*-1}\hat{G}^{-1}[\hat{F}-(\hat{\mu}-\rho)\hat{G}]=0,$$

where $y^1=x^1\{(t_i, t_j)\}$. This leads us to consider the equation

$$\hat{z}[\hat{F}-(\hat{\mu}-\rho)\hat{G}]=0. \tag{2.12}$$

Using the fact that $\hat{G}$ is nonsingular it follows that there is a one-to-one correspondence of solutions of (2.12) with those of (2.7) for each $\hat{\mu}$ different from ρ.

Assume for the moment that $m'=m$. Then the transformation of the matrix problem is completed by eliminating $\hat{A}$ from the matrices. This necessitates the expansion of the order of the system to $m+n$. From the definition of $\hat{A}$ and the choice of the basis vectors p_i and q_i we have

$$(\hat{A}-\rho)q_i=\begin{cases}(A-\rho)q_i+\sum_{k=1}^{n-r}c_{ik}p_{k+r}, & i=1, 2, \cdots, r, \\ (A-\rho)q_i, & i=r+1, r+2, \cdots, m,\end{cases}$$

where the constants c_{ik} satisfy

$$([A-\rho]q_i, p_{l+r})+\sum_{k=1}^{n-r}c_{ik}(p_{k+r}, p_{l+r})=0, \qquad i=1, 2, \cdots, r; \quad l=1, 2, \cdots, n-r.$$

Using this, the matrix equation (2.12) is transformed to the equivalent equation

$$z[F-(\hat{\mu}-\rho)G]=0 \tag{2.13}$$

where

$$z_j=\hat{z}_j, \qquad j=1, 2, \cdots, m+r,$$

and

$$z_j = \sum_{i=1}^{r} \hat{z}_i c_{ij-m-r}, \quad j = m+r+1, m+r+2, \cdots, m+n,$$

and F and G are the square matrices of order $m+n$ given by

$$F = \left(\begin{array}{c|c} ([A-\rho]q_i, [A-\rho]q_j) & ([A-\rho]q_i, p_{j-m}) \\ \hline (p_{i-m}, [A-\rho]q_j) & (p_{i-m}, p_{j-m}) \end{array} \right) \tag{2.14}$$

and

$$G = \left(\begin{array}{c|c} ([A-\rho]q_i, q_j) & (q_i, p_{j-m}) \\ \hline (p_{i-m}, q_j) & 0 \end{array} \right). \tag{2.15}$$

Clearly the expressions for F and G show that the equations (2.13) do not depend on the special bases for $\mathfrak{P}$ and $\mathfrak{Q}$ that we have used. Thus we see that the system (2.13) with F and G given by (2.14) and (2.15), which is precisely that given by Weinberger in [6] and [7], is equivalent to (2.7) for $\hat{A}_2$ on $\hat{\mathfrak{M}}_2$ for all eigenvalues $\hat{\mu} \neq \rho$.

When $m' < m$ the equations (2.12) can still be cast in the form (2.13)–(2.15), but at the cost of introducing an $m-m'$ manifold of solutions for each solution of (2.12). In fact, by augmenting $\hat{F}$ and $\hat{G}$ by $m-m'$ zero rows and columns corresponding to the q's on which $(\hat{A}-\rho)$ vanishes, and by adding $m-m'$ corresponding arbitrary components to $\hat{z}$, (2.12) is correct with $m = m'$. Then the transformation to (2.13)–(2.15) can be made as above. Of course the final F and G then have a common null space of dimension $m-m'$.

3. Monotone properties of the operators. In this section we show that the lower bounds obtained from this construction increase as the space $\mathfrak{Q}$ is enlarged, as $\mathfrak{P}$ is diminished, or as ρ is increased provided the inequality (1.1) continues to hold. These conclusions are in accord with the observation that expanding $\mathfrak{Q}$, contracting $\mathfrak{P}$, or increasing ρ gives more information about A; consequently the bounds should be better. In comparing the bounds obtained by enlarging $\mathfrak{Q}$ or restricting $\mathfrak{P}$ we use similar arguments. In each case the method gives rise to a new intermediate operator designated by $\hat{A}_2'$ on an enlarged subspace $\hat{\mathfrak{H}}'$. The analysis is carried out by constructing a new operator that is smaller than $\hat{A}_2'$ in $\hat{\mathfrak{H}}'$ and that has $\hat{A}_2$ as its part in $\hat{\mathfrak{H}}$.

3.1. Elementary comparisons for projections. The techniques of the demonstration use a number of facts about projections.

If $[\cdot, \cdot]$ is a positive quadratic form given by $[\cdot, \cdot] = (R\cdot, \cdot)$, R a densely defined symmetric operator on $\mathfrak{D}$, and $\mathfrak{S}$ is a finite-dimensional subspace in $\mathfrak{D}$, then we consider

$$\delta^2 = \min_{s \in \mathfrak{S}} [(u-s), (u-s)] = \min_{T} [(I-T)u, (I-T)u], \qquad u \in \mathfrak{D},$$

where the minimum on the right is over all linear operators T on $\mathfrak{D}$ with range in $\mathfrak{S}$. The minimizing T is a projection S on $\mathfrak{S}$ orthogonal with respect to $[\cdot, \cdot]$. From this characterization three easy inequalities follow:

1. If S is a projection on $\mathfrak{S}$ orthogonal with respect to $[\cdot,\cdot]$, and $\hat{S}$ is any bounded linear operator with range in $\mathfrak{S}$, then

$$R(I-S)=(I-S)^*R(I-S)\leqq(I-\hat{S})^*R(I-\hat{S}), \tag{3.1}$$

where the * indicates the adjoint in $\mathfrak{H}$.

2. If $\mathfrak{S}_1\subset\mathfrak{S}_2\subset\mathfrak{D}$ and S_1 and S_2 are projections on $\mathfrak{S}_1$ and $\mathfrak{S}_2$, respectively, then

$$R(I-S_1)\geqq R(I-S_2) \quad \text{and} \quad RS_1\leqq RS_2. \tag{3.2}$$

This is just Bessel's inequality.

3. If $R_1\leqq R_2$ and $\mathfrak{S}\subset\mathfrak{D}(R_2)$ and $[\cdot,\cdot]_1$ and $[\cdot,\cdot]_2$ are the quadratic forms constructed from R_1 and R_2, respectively, and if S_1 and S_2 are orthogonal projections on $\mathfrak{S}$ with respect to $[\cdot,\cdot]_1$ and $[\cdot,\cdot]_2$, respectively, then

$$R_1(I-S_1)\leqq R_2(I-S_2). \tag{3.3}$$

3.2. Enlargement of $\mathfrak{Q}$. We suppose that in place of the m-dimensional space $\mathfrak{Q}$ we use a *larger* m'-dimensional space $\mathfrak{Q}'$ that includes $\mathfrak{Q}$ to construct a new operator $\hat{A}_2'$. The space $\mathfrak{P}$ and the number ρ stay fixed. The notation will be the same as that used before but with a prime added. Note that $r\leqq r'=$ rank $(\mathfrak{P},\mathfrak{Q}')$. In parallel with (1.2) and (1.3) we write

$$\mathfrak{Q}'=\mathfrak{Q}_1'\vee\mathfrak{Q}_2', \qquad \mathfrak{Q}_2'=\mathfrak{Q}'\cap\mathfrak{P}^{\perp}, \tag{3.4}$$

and

$$\mathfrak{P}=\mathfrak{P}_1'\oplus\mathfrak{P}_2', \qquad \mathfrak{P}_2'=\mathfrak{P}\cap\mathfrak{Q}'^{\perp}. \tag{3.5}$$

Clearly $\mathfrak{Q}'\supset\mathfrak{Q}$ implies

$$\mathfrak{P}_2'\subset\mathfrak{P}_2 \quad \text{and} \quad \mathfrak{Q}_2\subset\mathfrak{Q}_2'. \tag{3.6}$$

Since $\mathfrak{P}_2'$ is a subspace of $\mathfrak{P}_2$, the operator $\hat{A}'$ is defined on $\hat{\mathfrak{H}}'=\mathfrak{P}_2'^{\perp}$, while $\hat{A}_2$ is defined on $\hat{\mathfrak{H}}=\mathfrak{P}_2^{\perp}$, which is a subspace of $\hat{\mathfrak{H}}'$ of deficiency $r'-r$. The operator $\hat{A}_2'$ is given by

$$\hat{A}_2'=\hat{A}_0'+\hat{A}_1'Q_2' \quad \text{on } \hat{\mathfrak{H}}_2'=\mathfrak{P}_2'^{\perp}, \tag{3.7}$$

where $\hat{A}_0'=\hat{A}'-\hat{A}_1'$ with $\hat{A}'$ the part of A in $\hat{\mathfrak{H}}'$ and $\hat{A}_1'=(I-Q_1')^*(\hat{A}'-\rho)(I-Q_1')$.

Define $\bar{Q}_2$ as the projection on $\mathfrak{Q}_2$ orthogonal with respect to the quadratic form of $\hat{A}_1'$ and let $\bar{A}_2$ be the bounded operator on $\hat{\mathfrak{H}}'$ given by

$$\bar{A}_2=\hat{A}_0'+\hat{A}_1'\bar{Q}_2. \tag{3.8}$$

Since $\mathfrak{Q}_2\subset\mathfrak{Q}_2'$, (3.2) implies $\hat{A}_1'\bar{Q}_2\leqq\hat{A}_1'Q_2'$, and consequently

$$\bar{A}_2\leqq\hat{A}_2' \quad \text{on } \hat{\mathfrak{H}}'. \tag{3.9}$$

Now we show that the part of $\bar{A}_2$ in $\hat{\mathfrak{H}}$ coincides with $\hat{A}_2$. The inclusions (3.6) allow us to write

$$\mathfrak{P}_1'=\mathfrak{P}_1\oplus\mathfrak{P}_1^+, \qquad \mathfrak{P}_1^+\subset\mathfrak{P}_2, \tag{3.10}$$

and to take $\mathfrak{Q}_1'$ in the form

$$\mathfrak{Q}_1' = \mathfrak{Q}_1 \vee \mathfrak{Q}_1^+, \qquad \mathfrak{Q}_1^+ \perp \mathfrak{P}_1. \tag{3.11}$$

With this choice of $\mathfrak{Q}_1'$ the operator Q_1' can be written in the form

$$Q_1' = Q_1 + Q_1^+,$$

where Q_1 is the operator used in constructing $\hat{A}_1$, and Q_1^+ satisfies

$$\mathfrak{R}(Q_1^+) \subset \mathfrak{Q}_1^+, \qquad \mathfrak{R}(I - Q_1^+) \perp \mathfrak{P}_1^+.$$

It follows that $\mathfrak{R}(Q_1^+) \subset \mathfrak{P}_1^{+\perp}$. A short calculation using this shows that the part of $\hat{A}_1'$ in $\hat{\mathfrak{H}}$ is $\hat{A}_1$ and the part of $\hat{A}_1'\bar{Q}_2$ is $\hat{A}_1 Q_2$, so that the part of $\bar{A}_2$ in $\hat{\mathfrak{H}}$ is exactly $\hat{A}_2$, as we wished to demonstrate. From this and (3.9) it is immediate that the eigenvalues of $\hat{A}_2$ and $\hat{A}_2'$ satisfy

$$\hat{\mu}_\nu \leqq \hat{\mu}'_{\nu+r'-r}, \qquad \nu = 1, 2, \cdots, \tag{3.12}$$

so that the enlargement of $\mathfrak{Q}$ to $\mathfrak{Q}'$ gives lower bounds to eigenvalues starting with the smaller eigenvalue $\lambda_{n-r'+1}$, and when $\hat{A}_2'$ and $\hat{A}_2$ both give lower bounds to the same eigenvalue, beginning with λ_{n-r+1}, those given by $\hat{A}_2'$ are at least as good as those from $\hat{A}_2$. This is summarized by

$$\begin{aligned} &\hat{\mu}'_\nu \leqq \lambda_{\nu+n-r'}, && \nu = 1, 2, \cdots, r'-r, \\ &\hat{\mu}_{\nu-r'+r} \leqq \hat{\mu}'_\nu \leqq \lambda_{\nu+n-r'}, && \nu = r'-r+1, r'-r+2, \cdots. \end{aligned} \tag{3.13}$$

3.3. Contraction of $\mathfrak{P}$. Now we compare the operator $\hat{A}_2'$ constructed using a *smaller* space $\mathfrak{P}'$ for which (1.1) is valid and using the same ρ and $\mathfrak{Q}$. Note that $r' = \text{rank } (\mathfrak{B}', \mathfrak{Q})$ satisfies $r' \leqq r$, and that $n' - r' \leqq n - r$, where $n' = \dim \mathfrak{B}'$. Here we use the notation

$$\mathfrak{Q} = \mathfrak{Q}_1' \vee \mathfrak{Q}_2', \quad \text{where} \quad \mathfrak{Q}_2' = \mathfrak{Q} \cap \mathfrak{P}'^{\perp} \tag{3.14}$$

and

$$\mathfrak{P}' = \mathfrak{P}_1' \oplus \mathfrak{P}_2', \quad \text{where} \quad \mathfrak{P}_2' = \mathfrak{P}' \cap \mathfrak{Q}^{\perp}. \tag{3.15}$$

Clearly

$$\mathfrak{Q}_2' \supset \mathfrak{Q}_2 \quad \text{and} \quad \mathfrak{P}_2' \subset \mathfrak{P}_2. \tag{3.16}$$

Since $\mathfrak{P}_2'$ is a subspace of $\mathfrak{P}_2$, the operator $\hat{A}_2'$ is defined on the space $\hat{\mathfrak{H}}' = \mathfrak{P}_2'^{\perp}$ while $\hat{A}_1$ is given on the smaller space $\hat{\mathfrak{H}} = \mathfrak{P}_2^{\perp}$. Since we have $\mathfrak{Q} = \mathfrak{Q}_1' \vee \mathfrak{Q}_2'$ and $\mathfrak{Q}_2' \supset \mathfrak{Q}_2$, we may write $\mathfrak{Q}_2' = \mathfrak{Q}_1^+ \vee \mathfrak{Q}_2$ and take

$$\mathfrak{Q}_1 = \mathfrak{Q}_1' \vee \mathfrak{Q}_1^+ \quad \text{and} \quad \mathfrak{Q}_1^+ \perp \mathfrak{P}_1'. \tag{3.17}$$

With this choice of $\mathfrak{Q}_1$ we can write

$$Q_1 = Q_1' + Q_1^+, \qquad \mathfrak{R}(Q_1^+) \subset \mathfrak{Q}_1^+.$$

From the second of (3.17) and the definition of Q_1' we have $Q_1' Q_1^+ = 0$. In addition $Q_1^+ Q_2 = 0$ since $\mathfrak{Q}_2 \subset \mathfrak{R}(Q_1)$ so that we have

$$(I - Q_1)(I - Q_2) = (I - Q_1')(I - Q_1^+)(I - Q_2) = (I - Q_1')(I - Q_1^+ - Q_2). \tag{3.18}$$

Using (3.18) and the notation $\tilde{Q}_2 = Q_1^+ + Q_2$ we can write $\hat{A}_2$ in the form

$$\hat{A}_2 = \hat{A} - (I-\tilde{Q}_2)^*(I-Q_1')^*(\hat{A}-\rho)(I-Q_1')(I-\tilde{Q}_2).$$

For the comparison we consider the operator $\tilde{A}_2$ defined on the larger space $\hat{\mathfrak{H}}'$ by

$$\tilde{A}_2 = \hat{A}' - (I-\tilde{Q}_2)^*(I-Q_1')^*(\hat{A}'-\rho)(I-Q_1')(I-\tilde{Q}_2).$$

Two observations about $\hat{A}_2$ complete the argument: First, the part of $\tilde{A}_2$ in $\hat{\mathfrak{H}}$ is exactly $\hat{A}_2$; second, we have

$$\tilde{A}_2 \leqq \hat{A}_2',$$

which is an immediate consequence of the inequality (3.1). Thus $\hat{A}_2$ is smaller than the part of $\hat{A}_2'$ in $\hat{\mathfrak{H}}$. As a consequence $\hat{A}_2'$ gives more lower bounds and better lower bounds than those of $\hat{A}_2$. This is summarized by

$$\begin{aligned} &\hat{\mu}_\nu' \leqq \lambda_{\nu-n'+r'}, && \nu = 1, 2, \cdots, n-r, \\ &\hat{\mu}_{\nu-n+r} \leqq \hat{\mu}_\nu' \leqq \lambda_{\nu-n'+r'}, && \nu = n-r+1, n-r+2, \cdots. \end{aligned} \tag{3.19}$$

3.4. Increase in ρ. If ρ is increased to $\rho\upsilon$ while (1.1) continues to hold and the spaces $\mathfrak{P}$ and $\mathfrak{Q}$ are not changed, the spaces $\mathfrak{Q}_1$, $\mathfrak{Q}_2$, $\mathfrak{P}_1$, and $\mathfrak{P}_2$ can be kept the same. Let $\hat{A}_2'$ be the new intermediate operator constructed using ρ', i.e.,

$$\hat{A}_2' = \hat{A}_0' + \hat{A}_1' Q_2' = \hat{A} - \hat{A}_1'(I-Q_2') \quad \text{on } \hat{\mathfrak{H}},$$

where

$$\hat{A}_1' = (I-Q_1)^*(\hat{A}-\rho')(I-Q_1).$$

Since $\rho' \geqq \rho$ we have $\hat{A}_1' \leqq \hat{A}_1$, and by (3.3) it follows that

$$\hat{A}_1(I-Q_2) \geqq \hat{A}_1'(I-Q_2') \quad \text{in } \hat{\mathfrak{H}},$$

and hence

$$\hat{A}_2 \leqq \hat{A}_2'. \tag{3.20}$$

From this it follows that the eigenvalues $\hat{\mu}'$ of $\hat{A}_2'$ are better lower bounds than those of $\hat{A}_2$, i.e.,

$$\hat{\mu}_\nu \leqq \hat{\mu}_\nu' \leqq \lambda_{\nu+n-r}, \qquad \nu = 1, 2, \cdots.$$

REFERENCES

[1] N. ARONSZAJN, *Approximation methods for eigenvalues of completely continuous symmetric operators*, Proc. Symp. Spectral Theory and Differential Problems, Mathematics Dept. Oklahoma Agricultural and Mechanical College, Stillwater, 1951, pp. 179–202.

[2] N. BAZLEY AND D. W. FOX, *Comparison operators for lower bounds to eigenvalues*, J. Reine Agnew Math., 223 (1966), pp. 142–149.

[3] ———, *Truncations in the method of intermediate problems for lower bounds to eigenvalues*, J. Res. Nat. Bur. Standards Sect. B, 65B (1961), pp. 105–111.

[4] 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.

[5] S. GOULD, *Variational Methods for Eigenvalue Problems*, 2nd ed., University of Toronto Press, 1966.

[6] H. F. WEINBERGER, *A theory of lower bounds for eigenvalues*, Tech. Note BN-183, Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, College Park, 1959.

[7] ———, *Variational Methods for Eigenvalue Approximation*, SIAM Regional Conf. Series in Appl. Math., Society for Industrial and Applied Mathematics, Philadelphia, 1974.

[8] A. WEINSTEIN, *Étude des spectres des équations aux derivées partielles*, Mémor. Sci. Math., 88 (1937).

[9] A. WEINSTEIN AND W. STENGER, *Methods of Intermediate Problems for Eigenvalues: Theory and Ramifications*, Academic Press, New York, 1972.

SIAM J. MATH. ANAL.
Vol. 8, No. 3, May 1977

ON THE TREATMENT OF A DIRICHLET–NEUMANN MIXED BOUNDARY VALUE PROBLEM FOR HARMONIC FUNCTIONS BY AN INTEGRAL EQUATION METHOD*

R. GONZÁLEZ AND R. KRESS†

Abstract. Using an approach which extends the well-known classical integral equation methods, we reduce a mixed boundary value problem for harmonic functions to a system consisting of two integral equations of the second kind. Existence is proved by the Fredholm alternative for compact operators. The integral equations can be solved approximately by successive iterations. Further investigations are made on the spectrum of the boundary integral operator.

1. Introduction. Let B be an open domain of $\mathbb{R}^m$, $m \geqq 3$, with boundary ∂B which is decomposable in the form

$$\partial B = \partial B_D \cup \partial B_N.$$

∂B_D and ∂B_N are nonempty sets and consist of a finite number of disjoint, closed, bounded Lyapunov surfaces and on which we shall impose Dirichlet and Neumann boundary conditions, respectively. The complement of $\bar{B}$ in $\mathbb{R}^m$ is denoted by $\hat{B}$. The union of all the components of $\hat{B}$ which have boundaries in $\partial B_D(\partial B_N)$ will be denoted by $\hat{B}_D(\hat{B}_N)$. If $\hat{B}$ is unbounded we denote the unbounded component of $\hat{B}$ by E with boundary ∂E.

In this paper we shall study the following Dirichlet–Neumann mixed boundary value problem for the Laplace equation:

$$\Delta u = 0 \quad \text{in } B, \tag{1.1}$$

$$u = f_D \quad \text{on } \partial B_D, \tag{1.2}$$

$$\frac{\partial u}{\partial n} = f_N \quad \text{on } \partial B_N, \tag{1.3}$$

where f_D and f_N are given continuous complex functions on ∂B_D and ∂B_N, and n is the outward drawn unit normal to ∂B. We shall confine ourselves to classical solutions, that is, we shall consider solutions which belong to $C^2(B) \cap C(\bar{B})$ and for which the normal derivative $\partial u/\partial n$ exists on ∂B_N in the sense of uniform convergence by approaching ∂B_N from inside B along the normal direction. In the case when B is unbounded we impose the additional restriction

$$u(x) \to 0, \qquad |x| \to \infty, \tag{1.4}$$

uniformly. Mixed boundary value problems of this type arise in many fields of mathematical physics, for instance in elasticity, heat conduction and eletrostatics.

By Hayes and Kellner [2] the mixed boundary value problem (1.1)–(1.3) for harmonic functions in $\mathbb{R}^2$ was reduced to a pair of coupled integral equations consisting of one equation of the first kind and one equation of the second kind. In general, however, integral equations of the second kind seem to be more

* Received by the editors October 21, 1975, and in revised form April 12, 1976.

† Lehrstühle für Numerische und Angewandte Mathematik, Universität Göttingen, D-34 Göttingen, West Germany.

advantageous in order to obtain existence and uniqueness theorems by the Fredholm alternative for compact operators.

Therefore in this paper we shall present an approach which extends the well-known classical integral equation method of Fredholm for the case when either ∂B_D or ∂B_N is empty and which replaces the mixed boundary value problem by an equivalent pair of coupled boundary integral equations of the second kind from which an existence theorem for the mixed boundary value problem can be derived. Furthermore, this system of integral equations can be solved approximately by iteration methods. Additional investigations are made on the spectrum of the boundary integral operator. The eigenspaces and generalized eigenspaces are studied and by an example it is demonstrated that the spectrum of the boundary integral operator in general is not real as it is in the limiting cases $\partial B_D = \varnothing$ or $\partial B_N = \varnothing$.

Without going into the details we remark that all results remain valid in $\mathbb{R}^2$ with some modifications in the proofs due to the behavior of the fundamental solution $\ln 1/|x-x'|$ at infinity. In a similar way a Dirichlet–Neumann mixed boundary value problem for harmonic vectorfields was treated in [3].

We note that on each closed surface we have either the Dirichlet or the Neumann condition. The more difficult problem in which the Dirichlet conditions are imposed on part of a closed surface and Neumann conditions on the remaining part is not considered.

2. Equivalent boundary integral equation. Let $C(\partial B_D)(C(\partial B_N))$ be the Banach space of continuous functions $\phi_D\colon \partial B_D \to \mathbb{C}(\phi_N\colon \partial B_N \to \mathbb{C})$ with norm $\|\phi_D\| := \sup_{x\in\partial B_D} |\phi_D(x)|(\|\phi_N\| := \sup_{x\in\partial B_N} |\phi_N(x)|)$. We can identify the Banach space $C(\partial B)$ of continuous functions $\Phi\colon \partial B \to \mathbb{C}$ with the Cartesian product $C(\partial B) = C(\partial B_D) \times C(\partial B_N)$.

Define compact linear integral operators $A, A'\colon C(\partial B) \to C(\partial B)$ by

$$\Phi = \begin{pmatrix} \phi_D \\ \phi_N \end{pmatrix} \to A\Phi := \begin{pmatrix} \displaystyle\int_{\partial B_D} \phi_D(x') \frac{\partial \gamma(x, x')}{\partial n(x')}\, ds(x') \\ \quad + \displaystyle\int_{\partial B_N} \phi_N(x')\gamma(x, x')\, ds(x') \\ \displaystyle -\int_{\partial B_D} \phi_D(x') \frac{\partial^2 \gamma(x, x')}{\partial n(x)\, \partial n(x')}\, ds(x') \\ \displaystyle -\int_{\partial B_N} \phi_N(x') \frac{\partial \gamma(x, x')}{\partial n(x)}\, ds(x') \end{pmatrix} \tag{2.1}$$

and

$$\Psi = \begin{pmatrix} \psi_D \\ \psi_N \end{pmatrix} \to A'\Psi := \begin{pmatrix} \displaystyle\int_{\partial B_D} \psi_D(x') \frac{\partial \gamma(x, x')}{\partial n(x)}\, ds(x') \\ \displaystyle -\int_{\partial B_N} \psi_N(x') \frac{\partial^2 \gamma(x, x')}{\partial n(x)\, \partial n(x')}\, ds(x') \\ \displaystyle\int_{\partial B_D} \psi_D(x')\gamma(x, x')\, ds(x') \\ \displaystyle -\int_{\partial B_N} \psi_N(x') \frac{\partial \gamma(x, x')}{\partial n(x')}\, ds(x') \end{pmatrix} \tag{2.2}$$

where

$$\gamma(x, x') := \frac{2}{(m-2)\omega_m} \frac{1}{|x-x'|^{m-2}}, \qquad x \neq x',$$

denotes the fundamental solution of the Laplace equation in $\mathbb{R}^m$ with ω_m being the area of the unit sphere. The compactness of A and A' follows from the weak singularity of the kernels on $\partial B_D \times \partial B_D$ and $\partial B_N \times \partial B_N$ and the continuity of the kernels on $\partial B_D \times \partial B_N$ and $\partial B_N \times \partial B_D$. As is easily seen by interchange of order of integrations, A and A' are adjoint with respect to the bilinear form $(\cdot, \cdot)$: $C(\partial B) \times C(\partial B) \to \mathbb{C}$ defined by

$$(\Phi, \Psi) := \int_{\partial B} \Phi\Psi \, ds = \int_{\partial B_D} \phi_D \psi_D \, ds + \int_{\partial B_N} \phi_N \psi_N \, ds; \tag{2.3}$$

that means there holds

$$(A\Phi, \Psi) = (\Phi, A'\Psi), \qquad \Phi, \Psi \in C(\partial B). \tag{2.4}$$

Extending the classical approach due to Fredholm we seek a solution u of the Dirichlet–Neumann mixed boundary value problem in the form of a double layer potential on ∂B_D and a single layer potential on ∂B_N

$$\begin{aligned} u(x) = &\int_{\partial B_D} \phi_D(x') \frac{\partial \gamma(x, x')}{\partial n(x')} \, ds(x') \\ &+ \int_{\partial B_N} \phi_N(x') \gamma(x, x') \, ds(x') \end{aligned} \tag{2.5}$$

with continuous densities ϕ_D and ϕ_N. The potential u clearly satisfies the Laplace equation in both B and $\hat{B}$ and furthermore $u(x) \to 0$, $|x| \to \infty$. From the well-known jump conditions of potential theory [6, p. 32] we get

$$\begin{aligned} u_\pm = &\int_{\partial B_D} \phi_D(x') \frac{\partial \gamma(x, x')}{\partial n(x')} \, ds(x') \\ &+ \int_{\partial B_N} \phi_N(x') \gamma(x, x') \, ds(x') \pm \phi_D \quad \text{on } \partial B_D, \end{aligned} \tag{2.6}$$

$$\frac{\partial u_+}{\partial n} = \frac{\partial u_-}{\partial n} \quad \text{on } \partial B_D, \tag{2.7}$$

$$\begin{aligned} \frac{\partial u_\pm}{\partial n} = &\int_{\partial B_D} \phi_D(x') \frac{\partial^2 \gamma(x, x')}{\partial n(x) \, \partial n(x')} \, ds(x') \\ &+ \int_{\partial B_N} \phi_N(x') \frac{\partial \gamma(x, x')}{\partial n(x)} \, ds(x') \mp \phi_N \quad \text{on } \partial B_N, \end{aligned} \tag{2.8}$$

$$u_+ = u_- \quad \text{on } \partial B_N. \tag{2.9}$$

With the indices $+$ and $-$ we distinguish the limits obtained by approaching ∂B from inside $\hat{B}$ and B, respectively. Employing the above relations, we

immediately have

THEOREM 2.1. *The potential*

$$u(x)=\int_{\partial B_D}\phi_D(x')\frac{\partial\gamma(x,x')}{\partial n(x')}\,ds(x')+\int_{\partial B_N}\phi_N(x')\gamma(x,x')\,ds(x')$$

with continuous densities ϕ_D and ϕ_N solves the mixed boundary value problem

$$\Delta u=0 \quad \text{in } B,$$

$$u=f_D \quad \text{on } \partial B_D,$$

$$\frac{\partial u}{\partial n}=f_N \quad \text{on } \partial B_N,$$

(and $u(x)\to 0, |x|\to\infty$ if B is unbounded) iff $\Phi:=\begin{pmatrix}\phi_D\\ \phi_N\end{pmatrix}$ solves the integral equation

$$\Phi-A\Phi=F$$

where $F:=\begin{pmatrix}-f_D\\ f_N\end{pmatrix}$.

Analogously the adjoint operator A' corresponds to the mixed boundary value problem in the complement $\hat{B}$ with Dirichlet conditions on ∂B_N and Neumann conditions on ∂B_D.

THEOREM 2.2. *The potential*

$$v(x)=\int_{\partial B_D}\psi_D(x')\gamma(x,x')\,ds(x')-\int_{\partial B_N}\psi_N(x)\frac{\partial\gamma(x,x')}{\partial n(x')}\,ds(x')$$

with continuous densities ψ_D and ψ_N solves the mixed boundary value problem

$$\Delta v=0 \quad \text{in } \hat{B},$$

$$\frac{\partial v}{\partial n}=g_D \quad \text{on } \partial B_D,$$

$$v=g_N \quad \text{on } \partial B_N,$$

(and $v(x)\to 0, |x|\to\infty$ if $\hat{B}$ is unbounded) iff $\Psi:=\begin{pmatrix}\psi_D\\ \psi_N\end{pmatrix}$ solves the integral equation

$$\Psi-A'\Psi=G$$

where $G:=\begin{pmatrix}-g_D\\ -g_N\end{pmatrix}$.

As is easily seen from the signs in the jump conditions (2.6) and (2.8), the operator $-A$ with respect to the potential u from (2.5) corresponds to the mixed boundary value problem in $\hat{B}$ with Dirichlet conditions on ∂B_D and Neumann conditions on ∂B_N. Thus, we summarize the correspondence of boundary integral operators and mixed boundary value problems in Table 1.

TABLE 1

Operator	A	A'
Region	B	$\hat{B}$
Dirichlet condition	∂B_D	∂B_N
Neumann condition	∂B_N	∂B_D
Operator	$-A$	$-A'$
Region	$\hat{B}$	B
Dirichlet condition	∂B_D	∂B_N
Neumann condition	∂B_N	∂B_D

3. Uniqueness of the mixed boundary value problem.

THEOREM 3.1. *The Dirichlet–Neumann mixed boundary value problem has not more than one solution.*

Proof. Let u be a solution of the homogeneous mixed boundary value problem. By the maximum principle its greatest and least values are attained on the boundary ∂B or at infinity if B is unbounded. If the maximum or minimum is attained on ∂B_N by a theorem of Giraud [6, p. 7] from $\partial u/\partial n = 0$ on ∂B_N we deduce $u = \text{const.}$ in B and then from $u = 0$ on ∂B_D we get $u = 0$ in B. If maximum and minimum are attained on ∂B_D or at infinity we immediately have $u = 0$ in B. Hence uniqueness is proved. We remark that Green's theorem cannot be applied because we have not assumed the normal derivative $\partial u/\partial n$ to exist on ∂B_D.

4. Homogeneous boundary integral equations.

THEOREM 4.1. *Let p_D be the number of bounded components of $\hat{B}_D$. Then for the null space of $I - A$ (I identity) there holds*

$$\dim N(I - A) = p_D.$$

Remark. This result also holds in the case $\partial B_N = \varnothing$. In the case $\partial B_D = \varnothing$ we have $\dim N(I - A) = 1$.

Proof. Let Φ be any solution of the equation

$$\Phi - A\Phi = 0. \tag{4.1}$$

Then by Theorem 2.1 the potential u defined by (2.5) solves the homogeneous mixed boundary value problem. Hence from our uniqueness theorem there follows

$$u = 0 \quad \text{in } B. \tag{4.2}$$

Making use of the continuity (2.9), we now get $u_+ = 0$ on ∂B_N and the maximum principle yields

$$u = 0 \quad \text{in } \hat{B}_N. \tag{4.3}$$

From the continuity (2.7) we have $\partial u_+/\partial n = 0$ on ∂B_D. Thus, employing Green's theorem, we find

$$\operatorname{grad} u = 0 \quad \text{in } \hat{B}_D, \tag{4.4}$$

which means $u \in H(\hat{B}_D)$, where $H(\hat{B}_D)$ denotes the linear space of all functions which are constant on the components of $\hat{B}_D$ and vanish in the unbounded component E if $\hat{B}_D$ is unbounded. It is obvious that $\dim H(\hat{B}_D) = p_D$. From $2\phi_D = u_+ - u_-$ on ∂B_D we derive $\phi_D \in K(\partial B_D)$, where $K(\partial B_D) := H(\hat{B}_D)|_{\partial B_D}$ is the linear space of all functions which are constant on the components of ∂B_D and vanish on ∂E if $\hat{B}_D$ is unbounded. Finally, $2\phi_N = \partial u_-/\partial n - \partial u_+/\partial n$ on ∂B_N yields $\phi_N = 0$. Therefore, any solution Φ of (4.1) must be of the form $\Phi = \begin{pmatrix} \kappa \\ 0 \end{pmatrix}$ with $\kappa \in (\partial B_D)$.

Conversely, let Φ be of the form $\Phi = \begin{pmatrix} \kappa \\ 0 \end{pmatrix}$ where $\kappa \in K(\partial B_D)$. Then by Gauss' theorem we deduce

$$\int_{\partial B_D} \kappa(x') \frac{\partial \gamma(x, x')}{\partial n(x')} ds(x') = \begin{cases} 0, & x \in B \cup \bar{\hat{B}}_N, \\ 2\kappa(x), & x \in \hat{B}_D. \end{cases}$$

Hence, using the jump conditions, it follows that

$$\int_{\partial B_D} \kappa(x') \frac{\partial \gamma(x, x')}{\partial n(x')} ds(x') = \kappa(x'), \qquad x \in \partial B_D, \tag{4.5}$$

and

$$\int_{\partial B_D} \kappa(x') \frac{\partial^2 \gamma(x, x')}{\partial n(x)\, \partial n(x')} ds(x') = 0, \qquad x \in \partial B_N, \tag{4.6}$$

that is, $\Phi - A\Phi = 0$. Thus, we have established

$$N(I - A) = K(\partial B_D) \times \{0\}, \tag{4.7}$$

and therefore $\dim N(I - A) = p_D$.

THEOREM 4.2. *Let p_N be the number of bounded components of $\hat{B}_N$. Then for the null space of $I + A$ there holds*

$$\dim N(I + A) = p_N.$$

Remark. This result holds in the case $\partial B_D = \varnothing$. In the case $\partial B_N = \varnothing$, we have $\dim N(I + A) = 1$.

Proof. Using Table 1 at the end of § 2 and applying Theorem 4.1 to $I + A'$, we get $\dim N(I + A') = p_N$. Hence by the Fredholm alternative $\dim N(I + A) = p_N$.

THEOREM 4.3. *If $\hat{B}_N$ is bounded the generalized null space of $I - A$ is equal to $N(I - A)$ (the Riesz number is one). If $\hat{B}_N$ is unbounded the generalized null space of $I - A$ is equal to $N(I - A)^2$ (the Riesz number is two) and has dimension $p_D + 1$.*

Proof. Let $\Phi_2 \in N(I - A)^2$ and define $\Phi_1 := \Phi_2 - A\Phi_2$. Then

$$\Phi_1 - A\Phi_1 = 0, \qquad \Phi_2 - A\Phi_2 = \Phi_1. \tag{4.8}$$

Define potentials u_1 and u_2 with layers Φ_1 and Φ_2 by (2.5). From the proof of Theorem 4.1 we already know

$$u_1 = 0 \quad \text{in } B \cup \bar{\hat{B}}_N, \tag{4.9}$$

$$\text{grad } u_1 = 0 \quad \text{in } \hat{B}_D. \tag{4.10}$$

Applying Theorem 2.1 to the second equation of (4.8) we then find

$$u_{2-} = -\phi_{D,1} = -\tfrac{1}{2}u_{1+} \quad \text{on } \partial B_D, \tag{4.11}$$

$$\frac{\partial u_2}{\partial n} = \phi_{N,1} = 0 \qquad \text{on } \partial B_N. \tag{4.12}$$

With the aid of $\partial u_{2+}/\partial n = \partial u_{2-}/\partial n$ on ∂B_D, Green's theorem yields

$$2\int_B |\text{grad } u_2|^2 \, dx = 2\int_{\partial B} u_{2-} \overline{\frac{\partial u_{2-}}{\partial n}} \, ds$$

$$= -\int_{\partial B_D} u_{1+} \overline{\frac{\partial u_{2+}}{\partial n}} \, ds = \int_{B_D} (\text{grad } u_1, \text{grad } \overline{u_2}) \, dx = 0.$$

Hence

$$\text{grad } u_2 = 0 \quad \text{in } B. \tag{4.13}$$

We now have to distinguish between the two cases where $\hat{B}_N$ is bounded or unbounded.

(i) Let $\hat{B}_N$ be bounded. In the case B is unbounded, from (4.13) there follows $u_2 = 0$ in B. Now (4.11) yields $u_{1+} = 0$ on ∂B_D, and by the maximum principle $u_1 = 0$ in $\hat{B}_D$. Hence, from $\phi_{D,1} = \frac{1}{2}(u_{1+} - u_{1-})$ on ∂B_D we get $\phi_{D,1} = 0$ which means $\Phi_1 = 0$.

In the case $\hat{B}_D$ is unbounded we have $u_1 = 0$ in the unbounded component E of $\hat{B}_D$. Then from (4.11) and (4.13) we again get $u_2 = 0$ in B and, arguing as in the previous case, we derive $\Phi_1 = 0$. Thus, $\Phi_2 \in N(I - A)$ and if $\hat{B}_N$ is bounded we indeed have $N(I - A) = N(I - A)^2$.

(ii) Let $\hat{B}_N$ be unbounded. From (4.13) there follows $u_2 = \text{const.}$ in B. Using $\partial u_{2+}/\partial n = \partial u_{2-}/\partial n$ on ∂B_D we get $\partial u_{2+}/\partial n = 0$ on ∂B_D. Hence $u_2 \in H(\hat{B}_D)$ and $\phi_{D,2} = \frac{1}{2}(u_{2+} - u_{2-}) \in K(\partial B_D)$. From $u_{2+} = u_{2-}$ on ∂B_N we deduce $u_{2+} = \text{const.}$ on ∂B_N and therefore $u_2 = \text{const.}$ in $\hat{B}_N \backslash E$. Thus $\phi_{N,2} = -\frac{1}{2}(\partial u_{2+}/\partial n - \partial u_{2-}/\partial n) = 0$ on $\partial B_N \backslash \partial E$. In view of (4.6), the second component of the equation $\Phi_2 - A\Phi_2 = \Phi_1$ now reads

$$\phi_{N,2}(x) + \int_{\partial E} \phi_{N,2}(x') \frac{\partial \gamma(x, x')}{\partial n(x)} \, ds(x') = 0, \qquad x \in \partial E. \tag{4.14}$$

This integral equation corresponds to the Neumann problem in $\hat{E} := \mathbb{R}^m \backslash \bar{E}$ and, as is well known, it has one linearly independent solution [6, p. 82]. Let $N(\partial B_N)$ be the one-dimensional linear space of all functions χ defined on ∂B_N which vanish on $\partial B_N \backslash \partial E$ and for which $\chi|_{\partial E}$ is a solution of (4.14). Then we have established that any $\Phi_2 \in N(I - A)^2$ must be of the form $\Phi_2 = \begin{pmatrix} \kappa \\ \chi \end{pmatrix}$ with $\kappa \in K(\partial B_D)$ and $\chi \in N(\partial B_N)$.

Conversely, let Φ_2 be of the form $\Phi_2 = \begin{pmatrix} \kappa \\ \chi \end{pmatrix}$ where $\kappa \in K(\partial B_D)$ and $\chi \in N(\partial B_N)$. Then for the potential

$$w(x) = \int_{\partial E} \chi(x')\gamma(x, x')\, ds(x'),$$

by the jump conditions and the integral equation (4.14) we get $\partial w_-/\partial n = 0$ on ∂E. Hence $w = \text{const.}$ in $\hat{E}$, and in particular,

$$\int_{\partial B_N} \chi(x')\gamma(x, x')\, ds(x') = \text{const.}, \qquad x \in \partial B_D.$$

Combining this with (4.5), (4.6) and (4.14) we conclude $\Phi_2 - A\Phi_2 = \Phi_1$ with $\phi_{D,1} = \text{const.}$ on ∂B_D and $\phi_{N,1} = 0$ on ∂B_N. Hence $(I-A)^2\Phi_2 = (I-A)\Phi_1 = 0$. Thus, we have established

$$N(I-A)^2 = K(\partial B_D) \times N(\partial B_N), \tag{4.15}$$

and hence $\dim N(I-A)^2 = p_D + 1$.

To complete the proof it remains to show that $N(I-A)^3 = N(I-A)^2$. Let $\Phi_3 \in N(I-A)^3$ and define $\Phi_2 := \Phi_3 - A\Phi_3$, $\Phi_1 := \Phi_2 - A\Phi_2$. Then

$$\Phi_1 - A\Phi_1 = 0, \quad \Phi_2 - A\Phi_2 = \Phi_1, \quad \Phi_3 - A\Phi_3 = \Phi_2. \tag{4.16}$$

Define potentials u_1, u_2 and u_3 with layers Φ_1, Φ_2 and Φ_3 by (2.5). Then in addition to the above results on u_1 and u_2, applying Theorem 2.1 to the third equation of (4.16), we get

$$u_{3-} = -\phi_{D,2} = -\tfrac{1}{2}(u_{2+} - u_{2-}) \quad \text{on } \partial B_D, \tag{4.17}$$

$$\frac{\partial u_{3-}}{\partial n} = \phi_{N,2} = -\tfrac{1}{2}\left(\frac{\partial u_{2+}}{\partial n} - \frac{\partial u_{2-}}{\partial n}\right) \quad \text{on } \partial B_N. \tag{4.18}$$

Since $u_2 = \text{const.}$ in B, using Green's theorem and $u_{2+} = u_{2-}$ on ∂B_N, we derive

$$\begin{aligned} \int_{B_N} |\operatorname{grad} u_2|^2\, dx &= -\int_{\partial B_N} u_{2+} \overline{\frac{\partial u_{2+}}{\partial n}}\, ds \\ &= 2\int_{\partial B_N} u_{2-} \overline{\frac{\partial u_{3-}}{\partial n}}\, ds = 2\int_{\partial B} u_{2-} \overline{\frac{\partial u_{3-}}{\partial n}}\, ds \\ &= 2\int_B (\operatorname{grad} u_2, \operatorname{grad} \overline{u_3})\, dx = 0. \end{aligned}$$

The integral

$$\int_{\partial B_D} u_{2-} \overline{\frac{\partial u_{3-}}{\partial n}}\, ds = \int_{\partial B_D} u_{2-} \overline{\frac{\partial u_{3+}}{\partial n}}\, ds$$

vanishes because $u_2 = \text{const.}$ in B and $\hat{B}_D$ is bounded. Now we conclude

$$\operatorname{grad} u_2 = 0 \quad \text{in } \hat{B}_N,$$

and therefore $u_2 = 0$ in the unbounded component E. Therefore $\phi_{N,2} = -\frac{1}{2}(\partial u_{2+}/\partial n - \partial u_{2-}/\partial n) = 0$ on ∂E. Summarizing all results on Φ_2, we find $\Phi_2 \in K(\partial B_D) \times \{0\} = N(I-A)$. Thus, $\Phi_3 \in N(I-A)^2$ and we indeed have $N(I-A)^2 = N(I-A)^3$.

Arguing similarly as in Theorem 4.2 we deduce

THEOREM 4.4. *If $\hat{B}_D$ is bounded, the generalized null space of $I+A$ is equal to $N(I+A)$ (the Riesz number is one). If $\hat{B}_D$ is unbounded the generalized null space of $I+A$ is equal to $N(I+A)^2$ (the Riesz number is two) and has dimension p_N+1.*

Remark. The above results on the generalized null spaces are different from the corresponding results in the limiting cases $\partial B_D = \emptyset$ or $\partial B_N = \emptyset$, where we always have Riesz number one. In the case of the Neumann problem where $\partial B_D = \emptyset$ there is $p_D = 0$ but we still have dim $N(I-A) = 1$, as mentioned in the proof of Theorem 4.3. Thus our results are in agreement with the fact of convergence of generalized eigenspaces for compact operators.

5. Existence of the mixed boundary value problem.

THEOREM 5.1. *The Dirichlet-Neumann mixed boundary value problem has a unique solution.*

Proof. In the case $p_D = 0$ the homogeneous boundary integral equation $\Phi - A\Phi = 0$ has only the trivial solution. Hence by the Fredholm alternative the inhomogeneous boundary integral equation $\Phi - A\Phi = F$ has a unique solution Φ for all F. By Theorem 2.1 we then get a solution u of the mixed boundary value problem.

In the case $p_D > 0$ let Ψ_k, $k = 1, \cdots, p_D$, be a basis of the adjoint null space $N(I-A')$. Define potentials

$$
\begin{aligned}
v_k(x) := & \int_{\partial B_D} \psi_{D,k}(x')\gamma(x, x')\, ds(x') \\
& - \int_{\partial B_N} \psi_{N,k}(x') \frac{\partial \gamma(x, x')}{\partial n(x')}\, ds(x'), \qquad k = 1, \cdots, p_D.
\end{aligned} \tag{5.1}
$$

These functions are harmonic in both B and $\hat{B}$, and by Theorem 2.2 they satisfy

$$
\frac{\partial v_{k+}}{\partial n} = 0 \quad \text{on } \partial B_D, \tag{5.2}
$$

$$
v_{k+} = 0 \quad \text{on } \partial B_N. \tag{5.3}
$$

From these boundary conditions we conclude $v_k \in H(\hat{B}_D)$ and $v_k = 0$ in $\hat{B}_N$.

Let $\hat{B}_{D,i}$, $i = 1, \cdots, p_D$, be the bounded components of $\hat{B}_D$. Choose p_D points $x_i \in \hat{B}_{D,i}$, $i = 1, \cdots, p_D$, and define potentials

$$
w_i(x) := \gamma(x, x_i), \qquad x \neq x_i. \tag{5.4}
$$

Then for the matrix

$$a_{ik} := (W_i, \Psi_k) \tag{5.5}$$

where

$$W_i := \begin{pmatrix} w_i|_{\partial B_D} \\ -\dfrac{\partial w_i}{\partial n}\Big|_{\partial B_N} \end{pmatrix} \tag{5.6}$$

we calculate

$$a_{ik} = v_k(x_i). \tag{5.7}$$

Let λ_k be a solution of the homogeneous system

$$\sum_{k=1}^{p_D} \lambda_k a_{ik} = 0, \qquad i = 1, \cdots, p_D, \tag{5.8}$$

and define

$$v := \sum_{k=1}^{p_D} \lambda_k v_k.$$

Then (5.8) becomes $v(x_i) = 0$, $i = 1, \cdots, p_D$, and in view of $v \in H(\hat{B}_D)$ we get $v = 0$ in $\hat{B}_D$. Remembering $v = 0$ in $\hat{B}_N$ and using $v_+ = v_-$ on ∂B_D and $\partial v_+/\partial n = \partial v_-/\partial n$ on ∂B_N, we derive $v_- = 0$ on ∂B_D and $\partial v_-/\partial n = 0$ on ∂B_N. Hence by the uniqueness theorem we conclude $v = 0$ in B. Thus we have

$$\sum_{k=1}^{p_D} \lambda_k v_k = 0 \quad \text{in } \mathbb{R}^m.$$

By the jump conditions the potentials v_k are linearly independent since the Ψ_k's are. Therefore $\lambda_k = 0$, $k = 1, \cdots, p_D$, and the matrix a_{ik} is regular.

Now the inhomogeneous system

$$\sum_{i=1}^{p_D} \lambda_i a_{ik} = (F, \Psi_k), \qquad k = 1, \cdots, p_D, \tag{5.9}$$

has a unique solution λ_i for all inhomogenities $F = \begin{pmatrix} -f_D \\ f_N \end{pmatrix}$. We define

$$\tilde{F} := F - \sum_{i=1}^{p_D} \lambda_i W_i.$$

Then $(\tilde{F}, \Psi_k) = 0$, $k = 1, \cdots, p_D$, and by the Fredholm alternative the boundary integral equation $\Phi - A\Phi = \tilde{F}$ has a solution Φ. By Theorem 2.1 this solution of the integral equation yields a solution $\tilde{u}$ of the mixed boundary value problem with boundary data

$$\tilde{u} = \tilde{f}_D = f_D + \sum_{i=1}^{p_D} \lambda_i w_i \qquad \text{on } \partial B_D,$$

$$\frac{\partial \tilde{u}}{\partial n} = \tilde{f}_N = f_N + \sum_{i=1}^{p_D} \lambda_i \frac{\partial w_i}{\partial n} \qquad \text{on } \partial B_N.$$

Finally,

$$u := \tilde{u} - \sum_{i=1}^{p_D} \lambda_i w_i$$

solves the original boundary value problem.

6. On the spectrum of the boundary integral operator.

LEMMA 6.1. *Let $\lambda \in \mathbb{C}$ be an eigenvalue of A with eigenfunction Φ. Then the corresponding potential defined by*

$$u(x) := \int_{\partial B_D} \phi_D(x') \frac{\partial \gamma(x, x')}{\partial n(x')}\, ds(x') + \int_{\partial B_N} \phi_N(x') \gamma(x, x')\, ds(x') \tag{6.1}$$

satisfies

$$\Delta u = 0 \qquad \text{in } \mathbb{R}^m \backslash \partial B, \tag{6.2}$$

$$(\lambda - 1) u_+ = (\lambda + 1) u_- \qquad \text{on } \partial B_D, \tag{6.3}$$

$$\frac{\partial u_+}{\partial n} = \frac{\partial u_-}{\partial n} \qquad \text{on } \partial B_D, \tag{6.4}$$

$$(\lambda - 1) \frac{\partial u_+}{\partial n} = (\lambda + 1) \frac{\partial u_-}{\partial n} \qquad \text{on } \partial B_N, \tag{6.5}$$

$$u_+ = u_- \qquad \text{on } \partial B_N, \tag{6.6}$$

$$u(x) \to 0, \qquad |x| \to \infty. \tag{6.7}$$

Conversely, let λ and u be such that (6.2)–(6.7) *hold. Then λ is an eigenvalue of A with eigenfunction Φ where $\phi_D := \frac{1}{2}(u_+ - u_-)$ on ∂B_D and $\phi_N := -\frac{1}{2}(\partial u_+/\partial n - \partial u_-/\partial n)$ on ∂B_N.*

Proof. Let λ be an eigenvalue of A with eigenfunction Φ. Then the potential u defined by (6.1) clearly satisfies (6.2), (6.4), (6.6) and (6.7). Using $A\Phi = \lambda\Phi$ from the jump conditions we obtain

$$u_\pm = (\lambda \pm 1)\phi_D \qquad \text{on } \partial B_D,$$

$$\frac{\partial u_\pm}{\partial n} = -(\lambda \pm 1)\phi_N \qquad \text{on } \partial B_N,$$

whence (6.3) and (6.5) follow.

Conversely, let λ and u satisfy (6.2)–(6.7) and define $\phi_D := \frac{1}{2}(u_+ - u_-)$ and $\phi_N := -\frac{1}{2}(\partial u_+/\partial n - \partial u_-/\partial n)$. Inserting (6.4) and (6.6) into Green's representation formula

$$u(x) = -\frac{1}{2} \int_{\partial B} \left(\frac{\partial u_+(x')}{\partial n(x')} - \frac{\partial u_-(x')}{\partial n(x')} \right) \gamma(x, x')\, ds(x')$$

$$+ \frac{1}{2} \int_{\partial B} [u_+(x') - u_-(x')] \frac{\partial \gamma(x, x')}{\partial n(x')}\, ds(x'), \qquad x \in \mathbb{R}^m \backslash \partial B,$$

we obtain

$$u(x)=\int_{\partial B_D}\phi_D(x')\frac{\partial\gamma(x,x')}{\partial n(x')}\,ds(x')+\int_{\partial B_N}\phi_N(x')\gamma(x,x')\,ds(x').$$

Using the jump conditions we get

$$\begin{pmatrix} u_\pm|_{\partial B_D} \\ -\dfrac{\partial u_\pm}{\partial n}\Big|_{\partial B_N}\end{pmatrix}=A\Phi\pm\Phi.$$

Hence from (6.3) and (6.5) it follows that $A\Phi=\lambda\Phi$.

THEOREM 6.2. *The boundary integral operator A has spectral radius $\rho(A)=1$.*

Proof. Since A is compact, any nonzero number in the spectrum is an eigenvalue. Let λ be an eigenvalue. Then by Lemma 6.1 the following relationships hold for the corresponding potential:

$$(\bar\lambda+1)(\lambda-1)u_+\frac{\partial\bar u_+}{\partial n}=(\bar\lambda+1)(\lambda+1)u_-\frac{\partial\bar u_-}{\partial n}\quad\text{on }\partial B_D,$$

$$(\lambda+1)(\bar\lambda-1)u_+\frac{\partial\bar u_+}{\partial n}=(\lambda+1)(\bar\lambda+1)u_-\frac{\partial\bar u_-}{\partial n}\quad\text{on }\partial B_N.$$

Integrating over ∂B_D and ∂B_N and adding with the aid of Green's theorem we find

$$\begin{aligned}&(1+\bar\lambda)(1-\lambda)\int_{\hat B_D}|\operatorname{grad}u|^2\,dx+(1+\lambda)(1-\bar\lambda)\int_{\hat B_N}|\operatorname{grad}u|^2\,dx\\&\quad=|1+\lambda|^2\int_B|\operatorname{grad}u|^2\,dx.\end{aligned}$$

The real part of this equation gives

$$(1-|\lambda|^2)\int_{\hat B}|\operatorname{grad}u|^2\,dx=|1+\lambda|^2\int_B|\operatorname{grad}u|^2\,dx. \tag{6.8}$$

Now assume $|\lambda|\geqq 1$, $\lambda\neq 1$, $\lambda\neq -1$. Then from (6.8) there follows

$$\int_{\hat B}|\operatorname{grad}u|^2\,dx=\int_B|\operatorname{grad}u|^2\,dx=0,$$

that is, grad $u=0$ in $\mathbb{R}^m$. Since u vanishes at infinity, from (6.3) and (6.6) we deduce $u=0$ in $\mathbb{R}^m$. Hence by the jump conditions we find $\Phi=0$. Thus, all eigenvalues of A except possibly 1 and -1 have absolute value less than one. By Theorems 4.1 and 4.2 we have 1 or -1 an eigenvalue if $p_D>0$ or $p_N>0$, respectively. Since $\partial B_D\neq\varnothing$ and $\partial B_N\neq\varnothing$ then $p_D+p_N>0$. Therefore the desired result $\rho(A)=1$ is established.

Remark. Since the spectral radius $\rho(A)=1$ the boundary integral equation can be solved approximately by successive iterations when relaxation methods (in

case -1 is an eigenvalue) and deflation methods (in case 1 is an eigenvalue) are combined. Using a modified boundary layer approach due to Brakhage, Leis and Werner [1], [5, p. 165] one can reduce the mixed boundary value problem to an integral equation which is uniquely solvable. For the corresponding mixed boundary value problem for the Helmholtz equation this approach is investigated in [4]. However, in this case the spectral radius of the boundary integral operator does not remain less than or equal to one; consequently in the modified approach successive approximations do not converge.

By well-known results of Plemelj [7] the spectrum of the boundary integral operators in the limiting cases $\partial B_D = \varnothing$ or $\partial B_N = \varnothing$ is real. This in general is not true for the mixed problem. To show this we conclude with the following problem.

Example. Let B be the domain between two concentric spheres. ∂B_D is the interior sphere of radius 1, and ∂B_N is the exterior sphere of radius R. In spherical coordinates ρ, Θ the general function u harmonic in $\mathbb{R}^m \backslash \partial B$ and vanishing at infinity is of the form

$$u(\rho,\Theta)=\begin{cases}\displaystyle\sum_{n=0}^{\infty}\rho^n\sum_{k=1}^{k_{n,m}}a_n^{(k)}Y_{n,m}^{(k)}(\Theta), & 0\leqq\rho\leqq 1,\\ \displaystyle\sum_{n=0}^{\infty}\sum_{k=1}^{k_{n,m}}\left(\rho^n b_n^{(k)}+\frac{1}{\rho^{n+m-2}}c_n^{(k)}\right)Y_{n,m}^{(k)}(\Theta), & 1\leqq\rho\leqq R,\\ \displaystyle\sum_{n=0}^{\infty}\frac{1}{\rho^{n+m-2}}\sum_{k=1}^{k_{n,m}}d_n^{(k)}Y_{n,m}^{(k)}(\Theta), & R\leqq\rho<\infty,\end{cases}$$

where $Y_{n,m}^{(k)}$, $k=1,\cdots,k_{n,m}$, denote the linearly independent spherical harmonics of order n in $\mathbb{R}^m$. Comparing the coefficients of the three series we easily derive that (6.3)–(6.6) are satisfied iff

$$\begin{aligned}(1-\lambda)a_n^{(k)}+(1+\lambda)b_n^{(k)}+(1+\lambda)c_n^{(k)}&=0,\\ na_n^{(k)}-nb_n^{(k)}+(n+m-2)c_n^{(k)}&=0,\\ (1+\lambda)nb_n^{(b)}-\frac{(1+\lambda)(n+m-2)}{R^{2n+m-2}}c_n^{(k)}-\frac{(1-\lambda)(n+m-2)}{R^{2n+m-2}}d_n^{(k)}&=0,\\ b_n^{(k)}+\frac{1}{R^{2n+m-2}}c_n^{(k)}-\frac{1}{R^{2n+m-2}}d_n^{(k)}&=0,\end{aligned}$$

$k=1,\cdots,k_{n,m}$, $n=0,1,2,\cdots$. These equations have a nontrivial solution iff the determinate vanishes

$$(1-\lambda)\left\{[\lambda(2n+m-2)-(m-2)]^2+\frac{4n(n+m-2)}{R^{2n+m-2}}\right\}=0.$$

Hence, the spectrum of A contains the complex conjugate eigenvalues

$$\lambda_n=\frac{1}{2n+m-2}\left[m-2\pm\frac{2\sqrt{n(n+m-2)}}{R^{n-1+m/2}}i\right],\qquad n=0,1,2,\cdots.$$

REFERENCES

[1] H. BRAKHAGE AND P. WERNER, *Über das Dirichlet'sche Außenraumproblem für die Helmholtz'sche Schwingungsgleichung*, Arch. Math., 16 (1965), pp. 325–329.

[2] J. HAYES AND R. KELLNER, *The eigenvalue problem for a pair of coupled integral equations arising in the numerical solution of Laplace's equation*, SIAM J. Appl. Math., 22 (1972), pp. 503–513.

[3] R. KRESS, *Ein kombiniertes Dirichlet-Neumann'sches Randwertproblem bei harmonischen Vektorfeldern*, Arch. Rational Mech. Anal., 42 (1971), pp. 40–49.

[4] R. KRESS AND G. F. ROACH, *On mixed boundary value problems for the Helmholtz equation*, Proc. Roy. Soc. Edinburgh, 76A (1977), to appear.

[5] R. LEIS, *Vorlesungen über partielle Differentialgleichungen zweiter Ordnung*, Bibliographisches Institut, 1967.

[6] C. MIRANDA, *Partial Differential Equations of Elliptic Type*, Springer-Verlag, Berlin, 1970.

[7] J. PLEMELJ, *Potentialtheoretische Untersuchungen*, Preisschrift der Fürstlich Jablononwskischen Gesellschaft, Leipzig, 1911.

SIAM J. MATH. ANAL.
Vol. 8, No. 3, May 1977

A CLASS OF GENERATING FUNCTIONS*

L. CARLITZ†

Abstract. Generating functions of the type $\sum_0^\infty L_n^{(\alpha+\beta x)}(x)z^n$, $P_n^{(\alpha+\gamma n,\beta+\delta n)}(x)z^n$ have been obtained recently. In the present paper a general result of this kind is derived making use of the Lagrange expansion. A number of applications are given. These include, in addition to the Laguerre and Jacobi polynomials, the Hermite, Legendre, Bernoulli, Euler and Bell polynomials.

1. The Laguerre polynomial

$$L_n^{(\alpha)}(x)=\sum_{k=0}^{n}(-1)^k\frac{(\alpha+1)_n x^k}{k!(n-k)!(\alpha+1)_k}$$

has the familiar generating function [11, p. 100]

$$\sum_{n=0}^{\infty}L_n^{(\alpha)}(x)z^n=(1-z)^{-\alpha-1}\exp\left(\frac{-xz}{1-z}\right). \tag{1.1}$$

Also it has been shown recently [2], [3] that

$$\sum_{n=0}^{\infty}L_n^{(\alpha+\beta n)}(x)u^n=\frac{(1+v)^{\alpha+1}}{1-\beta v}e^{-xv}, \tag{1.2}$$

where α, β are arbitrary complex numbers and v is defined by

$$u=v(1+v)^{-\beta-1},\qquad v(0)=0. \tag{1.3}$$

Moreover if (1.2) is written in the form

$$\sum_{n=0}^{\infty}L_n^{(\alpha+\beta n)}(x)u^n=A(u,\alpha,\beta)\exp\{xB(u,\beta)\}, \tag{1.4}$$

then

$$\sum_{n=0}^{\infty}L_n^{(-\alpha-(\beta+1)n)}(x)u^n=\frac{A(-u,\alpha,\beta)}{1-B(-u,\beta)}\exp\left\{\frac{-xB(-u,\beta)}{1-B(-u,\beta)}\right\}, \tag{1.5}$$

where

$$A(u,\alpha,\beta)=\frac{(1+v)^{\alpha+1}}{1-\beta v},\qquad B(u,\beta)=-v. \tag{1.6}$$

It is also shown in [3] that

$$B(u,\beta)=-\sum_{n=1}^{\infty}\binom{(\beta+1)n}{n-1}\frac{u^n}{n} \tag{1.7}$$

and

$$B(-u,-\beta-1)=\frac{-B(u,\beta)}{1-B(u,\beta)}. \tag{1.8}$$

* Received by the editors July 22, 1975, and in revised form February 28, 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.

Thus when u is replaced by $-u$ and β is replaced by $-\beta-1$,

$$v \to \frac{-v}{1+v}.$$

From these observations the equivalence of (1.4) and (1.5) is immediate.

The Jacobi polynomial

$$P_n^{(\alpha,\beta)}(x) = \sum_{k=0}^{n} \binom{n+\alpha}{k}\binom{n+\beta}{n-k}\left(\frac{x-1}{2}\right)^{n-k}\left(\frac{x+1}{2}\right)^{k}$$

has the familiar generating function [11, p. 69]

$$\sum_{n=0}^{\infty} P_n^{(\alpha,\beta)}(x)z^n = 2^{\alpha+\beta}\rho^{-1}(1-z+\rho)^{-\alpha}(1+z+\rho)^{-\beta}, \tag{1.9}$$

where

$$\rho = (1-2xz+z^2)^{1/2}.$$

Also it has been proved by Srivastava and Singhal [9] that

$$\sum_{n=0}^{\infty} P_n^{(\alpha-\lambda n,\beta-\mu n)}(x)z^n = (1+u)^{-\alpha}(1+v)^{-\beta}[1+(1-\lambda)u+(1-\mu)v]^{-1} \tag{1.10}$$

where u, v satisfy

$$\begin{aligned} u &= -\tfrac{1}{2}(x+1)z(1+u)^{\lambda}(1+v)^{\mu-1}, \\ v &= -\tfrac{1}{2}(x-1)z(1+u)^{\lambda-1}(1+v)^{\mu}. \end{aligned} \tag{1.11}$$

In another paper ([10]; see also [1]) Srivastava and Singhal have obtained a result like (1.8) for a more general class of polynomials.

The above results suggest the following more general situation. Let

$$A(z) = 1+\sum_{n=1}^{\infty} a_n z^n/n!, \qquad B(z) = 1+\sum_{n=1}^{\infty} b_n z^n/n! \tag{1.12}$$

denote arbitrary functions that are analytic about the origin and put

$$A(z)(B(z))^{\lambda} = \sum_{n=0}^{\infty} c_n^{(\lambda)} z^n/n!, \tag{1.13}$$

thus defining the sequence $\{c_n^{(\lambda)}\}$ depending on the parameter λ. We seek a generating function for

$$\sum_{n=0}^{\infty} c_n^{(\lambda+\mu n)} z^n/n!. \tag{1.14}$$

This is carried out by making use of the Lagrange expansion [5, p. 125]. We show that

$$\sum_{n=0}^{\infty} c_n^{(\lambda+\mu n)} u^n/n! = \frac{A(z)(B(z))^{\lambda+1}}{B(z)-\mu z B'(z)}, \tag{1.15}$$

where

$$u = z(B(z))^{-\mu}. \tag{1.16}$$

We then make a number of applications of the general result. In addition to the polynomials of Laguerre and Jacobi we treat also the Bell polynomials and the Bernoulli polynomials of arbitrary order. We remark that in the case of certain polynomial sets it may be possible to identify the parameter λ of (1.10) with the "variable". For example, we show that the Hermite polynomial defined by

$$\sum_{n=0}^{\infty} H_n(x)\frac{z^n}{n!} = e^{2xz-z^2}$$

also satisfies

$$\sum_{n=0}^{\infty} H_n(x+ny)\frac{u^n}{n!} = \frac{e^{2xz-z^2}}{1-2yz}, \tag{1.17}$$

where $u = ze^{-yz}$. Also a generating function of this kind is obtained for a polynomial closely related to the Legendre polynomial and, more generally, the ultraspherical polynomial (see (5.11) and (5.13) below).

It should be noted that more than one parameter can be varied—as for example in the case of the Jacobi polynomial. Indeed, for the general Bell polynomial, we may have a countable number of parameters (see (7.5) below). We remark that the Hermite polynomial is a very special instance of a Bell polynomial.

2. Let $c_n^{(\lambda)}$ be defined by (1.12) and (1.13). Then by Taylor's theorem

$$c_n^{(\lambda)} = D^n\{A(x)(B(x))^{\lambda}\}_{x=0} \qquad \left(D \equiv \frac{d}{dx}\right),$$

so that

$$c_n^{(\lambda+\mu n)} = [D^n\{A(x)(B(x))^{\lambda+\mu n}\}]_{x=0}. \tag{2.1}$$

We now apply the Lagrange expansion [5, p. 125]. Put

$$u = z/(\phi(z)) \qquad (\phi(0) \neq 0), \tag{2.2}$$

where $\phi(z)$ is analytic about the origin. Let $f(z)$ denote an arbitrary function that is analytic about the origin. Then

$$\frac{f(z)}{1-u\phi'(z)} = \sum_{n=0}^{\infty} \frac{u^n}{n!}[D^n\{f(x)(\phi(x))^n\}]_{x=0}. \tag{2.3}$$

Comparison of (2.3) with (2.1) suggests that we take

$$f(x) = A(x)(B(x))^{\lambda}, \qquad \phi(x) = (B(x))^{\mu}. \tag{2.4}$$

Thus (2.2) becomes

$$u = z(B(z))^{-\mu}, \tag{2.5}$$

so that

$$1-u\phi'(z)=\frac{B(z)-\mu zB'(z)}{B(z)}.$$

Hence (2.3) yields

$$\frac{A(z)(B(z))^{\lambda+1}}{B(z)-\mu zB'(z)}=\sum_{n=0}^{\infty}c_n^{(\lambda+\mu n)}\frac{u^n}{n!}. \tag{2.6}$$

We may state the following

THEOREM 1. *Let $A(z)$, $B(z)$ be arbitrary functions that are analytic about the origin and such that*

$$A(0)=B(0)=1.$$

Define the coefficients $\{c_n^{(\lambda)}\}$ by means of

$$A(z)(B(z))^{\lambda}=\sum_{n=0}^{\infty}c_n^{(\lambda)}z^n/n!, \tag{2.7}$$

where λ is independent of z but otherwise arbitrary. Then, for arbitrary μ independent of z, the generating function

$$\sum_{n=0}^{\infty}c_n^{(\lambda+\mu n)}\frac{u^n}{n!}$$

satisfies (2.6), *where*

$$u=z(B(z))^{-\mu}. \tag{2.8}$$

For several parameters, if, for example, $c_n^{(\lambda,\mu)}$ is defined by

$$\sum_{n=0}^{\infty}c_n^{(\lambda,\mu)}z^n/n!=A(z)(B(z))^{\lambda}(C(z))^{\mu}, \tag{2.9}$$

where $A(z)$, $B(z)$, $C(z)$ are analytic in a neighborhood of origin and

$$A(0)=B(0)=C(0)=1,$$

then we have

$$\sum_{n=0}^{\infty}c_n^{(\lambda+\lambda'n,\mu+\mu'n)}\frac{u^n}{n!}=\frac{A(z)(B(z))^{\lambda}(C(z))^{\mu}}{1-z\{\lambda'[B'(z)/B(z)]+\mu'[C'(z)/C(z)]\}}, \tag{2.10}$$

where

$$u=z(B(z))^{-\lambda'}(C(z))^{-\mu'}.$$

A greater number of parameters are treated in the same way.

3. We shall now examine some special cases of (2.6). In the first place, for $\mu=0$, (2.6) reduces to (2.7), as is to be expected.

Next take

$$A(z)=\exp\left(\frac{-xz}{1-z}\right), \qquad B(z)=1-z \tag{3.1}$$

and

$$\lambda = -\alpha - 1, \qquad \mu = -\beta. \tag{3.2}$$

Thus (2.8) becomes

$$u = z(1-z)^{\beta}. \tag{3.3}$$

Clearly, by (1.13) and (1.1),

$$c_n^{(\lambda)} \to n! L_n^{(\alpha)}(x).$$

It follows that (2.6) reduces to

$$\sum_{n=0}^{\infty} L_n^{(\alpha+\beta n)}(x) u^n = \frac{(1-z)^{-\alpha-1} \exp\left(-xz(1-z)^{-1}\right)}{1-\beta z(1-z)^{-1}}. \tag{3.4}$$

Put $v = z/(1-z)$, so that

$$z = \frac{v}{1+v}, \quad 1-z = \frac{1}{1+v}, \quad 1-\beta z(1-z)^{-1} = 1-\beta v.$$

Then (3.4) becomes

$$\sum_{n=0}^{\infty} L_n^{(\alpha+\beta n)}(x) u^n = \frac{(1+v)^{\alpha+1}}{1-\beta v} e^{-xv}, \tag{3.5}$$

where

$$u = v(1+v)^{-\beta-1}. \tag{3.6}$$

Turning next to the Jacobi polynomials, we have by (1.9),

$$P_n^{(\alpha,\beta)}(x) = n! 2^{\alpha+\beta} [D_z^n \{\rho^{-1}(1-z+\rho)^{-\alpha}(1+z+\rho)^{-\beta}\}]_{z=0},$$

so that

$$P_n^{(\alpha+\gamma n, \beta+\delta n)}(x)$$
$$= n! 2^{\alpha+\beta+(\gamma+\delta)n} [D_z^n \{\rho^{-1}(1-z+\rho)^{-\alpha-\gamma n}(1+z+\rho)^{-\beta-\delta n}\}]_{z=0}.$$

It is convenient to apply (2.3) directly. We accordingly take

$$f(z) = 2^{\alpha+\beta}\rho^{-1}(1-z+\rho)^{-\alpha}(1+z+\rho)^{-\beta},$$
$$\phi(z) = 2^{\gamma+\delta}(1-z+\rho)^{-\gamma}(1+z+\rho)^{-\delta}$$

and

$$u = 2^{-\gamma-\delta} z(1-z+\rho)^{\gamma}(1+z+\rho)^{\delta}. \tag{3.7}$$

Substituting in (2.3) we get

$$\sum_{n=0}^{\infty} P_n^{(\alpha+\gamma n, \beta+\delta n)}(x) u^n = \frac{2^{\alpha+\beta}(1-z+\rho)^{-\alpha}(1+z+\rho)^{-\beta}}{\rho - z\{\gamma[(x-z+\rho)/(1-z+\rho)] + \delta[(x-z-\rho)/(1+z+\rho)]\}}. \tag{3.8}$$

Put

$$r = \tfrac{1}{2}(-1-z+\rho),$$
$$s = \tfrac{1}{2}(-1+z+\rho),$$

so that

$$\rho = 1+r+s,$$
$$z = s-r.$$

Then (3.8) becomes

$$\sum_{n=0}^{\infty} P_n^{(\alpha+\gamma n,\beta+\delta n)}(x)\{z(1+r)^{\gamma}(1+s)^{\delta}\}^n = \frac{r^{-\alpha}s^{-\beta}}{1+r+s+\frac{1}{2}(r-s)\{\gamma[(x+1+2r)/(1+r)]+\delta[(x-1-2s)/(1+s)]\}}. \tag{3.9}$$

Now put

$$t = z(1+r)^{\gamma}(1+s)^{\delta} = (s-r)(1+r)^{\gamma}(1+s)^{\delta}. \tag{3.10}$$

To show the equivalence of (3.9) and (1.10), it suffices to show that

$$1+r+s+\tfrac{1}{2}(r-s)\left\{\gamma\frac{x+1+2r}{1+r}+\delta\frac{x-1-2s}{1+s}\right\} = 1+(1+\gamma)r+(1+\delta)s \tag{3.11}$$

and

$$r = \tfrac{1}{2}(x+1)t(1+r)^{-\gamma}(1+s)^{-\delta-1},$$
$$s = -\tfrac{1}{2}(x-1)t(1+r)^{-\gamma-1}(1+s)^{-\delta}. \tag{3.12}$$

By (3.10), (3.12) reduces to

$$r = \tfrac{1}{2}(x+1)(s-r)(1+s)^{-1},$$
$$s = -\tfrac{1}{2}(x-1)(s-r)(1+r)^{-1}, \tag{3.12$'$}$$

while (3.11) is

$$\gamma\left(\tfrac{1}{2}(x-1)\frac{r-s}{1+r}-s\right)+\delta\left(\tfrac{1}{2}(x+1)\frac{r-s}{1+s}-r\right) = 0. \tag{3.11$'$}$$

Since (3.11′) is implied by (3.12′), the verification is completed.

4. The Hermite polynomial $H_n(x)$ may be defined by

$$\sum_{n=0}^{\infty} H_n(x)\frac{z^n}{n!} = e^{2xz-z^2}, \tag{4.1}$$

so that

$$H_n(x) = [D_z^n e^{2xz-z^2}]_{z=0}.$$

Replacing x by $x+ny$ this becomes

$$H_n(x+ny) = [D_z^n e^{2xz-z^2+2nyz}]_{z=0}. \tag{4.2}$$

We accordingly take

$$f(z) = e^{2xz-z^2},$$
$$\phi(z) = e^{2yz}$$

and

$$u = z\,e^{-2yz}. \tag{4.3}$$

Hence (2.3) yields

$$\sum_{n=0}^{\infty} H_n(x+ny)\frac{u^n}{n!} = \frac{e^{2xz-z^2}}{1-2yz}. \tag{4.4}$$

This result is presumably new.

It is of some interest to give a direct proof of (4.4). Since

$$H_k(x+ky) = \sum_{n=0}^{\infty} \binom{k}{n}(2ky)^{k-n}H_n(x),$$

it follows that

$$\sum_{k=0}^{\infty} H_k(x+ky)\frac{u^k}{k!} = \sum_{k=0}^{\infty}\frac{z^k}{k!}e^{-2kyz}\sum_{n=0}^{k}\binom{k}{n}(2ky)^{k-n}H_n(x). \tag{4.5}$$

By (4.1) the right-hand side of (4.4) is equal to

$$\frac{1}{1-2yz}\sum_{n=0}^{\infty} H_n(x)\frac{z^n}{n!}. \tag{4.6}$$

Comparing (4.6) with (4.5) it is evident that (4.4) is equivalent to

$$\frac{1}{1-2yz}\frac{z^n}{n!} = \sum_{k=n}^{\infty}\frac{z^k}{k!}e^{-2kyz}\binom{k}{n}(2ky)^{k-n},$$

that is, to

$$\frac{e^{2nyz}}{1-2yz} = \sum_{k=0}^{\infty}\frac{(k+n)^k}{k!}(2yz\,e^{-2yz})^k. \tag{4.7}$$

Replacing $2yz$ by z and n by α this reduces to the known identity [5, p. 126, no. 214]

$$\frac{e^{\alpha z}}{1-z} = \sum_{k=0}^{\infty}\frac{(k+\alpha)^k}{k!}(z\,e^{-z})^k \tag{4.8}$$

which holds for all α.

5. For the Laguerre polynomial $L_n^{(\alpha)}(x)$ one can also obtain a generating function analogous to (4.4). Indeed one can now get a "mixed" generating function for $L_n^{(\alpha+\beta n)}(x+ny)$.

By (1.1) we have

$$L_n^{(\alpha)}(x) = n!\left[D_z^n\left\{(1-z)^{-\alpha-1}\exp\left(\frac{-xz}{1-z}\right)\right\}\right]_{z=0}.$$

Hence

$$L_n^{(\alpha+\beta n)}(x+ny) = n!\left[D_z^n\left\{(1-z)^{-\alpha-\beta n-1}\exp\left(\frac{-(x+ny)z}{1-z}\right)\right\}\right]_{z=0}. \tag{5.1}$$

We accordingly take

$$f(z)=(1-z)^{-\alpha-1}\exp\left(\frac{-xz}{1-z}\right),$$

$$\phi(z)=(1-z)^{-\beta}\exp\left(\frac{-yz}{1-z}\right).$$

Thus

$$\frac{\phi'(z)}{\phi(z)}=\frac{\beta}{1-z}-\frac{y}{(1-z)^2}$$

so that (2.3) yields

$$\sum_{n=0}^{\infty} L_n^{(\alpha+\beta n)}(x+ny)u^n=\frac{(1-z)^{-\alpha-1}\exp\left[-xz/(1-z)\right]}{1-[\beta z/(1-z)]+(yz/(1-z)^2]} \tag{5.2}$$

with

$$u=z(1-z)^{\beta}\exp\left(\frac{yz}{1-z}\right). \tag{5.3}$$

For $y=0$, (5.2) reduces to (3.4). For $\beta=0$, (5.2) becomes

$$\sum_{n=0}^{\infty} L_n^{(\alpha)}(x+ny)u^n=\frac{(1-z)^{-\alpha-1}\exp\left[-xz/(1-z)\right]}{1+[yz/(1-z)^2]}, \tag{5.4}$$

where u again is given by (5.3). A direct proof of (5.4) is not difficult.

If we put $v=z/(1-z)$, (5.2) becomes

$$\sum_{n=0}^{\infty} L_n^{(\alpha+\beta n)}(x+ny)u^n=\frac{(1+v)^{\alpha+1}\exp\left(-xv\right)}{1-\rho v+yv(1+v)}, \tag{5.5}$$

where now

$$u=v(1+v)^{-\beta-1}\exp\left(yv\right). \tag{5.6}$$

A generating function like (5.4) can also be obtained for a modified Legendre polynomial. Rainville [6] has noted that the polynomial

$$\phi_n(x)=(1-x^2)^{n/2}P_n((1-x^2)^{1/2})$$

satisfies

$$e^tI_0(tx)=\sum_{n=0}^{\infty}\phi_n(x)t^n/n!,$$

where

$$I_0(z)=\sum_{n=0}^{\infty}\frac{(z/2)^{2n}}{(n!)^2}.$$

Replacing x by $1/x$ and t by xz, it follows that

$$e^{xz}I_0(z)=\sum_{n=0}^{\infty}\bar{\phi}_n(x)z^n/n!, \tag{5.7}$$

where

$$\bar{\phi}_n(x) = x^n \phi_n\left(\frac{1}{x}\right) = (x^2-1)^{n/2} P_n\left(\frac{x}{\sqrt{x^2-1}}\right). \tag{5.8}$$

It follows from (5.8) and the familiar generating function

$$\sum_{n=0}^{\infty} P_n(x) z^n = (1-2xz+z^2)^{-1/2}$$

that

$$\sum_{n=0}^{\infty} \bar{\phi}_n(x) z^n = \{1-2xz+(x^2-1)z^2\}^{-1/2}. \tag{5.9}$$

Expanding the right-hand side, we get

$$\bar{\phi}_n(x) = \sum_{2r \leqq n} \binom{n+1}{2r+1} \binom{2r}{r} 2^{-2r} x^{n-2r}. \tag{5.10}$$

By (5.7),

$$\bar{\phi}_n(x+ny) = n![D_z^n\{e^{(x+ny)z} I_0(z)\}]_{z=0}$$

and therefore

$$\sum_{n=0}^{\infty} \bar{\phi}_n(x+ny) \frac{u^n}{n!} = \frac{e^{xz} I_0(z)}{1-yz}, \tag{5.11}$$

where

$$u = ze^{-yz}. \tag{5.12}$$

The generating function (5.11) is readily extended to

$$\sum_{n=0}^{\infty} \bar{\phi}_n^{(\alpha)}(x+ny) \frac{u^n}{n!} = \frac{\Gamma(\alpha+\frac{1}{2})\, e^{xz} (z/2)^{-\alpha+1/2} I_{\alpha-1/2}(z)}{1-yz}, \tag{5.13}$$

where

$$\phi_n^{(\alpha)}(x) = x^n \phi_n^{(\alpha)}\left(\frac{1}{x}\right) = (x^2-1)^{n/2} P_n^{(\alpha)}\left(\frac{x}{\sqrt{x^2-1}}\right) \tag{5.14}$$

and

$$I_\alpha(z) = \sum_{n=0}^{\infty} \frac{(z/2)^{\alpha+2n}}{n!\Gamma(\alpha+n+1)};$$

$P_n^{(\alpha)}(x)$ denotes the ultraspherical polynomial defined by

$$\sum_{n=0}^{\infty} P_n^{(\alpha)}(x) z^n = (1-2xz+z^2)^{-\alpha}.$$

6. We now construct some examples of a different nature. Let

$$S(n,k)=\frac{1}{k!}\sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j}j^n \tag{6.1}$$

denote the Stirling number of the second kind and put

$$A_n(x)=\sum_{k=0}^{n}S(n,k)x^k. \tag{6.2}$$

It is well known (and easily verified) that

$$\sum_{n=0}^{\infty}A_n(x)\frac{z^n}{n!}=\exp(x(e^z-1)). \tag{6.3}$$

It follows from (6.3) that

$$A_n(x)=n![D_z^n\exp(x(e^z-1))]_{z=0},$$

so that

$$A_n(x+ny)=n![D_z^n\exp((x+ny)(e^z-1))]_{z=0}. \tag{6.4}$$

We therefore take

$$f(z)=\exp(x(e^z-1)),$$
$$\phi(z)=\exp(y(e^z-1)).$$

Thus, by (2.3), we have

$$\sum_{n=0}^{\infty}A_n(x+ny)\frac{u^n}{n!}=\frac{\exp(x(e^z-1))}{1-yz\,e^z}, \tag{6.5}$$

where

$$u=z\exp(-y(e^z-1)). \tag{6.6}$$

Making use of (6.3), (6.5) becomes

$$\sum_{n=0}^{\infty}A_n(x+ny)\frac{u^n}{n!}=(1-yz\,e^z)^{-1}\sum_{k=0}^{\infty}A_k(x)\frac{z^k}{k!}. \tag{6.7}$$

Also it follows from (6.3) that

$$A_n(x+y)=\sum_{k=0}^{n}\binom{n}{k}A_k(x)A_{n-k}(y). \tag{6.8}$$

Thus the LHS (left-hand side) of (6.7) is equal to

$$\sum_{n=0}^{\infty}\frac{u^n}{n!}\sum_{k=0}^{n}\binom{n}{k}A_k(x)A_{n-k}(ny)=\sum_{k=0}^{\infty}A_k(x)\frac{u^k}{k!}\sum_{n=0}^{\infty}A_n((n+k)y)\frac{u^n}{n!}.$$

Since, by (6.6), u is independent of x, it follows that (6.7) is equivalent to

$$\sum_{n=0}^{\infty}A_n((n+k)y)\frac{u^n}{n!}=\left(\frac{z}{u}\right)^k(1-yz\,e^z)^{-1}. \tag{6.9}$$

By (6.6) and (6.3)

$$\left(\frac{z}{u}\right)^k = \exp(ky(e^z-1)) = \sum_{n=0}^{\infty} A_n(ky)\frac{z^n}{n!}$$

and so (6.9) becomes

$$(6.10)\quad \sum_{n=0}^{\infty} A_n((n+k)y)\frac{u^n}{n!} = (1-yz\,e^z)^{-1}\sum_{n=0}^{\infty} A_n(ky)\frac{z^n}{n!} \qquad (k=0,1,2,\cdots).$$

Presumably (6.10) holds for all k.

In particular, for $k=0$, (6.10) reduces to

$$(6.11)\quad \sum_{n=0}^{\infty} A_n(ny)\frac{u^n}{n!} = (1-yz\,e^z)^{-1}.$$

In the next place, the LHS of (6.10), by (6.6) and (6.3), is equal to

$$\sum_{n=0}^{\infty} A_n((n+k)y)\frac{z^n}{n!}\sum_{j=0}^{\infty} A_j(-ny)\frac{z^j}{j!} = \sum_{m=0}^{\infty}\frac{z^m}{m!}\sum_{n=0}^{m}\binom{m}{n}A_n((n+k)y)A_{m-n}(-ny).$$

Hence we get the rather curious identity

$$(6.12)\quad \sum_{m=0}^{\infty}\frac{z^m}{m!}\sum_{n=0}^{m}\binom{m}{n}A_n((n+k)y)A_{m-n}(-ny) = (1-yz\,e^z)^{-1}\sum_{m=0}^{\infty} A_m(ky)\frac{z^m}{m!} \qquad (k=0,1,2,\cdots).$$

Needless to say, (6.8) does not apply to the sum

$$\sum_{n=0}^{m}\binom{m}{n}A_n((n+k)y)A_{m-n}(-ny).$$

This sum suggests a connection with Abel type sums [7, Chap. 1] that we shall not pursue.

7. The polynomial $A_n(x)$ is sometimes called a single-variable Bell polynomial. The general Bell polynomial [8, Chap. 2] may be defined by

$$(7.1)\quad \sum_{n=0}^{\infty} A_n(x_1, x_2, \cdots, x_n)\frac{z^n}{n!} = \exp\{E(\mathbf{x}, z)\},$$

where

$$(7.2)\quad E(\mathbf{x}, z) = \sum_{k=1}^{\infty} x_k\frac{z^k}{k!}.$$

For brevity we shall write

$$(7.3)\quad A_n(\mathbf{x}) = A_n(x_1, x_2, \cdots, x_n).$$

By (7.1) and (7.3)

$$A_n(\mathbf{x}) = [D_z^n \exp\{E(\mathbf{x}, z)\}]_{z=0},$$

so that

$$A_n(\mathbf{x}+n\mathbf{y}) = [D_z^n \exp\{E(\mathbf{x}, z)+nE(\mathbf{y}, z)\}]_{z=0}. \tag{7.4}$$

We therefore take

$$f(z) = \exp\{E(\mathbf{x}, z)\},$$
$$\phi(z) = \exp\{E(\mathbf{y}, z)\}.$$

It follows that

$$\sum_{n=0}^{\infty} A_n(\mathbf{x}+n\mathbf{y})\frac{u^n}{n!} = \frac{\exp\{E(\mathbf{x}, z)\}}{1-zE'(\mathbf{y}, z)}, \tag{7.5}$$

where $E'(\mathbf{y}, z) = (\partial/\partial z)E(\mathbf{y}, z)$ and

$$u = z \exp\{-E(\mathbf{y}, z)\}. \tag{7.6}$$

Clearly, when

$$x = x_1 = x_2 = x_3 = \cdots,$$

$A_n(\mathbf{x})$ reduces to $A_n(x)$ and (7.5) reduces to (6.5). Moreover, (6.8) generalizes to

$$A_n(\mathbf{x}+\mathbf{y}) = \sum_{k=0}^{n} \binom{n}{k} A_k(\mathbf{x})A_{n-k}(\mathbf{y}) \tag{7.7}$$

and the remaining formulas of § 6 can also be generalized. A like remark applies to the formulas of § 4 since, by (4.1) and (7.1),

$$H_n(x) = A_n(2x, -2, 0, \cdots, 0).$$

A special case of $A_n(\mathbf{x})$ of some interest is defined by

$$\sum_{n=0}^{\infty} A_n(x_1, x_2)\frac{z^n}{n!} = \exp\{x_1 \sinh z + x_2(\cosh z - 1)\}. \tag{7.8}$$

It follows that

$$A_n(x_1, x_2) = A_n(x_1, x_2, x_1, x_2, \cdots). \tag{7.9}$$

Formula (7.5) becomes

$$\sum_{n=0}^{\infty} A_n(x_1+ny_1, x_2+ny_2)\frac{u^n}{n!} = \frac{\exp\{x_1 \sinh z + x_2(\cosh z - 1)\}}{1-z(y_1 \cosh z + y_2 \sinh z)}, \tag{7.10}$$

where now

$$u = z \exp\{-y_1 \sinh z - y_2(\cosh z - 1)\}. \tag{7.11}$$

8. The Stirling numbers of the first kind are defined by

$$(x)_n = x(x+1)\cdots(x+n-1) = \sum_{k=0}^{n} S_1(n, k)x^k. \tag{8.1}$$

Since

$$\sum_{n=0}^{\infty}(x)_n\frac{z^n}{n!}=(1-z)^{-x}=\exp\left\{x\sum_{k=1}^{\infty}\frac{z^k}{k}\right\}, \tag{8.2}$$

we again have a special case of a Bell polynomial. Thus specializing (7.5) or directly we have

$$\sum_{n=0}^{\infty}(x+ny)_n\frac{u^n}{n!}=\frac{(1-z)^{-x}}{1-(yz/(1-z))} \tag{8.3}$$

where

$$u=z(1-z)^y. \tag{8.4}$$

Formula (8.3) can also be thought of as a special case of the Laguerre polynomial identity (5.2).

9. A set of polynomials $\{f_n(x)\}$ is called an Appell set if

$$f_n'(x)=nf_{n-1}(x) \qquad (n=0,1,2,\cdots). \tag{9.1}$$

This is equivalent to

$$\sum_{n=0}^{\infty}f_n(x)\frac{z^n}{n!}=A(z)\,e^{xz}, \qquad A(0)\neq 0, \tag{9.2}$$

where $A(z)$ is an arbitrary power series in z.

Since

$$f_n(x)=[D_z^n\{A(z)\,e^{xz}\}]_{z=0}$$

and

$$f_n(x+ny)=[D_z^n\{A(z)\,e^{(x+ny)z}\}]_{z=0},$$

we take

$$f(z)=A(z)\,e^{xz}, \qquad \phi(z)=e^{yz}.$$

Therefore we have

$$\sum_{n=0}^{\infty}f_n(x+ny)\frac{u^n}{n!}=\frac{A(z)\,e^{xz}}{1-yz}, \tag{9.3}$$

where

$$u=z\,e^{-yz}. \tag{9.4}$$

Certain special Appell sets are of particular interest. The Hermite polynomials (with a slight change in notation) furnish an important example. Other instances that have received considerable study are the Bernoulli and Euler polynomials defined by [4, Chap. 2]

$$\sum_{n=0}^{\infty}B_n(x)\frac{z^n}{n!}=\frac{z\,e^{xz}}{e^z-1} \tag{9.5}$$

and

$$\sum_{n=0}^{\infty} E_n(x)\frac{z^n}{n!}=\frac{2e^{xz}}{e^z+1}, \tag{9.6}$$

respectively.

Specializing (9.3) we get

$$\sum_{n=0}^{\infty} B_n(x+ny)\frac{u^n}{n!}=\frac{1}{1-yz}\,\frac{ze^{xz}}{e^z-1} \tag{9.7}$$

and

$$\sum_{n=0}^{\infty} E_n(x+ny)\frac{u^n}{n!}=\frac{1}{1-yz}\,\frac{2e^{xz}}{e^z+1}; \tag{9.8}$$

in each case

$$u=ze^{-yz}. \tag{9.9}$$

However we can go considerably further. Bernoulli and Euler polynomials of higher order are defined by [4, Chap. 6]

$$\sum_{n=0}^{\infty} B_n^{(\alpha)}(x)\frac{z^n}{n!}=\left(\frac{z}{e^z-1}\right)^{\alpha} e^{xz} \tag{9.10}$$

and

$$\sum_{n=0}^{\infty} E_n^{(\alpha)}(x)\frac{z^n}{n!}=\left(\frac{2}{e^z+1}\right)^{\alpha} e^{xz}, \tag{9.11}$$

respectively.

For (9.10) we take

$$f(z)=\left(\frac{z}{e^z-1}\right)^{\alpha} e^{xz},$$

$$\phi(z)=\left(\frac{z}{e^z-1}\right)^{\beta} e^{yz}.$$

It follows that

$$\sum_{n=0}^{\infty} B_n^{(\alpha+n\beta)}(x+ny)\frac{u^n}{n!}=\frac{[z/(e^z-1)]^{\alpha}e^{xz}}{1-yz-\beta+[\beta z e^z/(e^z-1)]}, \tag{9.12}$$

where

$$u=z\left(\frac{z}{e^z-1}\right)^{\beta} e^{-yz}. \tag{9.13}$$

Similarly

$$\sum_{n=0}^{\infty} E_n^{(\alpha+n\beta)}(x+ny)\frac{u^n}{n!}=\frac{[2/(e^z+1)]^{\alpha}e^{xz}}{1-yz+[\beta z e^z/(e^z+1)]}, \tag{9.14}$$

where

$$u = z\left(\frac{2}{e^z+1}\right)^{-\beta} e^{-yz}. \tag{9.15}$$

REFERENCES

[1] J. W. BROWN, *New generating functions for classical polynomials*, Proc. Amer. Math. Soc., 21 (1969), pp. 263–268.

[2] ———, *On zero type sets of Laguerre polynomials*, Duke Math. J.,35 (1968), pp. 821–823.

[3] L. CARLITZ, *Some generating functions for Laguerre polynomials*, Ibid., 35 (1968), pp. 825–827.

[4] N. E. NÖRLUND, *Vorlesungen über Differzenrechnung*, Springer, Berlin, 1923.

[5] G. PÓLYA and G. SZEGÖ, *Aufgaben und Lehrsätze aus der Analysis, I*, Springer, Berlin, 1925.

[6] E. D. RAINVILLE, *Notes on Legendre polynomials*, Bull. Amer. Math. Soc., 51 (1945), pp. 268–271.

[7] J. RIORDAN, *Combinatorial Identities*, John Wiley, New York, 1968.

[8] ———, *An Introduction to Combinatorial Analysis*, John Wiley, New York, 1958.

[9] H. M. SRIVASTAVA AND J. P. SINGHAL, *New generating functions for Jacobi and related polynomials*, J. Math. Anal. Appl., 41 (1973), pp. 748–752.

[10] ———, *A unified treatment of certain classical polynomials*, Math. Comput., 26 (1972), pp. 969–975.

[11] G. SZEGÖ, *Orthogonal Polynomials*, rev. ed., American Mathematical Society, New York, 1959.

SIAM J. MATH. ANAL.
Vol. 8, No. 3, May 1977

A NOTE ON STIELTJES MOMENT SEQUENCES*

LEOPOLD SIMAR†

Abstract. In this note we derive some inequalities on Stieltjes moments of different distribution functions having some moments in common in the sequence and show how these inequalities can be used in the comparison of some compound processes.

1. Introduction. Analyzing the Stieltjes moment sequence of a distribution function we know from the Stieltjes moment problem (Shohat and Tamarkin [2]) whether its spectrum is reducible or not to a finite set of points. (The order of a distribution function is the number of points of its support.)

In this note we give two important relations between the moment sequences of distribution functions of different order having some common moments in their sequence, and we emphasize the importance of these theorems in the comparison of compound processes.

Consider a nondecreasing positive ($\geqq 0$) function $F(x)$ defined on $[0, \infty)$. The Stieltjes moment sequence of F is the sequence μ_k, $k = 0, 1, 2, \cdots$, where:

$$\mu_k = \int_0^\infty x^k \, dF(x).$$

We restrict attention throughout this note to those $F(x)$ whose moments all exist.

2. Two theorems.

THEOREM 1. *Let $F(x)$ be a nondecreasing, positive ($\geqq 0$) function defined on $[0, \infty)$ with Stieltjes moments $\mu_0, \mu_1, \cdots$, which all exist. Assume $F(x)$ is not reducible to a finite set of points.*

Let there be a nondecreasing, positive ($\geqq 0$) function $F^n(x)$ that is defined on $[0, \infty)$ and is purely discrete with n points of increase and with Stieltjes moments $\nu_0, \nu_1, \cdots$, which all exist and such that $\nu_k = \mu_k$ for $k = 0, 1, 2, \cdots, 2n-1$. Then, $\nu_k < \mu_k$ for all $k \geqq 2n$.

THEOREM 2. *Let $F^n(x)$ and $F^m(x)$ be two nondecreasing, positive ($\geqq 0$) functions that are defined on $[0, \infty)$ and are purely discrete of order n and m respectively, with $n < m$.*

Let $(\nu_0, \nu_1, \cdots)$ be the Stieltjes moment sequence of $F^n(x)$ and $(\mu_0, \mu_1, \cdots)$ that of $F^m(x)$.

If the sequences are such that $\nu_j = \mu_j$, $j = 0, 1, \cdots, 2n-1$, then $\nu_j \leqq \mu_j$ for all j. Strict inequality holds for at least $j = 2n$.

Those results are applications of the theory of principal representations for special Chebyshev systems. (See Karlin and Studden [1, Chap. 2].)

3. Comments. It may be noted that Theorems 1 and 2 are widely applicable because the functions need not be normalized to 1.

As an example, consider a compound process defined as

$$P(\tilde{\theta} = j) = \int_0^\infty f_j(x) \, dG(x) \tag{1}$$

* Received by the editors April 1, 1975, and in revised form February 16, 1976.

† Center for Operations Research and Econometrics, Université Catholique de Louvain, 3030 Heverlee, Belgium.

where G is an arbitrary probability distribution function. If $f_j(x)$ can be factorized as $x^j \cdot g(x) \cdot h(j)$, we can define

$$\mu_j = \int_0^\infty x^j \, dF(x) \tag{2}$$

where

$$F(x) = \int_0^x g(t) \, dG(t). \tag{3}$$

Since $\sum_{j=0}^{\infty} P(\tilde{\theta} = j) = 1$, the moments of $F(x)$ defined by (2) and (3) all exist. Therefore we can compare compound processes with different compounding distributions G using the two theorems since $P(\tilde{\theta} = j) = h(j) \cdot \mu_j$. This was done in Simar [3] for compound Poisson processes ($f_j(x) = x^j e^{-x}/j!$) in order to compute a maximum to the likelihood function of a given sample on $\tilde{\theta}$.

REFERENCES

[1] S. KARLIN AND W. J. STUDDEN, *Tchebycheff Systems: With Applications in Analysis and Statistics*, Interscience, New York, 1966.

[2] J. A. SHOHAT AND J. D. TAMARKIN, *The Problem of Moments*, Mathematical Surveys No. 1, American Mathematical Society, New York, 1943.

[3] L. SIMAR, *Maximum likelihood estimation of a compound Poisson process*, Ann. Statist., 4 (1976).

SIAM J. MATH. ANAL.
Vol. 8, No. 3, May 1977

THE ADDITION FORMULA FOR LAGUERRE POLYNOMIALS*

TOM KOORNWINDER†

Abstract. Bateman's addition formula for Laguerre polynomials of order zero is generalized to the case of order $\alpha > 0$. The result is obtained as a limit case of the addition formula for disk polynomials.

1. Introduction. This note answers a question posed by Askey [1, p. 83]. An addition formula for Laguerre polynomials $L_n^\alpha(x)$ ($\alpha > 0$) will be derived which reduces to Bateman's addition formula [2, p. 457] for $\alpha \downarrow 0$ and which leads by integration to Watson's integral representation [16] for the product $L_n^\alpha(x)L_n^\alpha(y)$ of two Laguerre polynomials.

This addition formula turns out to be a limit case of the addition formula for the so-called disk polynomials which are orthogonal polynomials in two variables on the unit disk. If r, ψ are polar coordinates on the unit disk then the addition formula is an orthogonal expansion of

$$L_n^\alpha(x^2+y^2-2xyr\cos\psi)\exp(xyre^{i\psi})$$

in terms of disk polynomials of order $\alpha-1$ depending on r and ψ.

2. Preliminaries. Let Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$, Laguerre polynomials $L_n^\alpha(x)$ and Bessel functions $J_\alpha(x)$ be as defined in Erdélyi et al. [6]. It will be convenient to use the slightly different functions $R_n^{(\alpha,\beta)}(x)$, $\mathcal{L}_n^\alpha(x)$ and $\mathcal{J}_\alpha(x)$, respectively, which are defined by

$$R_n^{(\alpha,\beta)}(x) = P_n^{(\alpha,\beta)}(x)/P_n^{(\alpha,\beta)}(1),$$
$$\mathcal{L}_n^\alpha(x) = e^{-(1/2)x}L_n^\alpha(x)/L_n^\alpha(0),$$
$$\mathcal{J}_\alpha(x) = \Gamma(\alpha+1)(\tfrac{1}{2}x)^{-\alpha}J_\alpha(x).$$

Laguerre polynomials are a confluent case of Jacobi polynomials by the limit formula

$$e^{(1/2)x}\mathcal{L}_n^\alpha(x) = \lim_{\beta\to\infty} R_n^{(\alpha,\beta)}(1-2\beta^{-1}x), \tag{2.1}$$

which holds uniformly for x in bounded sets. The functions $\mathcal{L}_n^\alpha(x)$ satisfy the inequality

$$|\mathcal{L}_n^\alpha(x)| \leqq 1 \qquad (\alpha \geqq 0, x \geqq 0), \tag{2.2}$$

(cf. Erdélyi et al. [6, 10.18(14)]).

Let $z = x+iy$, $\bar{z} = x-iy$, $x, y \in \mathbb{R}$. For $\alpha > -1$ and for nonnegative integers m, n the so-called disk polynomials $R_{m,n}^\alpha(z)$ are defined in terms of Jacobi polynomials by

$$R_{m,n}^\alpha(z) = \begin{cases} R_n^{(\alpha,m-n)}(2z\bar{z}-1)z^{m-n} & \text{if } m \geqq n, \\ R_m^{(\alpha,n-m)}(2z\bar{z}-1)\bar{z}^{n-m} & \text{if } m \leqq n. \end{cases} \tag{2.3}$$

* Received by the editors April 5, 1976, and in revised form July 16, 1976.
† Mathematical Centrum, Amsterdam, the Netherlands.

It is easily proved that the polynomials $R^{\alpha}_{m,n}(z)$ are orthogonal polynomials of degree $m+n$ in x and y on the unit disk with respect to the weight function $(1-x^2-y^2)^{\alpha}$. In fact, disk polynomials are characterized by the following properties:

(i) $R^{\alpha}_{m,n}(z) = \text{const. } z^m \bar{z}^n + \text{polynomial of degree less than } m+n$;

(ii) $\iint_{x^2+y^2<1} R^{\alpha}_{m,n}(x+iy)\overline{p(x,y)}(1-x^2-y^2)^{\alpha}\,dx\,dy = 0$ for every polynomial $p(x,y)$ of degree less than $m+n$;

(iii) $R^{\alpha}_{m,n}(1) = 1$.

These polynomials were first studied by Zernike and Brinkman [17]. The notation $R^{\alpha}_{m,n}(z)$ was introduced by the author [7, p. 18].

It can be proved that $|R^{\alpha}_{m,n}(z)| \leqq 1$ if $\alpha \geqq 0$, $|z| \leqq 1$. However, we shall only need the estimate

$$|R^{\alpha}_{m,n}(z)| = \mathcal{O}(m^n) \quad \text{for } m \to \infty, \tag{2.4}$$

uniformly for $|z| \leqq 1$, where $\alpha > -1$ and n are fixed. This estimate follows from Szegö [14, (7.32.2)] by using (2.3).

3. The addition formula for Laguerre polynomials. Let $\alpha > 0$. The formula

$$\begin{aligned} &R^{\alpha}_{m,n}(\cos\theta_1 e^{i\phi_1}\cos\theta_2 e^{i\phi_2} + \sin\theta_1 \sin\theta_2 r e^{i\psi}) \\ &= \sum_{k=0}^{m}\sum_{l=0}^{n} \frac{\alpha}{\alpha+k+l}\binom{m}{k}\binom{n}{l}\frac{(\alpha+n+1)_k(\alpha+m+1)_l}{(\alpha+l)_k(\alpha+k)_l} \\ &\cdot (\sin\theta_1)^{k+l} R^{\alpha+k+l}_{m-k,n-l}(\cos\theta_1 e^{i\phi_1}) \\ &\cdot (\sin\theta_2)^{k+l} R^{\alpha+k+l}_{m-k,n-l}(\cos\theta_2 e^{i\phi_2}) R^{\alpha-1}_{k,l}(r e^{i\psi}) \end{aligned} \tag{3.1}$$

is called the addition formula for disk polynomials; cf. Šapiro [13, (1, 21)] and Koornwinder [8, (5.4)]. For $\alpha = 1, 2, 3, \cdots$ both authors independently obtained this formula by interpreting disk polynomials $R^{\alpha}_{m,n}(z)$ as spherical functions on the homogeneous space $SU(\alpha+2)/SU(\alpha+1)$. Since both sides of (3.1) are rational functions in α, the case of general α then follows by analytic continuation.

By putting $\phi_1 = \phi_2 = 0$, $x = \sin\theta_1$, $y = \sin\theta_2$ in (3.1) and by substituting (2.3) in (3.1) we obtain for $m \geqq n$, $\alpha > 0$, $0 \leqq x \leqq 1$, $0 \leqq y \leqq 1$:

$$\begin{aligned} &R_n^{(\alpha,m-n)}(2(1-x^2)(1-y^2)+2x^2y^2r^2+4xy(1-x^2)^{1/2}(1-y^2)^{1/2}r\cos\psi - 1) \\ &\cdot ((1-x^2)^{1/2}(1-y^2)^{1/2} + xyr e^{i\psi})^{m-n} \\ &= \sum_{k=0}^{m}\sum_{l=0}^{n} \frac{\alpha}{\alpha+k+l}\binom{m}{k}\binom{n}{l}\frac{(\alpha+n+1)(\alpha+m+1)_l}{(\alpha+l)_k(\alpha+k)_l} \\ &\cdot x^{k+l} R^{(\alpha+k+l,|m-n-k+l|)}_{(m-k)\wedge(n-1)}(1-2x^2)(1-x^2)^{(1/2)|m-n-k+l|} \\ &\cdot y^{k+l} R^{(\alpha+k+l,|m-n-k+l|)}_{(m-k)\wedge(n-l)}(1-2y^2)(1-y^2)^{(1/2)|m-n-k+l|} R^{\alpha-1}_{k,l}(r e^{i\psi}). \end{aligned} \tag{3.2}$$

Here $m \wedge n$ denotes the minimum of m and n. Both in (3.1) and (3.2) the right hand side is an orthogonal expansion of the left hand side in terms of disk polynomials $R^{\alpha-1}_{k,l}(r e^{i\psi})$.

Let us next replace x by $m^{-1/2}x$ and y by $m^{-1/2}y$ in (3.2). Denote this new formula by (3.2′) and let $m \to \infty$. First we calculate the formal limit case of (3.2′) by

taking termwise limits. Using (2.1) we obtain

$$\begin{aligned}&\mathscr{L}_n^\alpha(x^2+y^2-2xyr\cos\psi)\exp(ixyr\sin\psi)\\&\quad=\sum_{k=0}^{\infty}\sum_{l=0}^{n}\frac{\alpha}{\alpha+k+l}\binom{n}{l}\frac{(\alpha+n+1)}{k!(\alpha+l)_k(\alpha+k)_l}\\&\qquad\cdot x^{k+l}\mathscr{L}_{n-l}^{\alpha+k+l}(x^2)y^{k+l}\mathscr{L}_{n-l}^{\alpha+k+l}(y^2)R_{k,l}^{\alpha-1}(re^{i\psi}),\end{aligned}\tag{3.3}$$

where $x\geqq 0, y\geqq 0, 0\leqq r\leqq 1, 0\leqq\psi<2\pi, \alpha>0, n=0,1,2,\cdots$. For fixed x, y, α, n the convergence of the left hand side of (3.2′) to the left hand side of (3.3) is uniform in r and ψ. Denote the right hand side by

$$\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}c_{k,l}(x,y,\alpha,n)R_{k,l}^{\alpha-1}(re^{i\psi}),$$

where $c_{k,l}\equiv 0$ if $l>n$. Then the coefficients $c_{k,l}$ denote the Fourier coefficients of the left hand side with respect to the orthogonal functions $R_{k,l}^{\alpha-1}(re^{i\psi})$. We shall prove that this Fourier series uniformly converges in r and ψ. Then the identity (3.3) actually holds.

Let α and n be fixed and let x and y be in bounded sets. Then, by (2.2) and (2.4) there is a constant $M>0$ such that

$$|c_{k,l}(x,y,\alpha,n)R_{k,l}^{\alpha-1}(re^{i\psi})|\leqq M^k/k!,$$

uniformly in r and ψ. Hence the Fourier series is uniformly convergent in r and ψ.

Integration of (3.3) gives the product formula

$$\begin{aligned}\mathscr{L}_n^\alpha(x^2)\mathscr{L}_n^\alpha(y^2)=2\alpha\pi^{-1}\int_0^1\int_0^\pi&\mathscr{L}_n^\alpha(x^2+y^2+2xyr\cos\psi)\\&\cdot\cos(xyr\sin\psi)r(1-r^2)^{\alpha-1}\,dr\,d\psi\\&(x,y\geqq 0,\alpha>0).\end{aligned}\tag{3.4}$$

By putting $r\cos\psi=\cos\theta$, $r\sin\psi=\sin\theta\cos\psi$ in (3.4) and by substituting Poisson's integral representation for Bessel functions we obtain

$$\begin{aligned}\mathscr{L}_n^\alpha(x^2)\mathscr{L}_n^\alpha(y^2)=\frac{\Gamma(\alpha+1)}{\Gamma(\alpha+\frac12)\Gamma(\frac12)}\int_0^\pi&\mathscr{L}_n^\alpha(x^2+y^2+2xy\cos\theta)\\&\cdot\mathscr{J}_{\alpha-1/2}(xy\sin\theta)(\sin\theta)^{2\alpha}\,d\theta\\&(x,y\geqq 0,\alpha>-\tfrac12).\end{aligned}\tag{3.5}$$

The case $-\frac12<\alpha\leqq 0$ follows by analytic continuation. This formula is due to Watson [16]. Askey [1, pp. 82, 83] applied this product formula to define a convolution structure for Laguerre series, thus extending earlier results of McCully [10] for the case $\alpha=0$. However, this convolution structure is not positive and it is not defined for all L^1-functions.

If we put $r=1$ in (3.3) and let $\alpha\downarrow 0$ then we obtain the addition formula

$$\begin{aligned}&\mathscr{L}_n^0(x^2+y^2-2xy\cos\psi)\exp(ixy\sin\psi)\\&\quad=\sum_{k=0}^{\infty}\frac{1}{k!}\binom{n+k}{k}x^k\mathscr{L}_n^k(x^2)y^k\mathscr{L}_n^k(y^2)\,e^{ik\psi}\\&\qquad+\sum_{l=1}^{n}\frac{1}{l!}\binom{n}{l}x^l\mathscr{L}_{n-l}^l(x^2)y^l\mathscr{L}_{n-l}^l(y^2)\,e^{-il\psi}.\end{aligned}\tag{3.6}$$

This formula was stated without proof by Bateman [2]. Later three different proofs were given by Buchholz [3, p. 144], Carlitz [4] and Miller [11, (3.14)], [12, (4.127)]. Miller's proof uses group theoretic methods.

4. Remarks.

Remark 4.1. For $x = y$, $r = 1$, $\psi = 0$ formula (3.3) implies the identity

$$1 = \sum_{k=0}^{\infty} \sum_{l=0}^{n} \frac{\alpha}{\alpha+k+l} \binom{n}{l} \frac{(\alpha+n+1)_k}{k!(\alpha+l)_k(\alpha+k)_l} (x^{k+l} \mathscr{L}_{n-l}^{\alpha+k+l}(x^2))^2.$$

Inequality (2.2) is contained in this identity. Expressions for $\mathscr{L}_n^{\alpha}((x+y)^2)$ and $\mathscr{L}_n^{\alpha}((x-y)^2)$ follow from (3.3) by putting $r = 1$ and $\psi = 0$ or π.

Remark 4.2. Askey [1, pp. 82, 83] applied the product formula (3.5) to define a convolution structure for Laguerre series, thus extending earlier results of McCully [10] for the case $\alpha = 0$. However, it was pointed out in Askey [1] that this convolution structure is not positive and that it is not defined for all L^1-functions. More satisfactory results might be obtained in terms of the functions $\Lambda_{n,\mu}^{\alpha}(x, t)$ which are defined by

$$\Lambda_{n,\mu}^{\alpha}(x, t) = \frac{L_n^{\alpha}(|\mu| x)}{L_n^{\alpha}(0)} e^{-(1/2)|\mu| x + i\mu t}, \tag{4.1}$$

$$x \geqq 0, \quad t \in \mathbb{R}, \quad n = 0, 1, 2, \cdots, \mu \in \mathbb{R}.$$

For suitable functions f on $[0, \infty) \times \mathbb{R}$ we can define a Fourier transform

$$\hat{f}(n, \mu) = \frac{1}{2\pi\Gamma(\alpha+1)} \int_{x=0}^{\infty} \int_{t=-\infty}^{\infty} f(x, t) \Lambda_{n,-\mu}^{\alpha}(x, t) x^{\alpha} \, dx \, dt \tag{4.2}$$

with (formal) inversion formula

$$f(x, t) = \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} \hat{f}(n, \mu) \Lambda_{n,\mu}^{\alpha}(x, t) |\mu|^{\alpha+1} \frac{(\alpha+1)_n}{n!} \, d\mu. \tag{4.3}$$

Then the product formula (3.4) takes the form

$$\begin{aligned} &\Lambda_{n,\mu}^{\alpha}(x^2, s) \Lambda_{n,\mu}^{\alpha}(y^2, t) \\ &\quad = \alpha\pi^{-1} \int_0^1 \int_0^{2\pi} \Lambda_{n,\mu}^{\alpha}(x^2 + y^2 + 2xyr \cos\psi, s + t + xyr \sin\psi) \\ &\qquad \cdot r(1-r^2)^{\alpha-1} \, dr \, d\psi, \qquad \alpha > 0, \end{aligned} \tag{4.4}$$

which implies a positive convolution structure associated with the above Fourier expansion.

Remark 4.3. For $\alpha = 0, 1, 2, \cdots$ the functions $\Lambda_{n,\mu}^{\alpha}(x^2, t)$ can be interpreted as zonal spherical functions on a certain homogeneous space. Let $\mathbb{R}^{2q-1}$ (with elements $(z, t) \in \mathbb{C}^{q-1} \times \mathbb{R}$) have the structure of a nilpotent Lie group N by the multiplication rule

$$(z_1, t_1) \cdot (z_2, t_2) = (z_1 + z_2, t_1 + t_2 + \operatorname{Im} \langle z_2, z_1 \rangle),$$

where $\langle \cdot, \cdot \rangle$ denotes the Hermitian inner product on $\mathbb{C}^{q-1}$. Let G be the semidirect product of $U(q-1)$ with N, where $u \in U(q-1)$ acts on N as the

automorphism $(z, t) \to (uz, t)$. This group G (the so-called group of bordered unitary matrices) and some of its representations are discussed by Vilenkin [15] and, in the case $q = 2$, for instance by Miller [11], [12]. The space $\mathbb{R}^{2q-1}$ can be considered as the homogeneous space $G/U(q-1)$. Zonal functions of $(z, t) \in \mathbb{R}^{2q-1}$ only depend on $|z|^2$ and t. It turns out that convolution for zonal functions on this space is commutative. The functional equation

$$f(x)f(y) = \int_k f(xky)\, dk, \qquad x, y \in G$$

for spherical functions f on the group G with respect to the compact subgroup K can be reduced in the above case to a formula with the structure of (4.4), $\alpha = q-2$. Thus the functions $(z, t) \to \Lambda_{n,\mu}^{q-2}(|z|^2, t)$ are identified as spherical functions on $\mathbb{R}^{2q-1}$ as a homogeneous space of G. It would be of interest to extend the results in Vilenkin [15] in such a way that the addition formula (3.3), $\alpha = q-2$, is obtained from the group theoretic interpretation, thus extending results by Miller [11], [12] in the case $q = 2$.

Remark 4.4. It was pointed out in Koornwinder [9, § 5, Remark 8] that an addition formula for Laguerre polynomials cannot be obtained as a limit case of the addition formula for Jacobi polynomials. However, in a very recent paper Durand [5] succeeded in deriving an addition formula for Laguerre polynomials in this way. His result is a finite expansion of

$$L_n^{\alpha}(x^2 + y^2 - 2xyrt - xyr^2)$$

in terms of certain polynomials in r and t which are products of Hermite and Bessel polynomials.

Acknowledgment. The author would like to thank the referee for calling his attention to [11], [12] and [15].

REFERENCES

[1] R. ASKEY, *Orthogonal polynomials and positivity*, Special Functions and Wave Propagation, Studies in Applied Mathematics 6, Society for Industrial and Applied Mathematics, Philadelphia, 1970, pp. 64–85.

[2] H. BATEMAN, *Partial Differential Equations of Mathematical Physics*, Cambridge University Press, Cambridge, England, 1932.

[3] H. BUCHHOLZ, *Die Konfluente Hypergeometrische Funktion*, Springer-Verlag, Berlin, 1953.

[4] L. CARLITZ, *A formula of Bateman*, Proc. Glasgow Math. Assoc., 3 (1957), pp. 99–101.

[5] L. DURAND, *A symmetrical addition formula for the Laguerre polynomials*, Rep. LA-UR-76-465, Los Alamos Scientific Laboratory, Los Alamos, NM, 1976.

[6] A. ERDÉLYI, W. MAGNUS, F. OBERHETTINGER AND F. G. TRICOMI, *Higher Transcendental Functions*, vol. II, McGraw-Hill, New York, 1953.

[7] T. H. KOORNWINDER, *The addition formula for Jacobi polynomials, II. The Laplace type integral representation and the product formula*, Rep. TW 133, Math. Centrum Afd. Toegepaste Wisk., 1972.

[8] ———, *The addition formula for Jacobi polynomials, III. Completion of the proof*, Rep. TW 135, Math. Centrum Afd. Tolgepaste Wisk., 1972.

[9] ———, *Jacobi polynomials, III. An analytic proof of the addition formula*, this Journal, 6 (1975), pp. 533–543.

[10] J. McCULLY, *The Laguerre transform*, SIAM Rev., 2 (1960), pp. 185–191.
[11] W. MILLER, JR., *On the special function theory of occupation number space*, Comm. Pure Appl. Math., 18 (1965), pp. 679–696.
[12] ———, *Lie Theory and Special Functions*, Academic Press, New York, 1968.
[13] R. L. ŠAPIRO, *Special functions related to representations of the group $SU(n)$, of class I with respect to $SU(n-1)$ ($n \geqq 3$)*, Izv. Vysš. Učebn. Zaved. Matematika, 71 (1968), pp. 97–107. (In Russian.)
[14] G. SZEGO, *Orthogonal Polynomials*, Colloquium Publications, vol. 23, American Mathematical Society, Providence, RI, 1972.
[15] N. JA. VILENKIN, *Laguerre polynomials, Whittaker functions and the representations of groups of bordered matrices*, Math. USSR-Sb., 4 (1968), pp. 399–410.
[16] G. N. WATSON, *Another note in Laguerre polynomials*, J. London Math. Soc., 14 (1939), pp. 19–22.
[17] F. ZERNIKE AND H. C. BRINKMAN, *Hypersphärische Funktionen und die in sphärischen Bereichen orthogonalen Polynome*, Proc. Royal Acad. Amsterdam, 38 (1935), pp. 161–170.

SIAM J. MATH. ANAL.
Vol. 8, No. 3, May 1977

A SYMMETRICAL ADDITION FORMULA FOR THE LAGUERRE POLYNOMIALS*

LOYAL DURAND†

Abstract. We present a symmetrical addition formula for the general Laguerre polynomials $L_n^\alpha(Z)$ with argument $Z = x+y+2\sqrt{xy}\,rt - xyr^2$. The formula involves a finite sum of Laguerre and Hermite polynomials, and can be integrated to give a new product formula for $L_n^\alpha(x)L_n^\alpha(y)$. This addition formula is obtained as a limiting case of Koornwinder's addition formula for the Jacobi polynomials.

1. Introduction. Koornwinder [1] recently derived a symmetric addition formula for the Laguerre polynomials $L_n^\alpha(x)$ as a limit of the addition formula for the disk polynomials [2], [3]. Koorwinder's result can be written as

$$\begin{aligned}&\exp(-\sqrt{xy}\,re^{-i\psi})L_n^\alpha(x+y+2\sqrt{xy}\,r\cos\psi)\\ &\quad=\sum_{k=0}^{\infty}\sum_{l=0}^{n}(\alpha+k+l)\frac{\Gamma(\alpha+k)\Gamma(n-l+1)}{\Gamma(k+1)\Gamma(n+\alpha+k+1)}(-1)^{k-l}\\ &\quad\cdot(xy)^{(1/2)(k+l)}L_{n-l}^{\alpha+k+l}(x)L_{n-l}^{\alpha+k+l}(y)P_l^{(\alpha-1,k-l)}(2r^2-1)r^{k-l}e^{-i(k-l)\psi}.\end{aligned}\tag{1}$$

This expression reduces for $r=1$ and $\alpha\to 0+$ to Bateman's addition formula for $L_n^0(x)$ [4, p. 457], [5] and may be integrated [1] to give Watson's product formula for $L_n^\alpha(x)L_n^\alpha(y)$ [6],

$$\begin{aligned}L_n^\alpha(x)L_n^\alpha(y)&=2^{\alpha-1/2}\frac{\Gamma(n+\alpha+1)}{\sqrt{\pi}\,\Gamma(n+1)}\int_0^\pi e^{-\sqrt{xy}\cos\psi}\frac{J_{\alpha-(1/2)}(\sqrt{xy}\sin\psi)}{(\sqrt{xy}\sin\psi)^{\alpha-1/2}}\\ &\quad\cdot L_n^\alpha(x+y+2\sqrt{xy}\cos\psi)(\sin\psi)^{2\alpha}\,d\psi.\end{aligned}\tag{2}$$

A different limit of the addition formula for the disk polynomials leads to Koornwinder's addition formula for the Jacobi polynomials [7]–[9],

$$\begin{aligned}&P_n^{(\alpha,\beta)}(xy+\sqrt{1-x^2}\sqrt{1-y^2}\,rt+\tfrac{1}{2}(1-x)(1-y)(r^2-1))\\ &\quad=\sum_{k=0}^{n}\sum_{l=0}^{k}c_{k,l}(n,\alpha,\beta)[(1-x)(1-y)]^{(1/2)(k+l)}[(1+x)(1+y)]^{(1/2)(k-l)}\\ &\quad\cdot P_{n-k}^{(\alpha+k+l,\beta+k-l)}(x)P_{n-k}^{(\alpha+k+l,\beta+k-l)}(y)\;\; r^{k-l}P_l^{(\alpha-\beta-1,\beta+k-l)}(2r^2-1)C_{k-l}^{\beta}(t).\end{aligned}\tag{3}$$

* Received by the editors March 11, 1976.

† Department of Physics, University of Wisconsin, Madison, Wisconsin 53706. This work was done while the author was at the Theoretical Division, Los Alamos Scientific Laboratory, University of California, Los Alamos, New Mexico. 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.

Here $P_n^{(\alpha,\beta)}(z)$ and $C_n^{\beta}(z)$ are the usual Jacobi and Gegenbauer polynomials [10], [11, especially Chap. 10]. The coefficients $c_{k,l}(n, \alpha, \beta)$ are given by

$$c_{k,l}(n,\alpha,\beta)=2^{-2k}(\alpha+k+l)(\beta+k-l) \cdot \frac{\Gamma(\beta)\Gamma(n+\beta+1)\Gamma(n-k+1)\Gamma(\alpha+k)\Gamma(n+\alpha+\beta+k+1)}{\Gamma(\beta+k+1)\Gamma(n+\beta-l+1)\Gamma(n+\alpha+l+1)\Gamma(n+\alpha+\beta+1)}. \tag{4}$$

2. Procedure and result. It is the purpose of this note to point out that a symmetrical addition formula for $L_n^{\alpha}(Z)$ different from (1) can be derived as a limiting case of (3). Our procedure is suggested by the observation that the Laguerre polynomial $L_n^{\alpha}(x)$ can be obtained as a confluent limit of the Jacobi polynomial $P_n^{(\alpha,\beta)}$ [10, § 5.3],

$$L_n^{\alpha}(x)=\lim_{\beta\to\infty} P_n^{(\alpha,\beta)}\left(1-\frac{2x}{\beta}\right). \tag{5}$$

We note also that [10, § 5.6]

$$\lim_{\beta\to\infty} \beta^{-n/2}\Gamma(n+1)C_n^{\beta}(x/\sqrt{\beta})=H_n(x), \tag{6}$$

where $H_n(t)$ is the usual Hermite polynomial [10], [11, especially Chap. 10]. If we replace x, y, r, and t in (3) by $1-2x/\beta$, $1-2y/\beta$, $r\sqrt{\beta}$, and $t/\sqrt{\beta}$, take the limit $\beta\to\infty$, and use (5), (6), and the relation

$$\begin{aligned}
&\lim_{\beta\to\infty} \beta^{-l}P_l^{(\alpha-\beta-1,\beta+k-l)}(2\beta r^2-1)\\
&\quad=\lim_{\beta\to\infty}\frac{\Gamma(l+\alpha-\beta)}{\Gamma(\alpha-\beta)\Gamma(l+1)}\beta^{-l}{}_2F_1(-l,\alpha+k;\alpha-\beta;1-\beta r^2)\\
&\quad=(-1)^l\frac{1}{\Gamma(l+1)}{}_2F_0(-l,\alpha+k,r^2)\\
&\quad=r^{2l}\frac{\Gamma(\alpha+k+l)}{\Gamma(l+1)\Gamma(\alpha+k)}{}_1F_1\left(-l,-\alpha-k-l+1,-\frac{1}{r^2}\right)\\
&\quad=(-r^2)^l L_l^{-\alpha-k-l}\left(-\frac{1}{r^2}\right),
\end{aligned} \tag{7}$$

we find that[1]

$$\begin{aligned}
L_n^{\alpha}(Z)=&\sum_{k=0}^{n}\sum_{l=0}^{k}(-1)^l(\alpha+k+l)\frac{\Gamma(\alpha+k)\Gamma(n-k+1)}{\Gamma(n+\alpha+l+1)\Gamma(k-l+1)}(xy)^{(1/2)(k+l)}\\
&\cdot L_{n-k}^{\alpha+k+l}(x)L_{n-k}^{\alpha+k+l}(y)r^{k+l}L_l^{-\alpha-k-l}\left(-\frac{1}{r^2}\right)H_{k-l}(t),
\end{aligned} \tag{8}$$

$$Z=x+y-2\sqrt{xy}\,rt-xyr^2.$$

[1] It is tempting simply to replace x and y in (3) by $1-2x/\beta$ and $1-2y/\beta$, and then to take the limit $\beta\to\infty$ with r and t fixed. The result is a well-defined double series, but one which cannot be integrated to obtain a product for $L_n^{\alpha}(x)L_n^{\alpha}(y)$.

Note that both sides of (3) and (8) are polynomials in x, y, r, and t, so there is no problem with the limiting process.

The addition formula can be put in a somewhat more useful form by making the substitution $r=-e^{i\psi}$ and summing on $m=k-l$ instead of l. This gives our basic result,

$$\begin{aligned} L_n^{\alpha}(Z)=&\sum_{k=0}^{n}\sum_{m=0}^{k}(-1)^k(\alpha+2k-m)\\ &\cdot\frac{\Gamma(\alpha+k)\Gamma(n-k+1)}{\Gamma(n+\alpha+k-m+1)\Gamma(m+1)}(xy)^{k-(1/2)m}\\ &\cdot L_{n-k}^{\alpha+2k-m}(x)L_{n-k}^{\alpha+2k-m}(y)\,e^{i(2k-m)\psi}L_{k-m}^{-\alpha-2k+m}(-e^{-2i\psi})H_m(t), \end{aligned} \tag{9}$$

$$Z=x+y+2\sqrt{xy}\,te^{i\psi}-xy\,e^{2i\psi}.$$

This expression can be inverted to obtain a product formula for $L_n^{\alpha}(x)L_n^{\alpha}(y)$ by using the orthogonality relations for the Hermite polynomials [10, § 5.5] and the following relation derived in the Appendix:

$$\begin{aligned} \int_0^{2\pi} e^{2i(k'-k)\psi}&{}_1F_1(k-m+1,\alpha+2k-m+1,e^{-2i\psi})\,{}_1F_1(-k'+m,-\alpha-2k'\\ &+m+1,-e^{-2i\psi})\,d\psi \\ &=2\pi\delta_{k,k'} \qquad k,k'\geqq m. \end{aligned} \tag{10}$$

The result is as follows:

$$\begin{aligned} &(xy)^{k-(1/2)m}L_{n-k}^{\alpha+2k-m}(x)L_{n-k}^{\alpha+2k-m}(y)\\ &\quad=(-1)^m\pi^{-3/2}2^{-m-1}\frac{\Gamma(k-m+1)\Gamma(n+\alpha+k-m+1)}{\Gamma(n-k+1)\Gamma(\alpha+2k-m+1)}\\ &\quad\cdot\int_0^{2\pi}d\psi\int_{-\infty}^{\infty}dt\,L_n^{\alpha}(Z)\,e^{-i(2k-m)\psi}\,{}_1F_1(k-m+1,\alpha+2k-m\\ &\qquad+1,e^{-2i\psi})\,e^{-t^2}H_m(t), \qquad 0\leqq k\leqq n, \quad 0\leqq m\leqq k. \end{aligned} \tag{11}$$

In particular, for $k=m=0$, we obtain a result quite different from (2),

$$\begin{aligned} L_n^{\alpha}(x)L_n^{\alpha}(y)=&\tfrac{1}{2}\pi^{-3/2}\frac{\Gamma(n+\alpha+1)}{\Gamma(n+1)\Gamma(\alpha+1)}\\ &\cdot\int_0^{2\pi}d\psi\int_{-\infty}^{\infty}dt\,L_n^{\alpha}(Z)\,{}_1F_1(1,\alpha+1,e^{-2i\psi})\,e^{-t^2}. \end{aligned}$$

We conclude by noting that Koornwinder's addition formula (1) gives a symmetrical double series expansion of

$$\exp(-\sqrt{xy}\,re^{-i\psi})L_n^{\alpha}(x+y+2\sqrt{xy}\,r\cos\psi).$$

Because of the exponential, one of the series does not terminate. In contrast, (9) gives a finite symmetrical expansion of the polynomial $L_n^{\alpha}(Z)$, and appears, therefore, to be the natural addition formula for the Laguerre polynomials. Other

addition formulas are known [12, Chap. 4], [13, Chap. 8, § 5], [14, § 5, Remark 8], but these are generally unsymmetrical, and with the exception of that given in [14, § 5, Remark 8], involve infinite series.

Appendix. The orthogonality relation for the confluent hypergeometric functions given in (10) can be derived as follows. We begin with the orthogonality relation for the Jacobi polynomials [10, 4.3.3],

$$\int_{-1}^{1} P_n^{(\alpha,\beta)}(t)P_{n'}^{(\alpha,\beta)}(t)(1-t)^{\alpha}(1+t)^{\beta}\,dt = h_n^{\alpha,\beta}\delta_{n',n}, \tag{A.1}$$

$$h_n^{\alpha,\beta} = \frac{2^{\alpha+\beta+1}\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{(2n+\alpha+\beta+1)\Gamma(n+1)\Gamma(n+\alpha+\beta+1)}. \tag{A.2}$$

This relation can be written as a contour integral by introducing the Jacobi functions of the second kind, $Q_n^{(\alpha,\beta)}(t)$, defined for general complex argument by [10, § 4.61]

$$\begin{aligned} Q_n^{(\alpha,\beta)}(t) &= 2^{n+\alpha+\beta}\frac{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{\Gamma(2n+\alpha+\beta+2)}(t-1)^{-n-\alpha-1}(t+1)^{-\beta} \\ &\quad \cdot {}_2F_1\left(n+1, n+\alpha+1; 2n+\alpha+\beta+2; \frac{2}{1-t}\right), \qquad |\arg(t\pm 1)| < \pi. \end{aligned} \tag{A.3}$$

For n an integer, the function $(t-1)^{\alpha}(t+1)^{\beta}Q_n^{(\alpha,\beta)}(t)$ is analytic in the complex t-plane cut from -1 to $+1$, and has a discontinuity across the cut which is just $-i\pi(1-t)^{\alpha}(1+t)^{\beta}P_n^{(\alpha,\beta)}(t)$, t real, $-1 \leqq t \leqq 1$. The Jacobi polynomial $P_{n'}^{(\alpha,\beta)}(t)$ is of course continuous across the cut. We can use the relation of $P_n^{(\alpha,\beta)}$ to $Q_n^{(\alpha,\beta)}$ to rewrite (A.1) as

$$\begin{aligned} &\int_{-1}^{1} P_n^{(\alpha,\beta)}(t)P_{n'}^{(\alpha,\beta)}(t)(1-t)^{\alpha}\,(1+t)^{\beta}\,dt \\ &\quad = -\frac{i}{\pi}\int_C Q_n^{(\alpha,\beta)}(t)P_{n'}^{(\alpha,\beta)}(t)(t-1)^{\alpha}(t+1)^{\beta}\,dt = h_n^{\alpha,\beta}\delta_{n',n}, \end{aligned} \tag{A.4}$$

where the contour C circles the interval $[-1, 1]$ in the positive sense. If we extend the contour to circle $[-1, 1]$ twice, and make the change of variable $t = 2r^2 - 1$, we obtain the relation

(A.5)

$$-\frac{i}{2\pi}\int_{C_r} Q_n^{(\alpha,\beta)}(2r^2-1)P_{n'}^{(\alpha,\beta)}(2r^2-1)(r^2-1)^{\alpha}r^{2\beta+1}\,dr = 2^{-\alpha-\beta-2}\,h_n^{\alpha,\beta}\delta_{n',n},$$

where the contour C_r circles the interval $[-1, 1]$ in the r-plane once in the positive sense. It will be convenient to rewrite this equation for the parameters of interest,

$\alpha \to \alpha-\beta-1$, $\beta \to \beta+m$, $n \to k-m$, $n' \to k'-m$ as

$$
\begin{aligned}
&[i\Gamma(\alpha-\beta+k-m)\Gamma(\beta+k+1)]^{-1}\int_{C_r} Q_{k-m}^{(\alpha-\beta-1,\beta+m)}(2r^2-1) \\
&\qquad \cdot P_{k'-m}^{(\alpha-\beta-1,\beta+m)}(2r^2-1)\quad (r^2-1)^{\alpha-\beta-1} r^{2\beta+2m+1}\,dr \\
&\qquad = \pi[(\alpha+2k-m)\Gamma(k-m+1)\Gamma(\alpha+k)]^{-1}\delta_{kk'}, \qquad k, k' \geqq m.
\end{aligned} \tag{A.6}
$$

Equation (A.6) is in a form suitable for application of the limiting procedure which led to (7) and (8). If we replace r by $r\sqrt{\beta}$ in (A.6) and take the limit $\beta \to \infty$, making use of (7) and the relation

$$
\begin{aligned}
&\lim_{\beta\to\infty} \beta^{\alpha+k}[\Gamma(\alpha-\beta+k-m)\Gamma(\beta+k+1)]^{-1} Q_{k-m}^{(\alpha-\beta-1,\beta+m)}(2\beta r^2-1) \\
&\qquad = \tfrac{1}{2}[\Gamma(\alpha+2k-m+1)]^{-1} r^{-2\alpha-2k}\, e^{-1/r^2}\, {}_1F_1\Big(k-m+1, \alpha+2k-m \\
&\qquad\qquad +1, \frac{1}{r^2}\Big),
\end{aligned} \tag{A.7}
$$

we find that

$$
\begin{aligned}
&-i\int_{0+} {}_1F_1\Big(k-m+1, \alpha+2k-m+1, \frac{1}{r^2}\Big)\, {}_1F_1\Big(-k'+m, -\alpha-2k'+m+1, \\
&\qquad\qquad -\frac{1}{r^2}\Big) r^{2k'-2k-1}\,dr \\
&\qquad = 2\pi\delta_{kk'}, \qquad k, k' \geqq m.
\end{aligned} \tag{A.8}
$$

Alternatively, with $r = e^{i\psi}$,

$$
\begin{aligned}
&\int_0^{2\pi} e^{2i(k'-k)\psi}\, {}_1F_1(k-m+1, \alpha+2k-m+1, e^{-2i\psi})\, {}_1F_1(-k'+m, -\alpha-2k' \\
&\qquad\qquad +m+1, -e^{-2i\psi})\,d\psi \\
&\qquad = 2\pi\delta_{kk'}, \qquad k, k' \geqq m.
\end{aligned} \tag{A.9}
$$

This is the relation needed to invert (9) and obtain the product formula (11) for $L_n^{\alpha}(x)L_n^{\alpha}(y)$. It is apparently new.

Acknowledgment. The author would like to thank the faculty of the School of Natural Sciences of the Institute for Advanced Study, Princeton, New Jersey, for the hospitality accorded him during the fall of 1975, when this work originated.

REFERENCES

[1] T. H. KOORNWINDER, *The addition formula for Laguerre polynomials*, this Journal, to appear.

[2] R. L. ŠAPIRO, *Special functions related to representations of the group SU(n), of class I with respect to SU(n−1)*, Izv. Vyss. Učebn. Zaved Matematika, 71 (1968), no. 4, pp. 97–107. (In Russian.)

[3] T. H. KOORNWINDER, *The addition formula for Jacobi polynomials, III. Completion of the proof*, Rep. TW 135, Mathematisch Centrum, Amsterdam, 1972.

[4] H. BATEMAN, *The Partial Differential Equations of Mathematical Physics*, Cambridge University Press, Cambridge, 1932.

[5] L. CARLITZ, *A formula of Bateman*, Proc. Glasgow Math. Assoc., 3 (1957), pp. 99–101.

[6] G. N. WATSON, *Another note on Laguerre polynomials*, J. London Math. Soc., 14 (1939), pp. 91–22.

[7] T. H. KOORNWINDER, *Yet another proof of the addition formula for Jacobi polynomials, Three nodes on classical orthogonal polynomials*, part I, Rep. TW 150, Mathematisch Centrum, Amsterdam, 1975.

[8] ———, *The addition formula for Jacobi polynomials, I. Summary of results*, Indag. Math., 34 (1972), pp. 188–191.

[9] ———, *The addition formula for Jacobi polynomials and spherical harmonics*, SIAM J. Appl. Math., 25 (1973), pp. 236–246.

[10] G. SZEGÖ, *Orthogonal Polynomials*, Colloquium Publications, vol. 23, American Mathematical Society, Providence, RI, 1972.

[11] A. ERDÉLYI, W. MAGNUS, F. OBERHETTINGER AND F. G. TRICOMI, *Higher Transcendental Functions*, vols. I, II, McGraw-Hill, New York, 1953.

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

[13] N. JA. VILENKIN, *Special Functions and the Theory of Group Representations*, American Mathematical Society, Providence, RI, 1968.

[14] T. H. KOORNWINDER, *Jacobi polynomials, III. An analytic proof of the addition formula*, this Journal, 6 (1975), pp. 533–543.

SIAM J. MATH. ANAL.
Vol. 8, No. 3, May 1977

L_2-STABILITY AND -INSTABILITY OF LARGE-SCALE SYSTEMS DESCRIBED BY INTEGRODIFFERENTIAL EQUATIONS*

R. K. MILLER† AND A. N. MICHEL‡

Abstract. New results for (i) L_2-stability, (ii) L_2-instability, and (iii) asymptotic stability, in the sense of Lyapunov for a large class of large-scale dynamical systems (also called composite systems, interconnected systems, and multiloop systems) described by Volterra integrodifferential equations are presented. These results allow usage of frequency domain techniques. A simple example is given which illustrates how the various theorems can be applied.

1. Introduction. New results for L_2-stability and L_2-instability for a class of large-scale systems (also called interconnected systems, composite systems, multiloop systems, and the like) described by Volterra integrodifferential equations are established. It is shown that when the large-scale systems considered are bounded input bounded output stable on L_2 (BIBO stable on L_2), then they are also asymptotically stable in the sense of Lyapunov. The present results are proved for several forms of the describing equations, including the case of multiple-input multiple-output systems (MIMO systems) with interconnections. For the case of MIMO systems, conditions are established which guarantee instability when at least one subsystem is unstable.

In the present approach, the objective is to analyze large-scale systems in terms of lower-order subsystems and in terms of the system interconnecting structure. The results are applied to an example to illustrate the use of the theory. We have chosen a well known integrodifferential model which has enough complexity to illustrate the various possibilities.

The present BIBO stability results are motivated by those of Sandberg [20], [21] and Zames [24], [25] for single loop systems while the Lyapunov stability result is motivated by a result of Grossman and Miller [5]. The instability result constitutes a generalization of work by Vidyasagar [22]. Recent related results for BIBO stability of large-scale systems are due to Porter and Michel [17], [18] and Lasley and Michel [8]–[10]. For a general introduction to Volterra integral equations, refer to the monograph by Miller [11]. For additional specific information on resolvents which will be needed, refer to Grossman and Miller [5], [6] and Miller [12]. For a comprehensive summary of BIBO stability results, refer to [2] and [23].

2. Notation and preliminaries. Let $C=[c_{ij}]$ denote an $m\times n$ matrix and let C' be the transpose of C. If C and D are real $m\times n$ matrices, then $C\geqq D$ means $c_{ij}\geqq d_{ij}$ for all i and j. Let I denote the identity matrix, let $R^+=[0,\infty)$ and let R^m be the Euclidean m-space. If $x=[x_1,\cdots,x_m]\in R_m$, then $|x|$ is the norm of x. Let L_p (or L_p^m if m needs to be emphasized) denote the set of all Lebesgue measurable

* Received by the editors October 21, 1975, and in revised form February 16, 1976.

† Department of Mathematics, Iowa State University, Ames, Iowa 50011. The work of this author was supported by the National Science Foundation under Grant MPS 75-07622.

‡ Electrical Engineering Department and Engineering Research Institute, Iowa State University, Ames, Iowa 50011. The work of this author was supported in part by the National Science Foundation under Grant ENG-14093 and in part by the Engineering Research Institute, Iowa State University.

functions $f: R^+ \to R^m$ such that $\|f\|_p = [\int_0^\infty |f(t)|^p\, dt]^{1/p}$ is finite. When $p = 2$, L_2 is a Hilbert space with inner product $\langle f, g\rangle = \int_0^\infty f(t)'g(t)\, dt$ and $\|f\|_2 = \langle f, f\rangle^{1/2}$. Given a subspace $X \subset L_2$, $X^\perp = \{f: \langle f, g\rangle = 0 \text{ for all } g \in X\}$. Also, let L_∞ (or L_∞^m if m needs to be emphasized) be the space of essentially bounded functions f: $R^+ \to R^m$ and let $\|f\|_\infty = \text{ess sup}_{t \geqq 0} |f(t)|$. Let $C[0, \infty]$ be the set of all continuous functions f: $R^+ \to R^m$ and $C_0 = \{f \in C[0, \infty) | f(t) \to 0 \text{ as } t \to \infty\}$, the subset of $C[0, \infty)$ which goes to zero as $t \to \infty$.

Given a function $f \in L_p$, let f_T denote the *truncation* of f at time T, that is $f_T(t) = f(t)$ on $0 \leqq t \leqq T$ and $f_T(t) = 0$ if $t > T$. Also, let L_{pe} denote the *extended space* of L_p (see, e.g., Zames [24], [25]):

$$L_{pe} = \{f: f_T \in L_p \text{ for all } T > 0\}.$$

Thus L_{pe} is the space of all locally L_p functions $f(t)$ defined on $0 \leqq t < \infty$.

If H: $L_{2e}^k \to L_{2e}^l$, H is said to be L_2-*stable* if H maps L_2^k into L_2^l. In this case the *gain* of H, written $g(H)$ is the smallest number K such that $\|(HX)_T\|_2 \leqq K\|x_T\|_2$ for all $x \in L_{2e}^k$ and all $T > 0$. If the stable set of H, $S(H) = \{x \in L_2^k: Hx \in L_2^l\}$ is a proper subset of L_2^l, then H is said to be L_2-*unstable*. In this case the *conditional gain* of H, written $g_c(H)$, is the smallest number K such that $\|(Hx)_T\|_2 \leqq K\|x_T\|_2$ for all $T > 0$ and all x in the stable set $S(H)$. In the special case where $S(H) = L_2$, H is stable and $g_c(H) = g(H)$. H is *interior conic* (C, r) if $\|(Hx)_T - Cx_T\|_2 \leqq r\|x_T\|$ for some real constant $r \geqq 0$ and some matrix C.

Let $\mathscr{L}_e$ denote the class of linear time-invariant operators on L_{pe} having the following properties: If $H \in \mathscr{L}_e$, then there is a function $h \in L_{1e}$ and two sequences $\{h_i\}$ and $\{t_i\}$, such that $t_i < t_{i+1}$ with $t_1 = 0$, $t_j \to \infty$ as $j \to \infty$ and

$$(Hx)(t) = \sum_{i=1}^{\infty} h_i x(t - t_i) + \int_0^t h(s)x(t - s)\, ds \tag{1}$$

for all $x \in L_{pe}$. The class $\mathscr{L}$ will consist of all $H \in \mathscr{L}_e$ such that the corresponding function $h(t)$ and sequence $\{h_i\}$ satisfy the conditions

$$\sum_{i=1}^{\infty} |h_i| < \infty, \qquad \int_0^\infty |h(t)|\, dt < \infty.$$

Let $H^*(s)$ and $h^*(s)$ denote the *Laplace transforms* of the operator H and of $h(t)$, respectively (see (1)), where $s = \sigma + j\tau$, $j = \sqrt{-1}$ and $\sigma, \tau \in R^1$. If $H \in \mathscr{L}$, then the representation $H^*(s)$ is guaranteed to converge for all s with $\text{Re } s = \sigma \geqq 0$. The function $H^*(j\omega)$ is essentially the *Fourier transform* or frequency response of H. The *resolvent* of H, denoted by R, is the operator $R \in \mathscr{L}_e$ given by

$$R^*(s) = (sI - H^*(s))^{-1}.$$

Given $H \in \mathscr{L}$, H is said to have *property* **F** when $\det(j\omega I - H^*(j\omega)) \neq 0$ for all $\omega \in R^1$. Moreover H will have *property* **L** if $\det(sI - H^*(s)) \neq 0$ for all s such that $\text{Re } s \geqq 0$. It is known (see, e.g., Grossman and Miller [6] and Jordan and Wheeler [7]) that if $h_i = 0$ for all $i > 1$, H has property L if and only if $R \in \mathscr{L}$. Also, H

has property F if and only if there is a finite set of points $\{s_1, s_2, \cdots, s_N\}$ in the half plane $\operatorname{Re} s > 0$ such that $\det(s_j I - H^*(s_j)) = 0$. For such an H one can find a set of matrices M_{jk}, nonnegative integers N_j and an operator $S \in \mathscr{L}$ such that

$$R^*(s) = S^*(s) + \sum_{j=1}^{N} \sum_{k=0}^{N_j} M_{jk}/(s-s_j)^k, \qquad (\operatorname{Re} s_j > 0). \tag{2}$$

The operator S is called the *residual resolvent* of H. (If R has property $\mathscr{L}$, then there are no such points s_j and so the resolvent and residual are identical.) The resolvent R and residual resolvent S are always of the special form

$$Rx(t) = \int_0^t r(t-u)x(u)\,du, \qquad Sx(t) = \int_0^t s(t-u)x(u)\,du,$$

respectively. The matrix functions $r(t) = [r_{ij}(t)]$ and $s(t) = [s_{ij}(t)]$ are called the *kernels* of the operators R and S, respectively. When $H \in \mathscr{L}$ and has property L (or F) then $r_{ij} \in C_0 \cap L_2$ (or $s_{ij} \in C_0 \cap L_2$) for all indices i and j.

It is well known (see [2] or [23]) that the gain H of $\mathscr{L}$ on the space L_2 is

$$g(H) = \operatorname*{ess\,sup}_{-\infty<\omega<\infty} |H^*(j\omega)|.$$

If $H \in \mathscr{L}$ and has property L then by the same reasoning the gain of the resolvent is

$$g(R) = \operatorname*{ess\,sup}_{-\infty<\omega<\infty} |(j\omega - H(j\omega))^{-1}|.$$

A *regular operator* H is an element of $\mathscr{L}_e$ such that for the corresponding sequence $\{h_i\}$, $h_i = 0$ for all $i > 1$. If H is regular, the resolvent kernel $r(t) = [r_{ij}(t)]$ has continuously differentiable elements $r_{ij}(t)$ on R^+ (see Grossman and Miller [6]). Also $\dot{r}_{ij} \in L_1 \cap L_2 \cap C_0$ when H and $R \in \mathscr{L}$.

A linear integrodifferential equation with kernel $H \in \mathscr{L}_e$ and forcing function $f \in L_{2e}$ is the equation

$$\dot{y}(t) = Hy(t) + f(t), \qquad t \geqq 0,$$

with initial condition $y(0) = y_0$. By a solution $y(t)$ we mean a function which is absolutely continuous on any finite interval $0 \leqq t \leqq T$ with derivative $\dot{y}(t)$ satisfying the equation at almost all points $t \in R^+$. If $H \in \mathscr{L}_e$ and $f \in \mathscr{L}_{2e}$ it is well known that this linear equation has a unique solution. Indeed, if $r(t)$ is the kernel of the resolvent R, then

$$y(t) = r(t)y_0 + \int_0^t r(t-s)f(s)\,ds, \tag{3}$$

or $y = R[\delta y_0 + f]$, where δ is the delta function.

All nonlinear integrodifferential equations considered herein are of the form

$$\dot{x}(t) = Hx(t) + B(t, x_t) + f(t), \qquad t \geqq 0, \tag{N}$$

with $x(0) = x_0$. Here $H \in \mathscr{L}_e$, $f \in L_{2e}$, x_t denotes the truncation of x at time t, and B is a continuous functional from $R^+ \times C[0, \infty)$ into R^n. These assumptions insure that (N) has a solution (which will be unique if the functional B satisfies the usual

Lipschitz condition). If one thinks of $B(t, x_t)+f(t)$ as known, then from (3) the following *variation of constants formula* follows:

$$x = R[\delta x_0 + f] + R[B(t, x_t)]. \tag{V}$$

Remark 1. Let $r(t)$ be the resolvent kernel of the regular resolvent R. If B has the *regular form*

$$B(t, x) = B_1(t, x(t)) + \int_0^t h(t-s)B_2(s, x(s))\, ds, \tag{4}$$

then (V) can be put into an alternate form. Since

$$\dot{r}(t) = r(t)h_1 + \int_0^t r(t-s)h(s)\, ds = r(t)h_1 + r * h(t),$$

where $*$ denotes convolution, then $r * h = \dot{r} - rh_1$. The last term in (V) is of the form

$$\begin{aligned} R[B_1 + h * B_2] &= R(B_1) + r * h * B_2 = R(B_1) + (\dot{r} - rh_1) * B_2 \\ &= R(B_1 - h_1 B_2) + \dot{r} * B_2. \end{aligned}$$

Thus, (V) can be written in the alternate form

$$x = R[\delta x_0 + f] + R[B_1(t, x(t)) - h_1 B_2(t, x(t))] + \dot{r} * B_2(t, x(t)). \tag{V'}$$

Equation (N) is called L_2-*stable* (i.e., BIBO stable on L_2) if for each $f \in L_2$ and initial conditions $x_0 \in R^n$, all solutions of (N) are in L_2. Otherwise (N) is called L_2-*unstable*. System (N) is called *stable in the sense of Lyapunov* if for any $\varepsilon > 0$ there is a $\delta > 0$ such that when $|x_0| \leqq \delta$ and $\|f\|_2 \leqq \delta$ then $|x(t)| < \varepsilon$ for all $t \geqq 0$. It is *asymptotically stable in the sense of Lyapunov* if in addition there is a $\delta_0 > 0$ such that $x(t) \in C_0$ when $|x_0| \leqq \delta_0$ and $\|f\|_2 \leqq \delta_0$.

3. Main results. Systems are considered which can be described by the set of nonlinear integrodifferential equations

$$\dot{x}_k(t) = \sum_{j=1}^{N} [H_{kj}x_j(t) + B_{kj}(t, x_{jt})] + f_k(t) \tag{5}$$

for $k = 1, 2, \cdots, N$ and $t \geqq 0$, with initial conditions $x_k(0) = x_{0k}$. Here $x_k : R^+ \to R^{n_k}$, $f_k \in L_2^{n_k}$, $H_{kj} : L_{2e}^{n_j} \to L_{2e}^{n_k}$ is in $\mathscr{L}_e$, $x_{jt} = (x_j)_t$ is the truncation of x_j and B_{ij} is a continuous nonlinear functional. Letting $x = [x_1, x_2, \cdots, x_n]'$, $f = [f_1, f_2, \cdots, f_N]'$, $B = [\sum B_{1j}, \cdots, \sum B_{Nj}]'$ and $H = [H_{jk}]$, system (5) assumes the form (N) with initial condition $x_0 = [x_{10}, \cdots, x_{N0}]' \in R^m$ where $m = \sum_{j=1}^{N} n_j$. The *matrix of gains* $g(B)$ is defined as $g(B) = [g(B_{ij})]$. The matrix of gains $g(C)$ of any other matrix C is defined similarly.

When (N) is of the form (5), one speaks of a *large-scale system* (N) *with decomposition* (5).

THEOREM 1. *If $g(B_{kj}) < \infty$ for all k and j, if $H = [H_{ij}] \in \mathscr{L}$, if H has property L with resolvent $R = [R_{ij}]$, and if the successive principal minors of the test matrix $I - g(R)G(B)$ are all positive, then* (5) *is L_2-stable.*

THEOREM 2. *Let B_{ij} be regular in the sense of* (4), *let $H = [H_{kj}] \in \mathscr{L}$ be regular, satisfy property L, have resolvent $R = [R_{kj}]$ with kernel $[r_{kj}]$, let T_{kj} be the operator*

defined by $tT_{kj}x = \dot{r}_{kj} * x$, *and let* $T = [T_{kj}]$. *If the successive principal minors of the test matrix* $I - g(R)g(B_1 - h_1B_2) - g(T)g(B_2)$ *are all positive, then* (5) *is* L_2*-stable.*

THEOREM 3. *If the hypotheses of Theorem* 1 *and H is regular or Theorem* 2 *are true, then* (5) *is asymptotically stable in the sense of Lyapunov.*

Remark 2. In some applications of Theorem 2, $B_1 = h_1B_2$. In this case the term $g(B_1 - h_1B_2)$ in Theorem 2 is zero.

Remark 3. Consider the system of equations

$$\dot{x}_k(t) = f_k(t) + \sum_{j=1}^{N} \left[G_{kj}(t, x_j(t)) + \int_0^t M_{kj}(t-s) N_{kj}(s, x_j(s))\, ds \right]$$

where G_{kj} is interior conic (m_{kj}, r_{kj}) and N_{kj} is interior conic (n_{kj}, w_{kj}). In this case one can rewrite the equation for $\dot{x}_k(t)$ as

$$\begin{aligned} \dot{x}_k(t) = f_k(t) &+ \sum_{j=1}^{N} m_{kj}x_j(t) + \int_0^t M_{kj}(t-s) n_{kj}x_j(s)\, ds \\ &+ \sum_{j=1}^{N} [G_{kj}(t, x_j(t)) - m_{kj}x_j] \\ &+ \int_0^t M_{kj}(t-s)[N_{kj}(s, x_j(s)) - n_{kj}x_j(s)]\, ds. \end{aligned}$$

This equation has the correct form for Theorem 2 since $B_1 = [G_{kj}(t, x_j(t)) - m_{kj}x_j]$, $g(B_1) \leqq [r_{kj}]$, $B_2 = [M_{kj} * (N_{kj}(t, x_j) - n_{kj}x_j)]$, $g(B_2) \leqq g([M_{kj}])[w_{kj}]$, $h_1 = [m_{kj}]$, and $h(t) = [M_{kj}(t) n_{kj}]$.

Of special interest in applications are systems described by (5) with $H_{ij} = 0$ for all $i \neq j$, i.e., MIMO systems of the form

$$\dot{x}_k(t) = H_k x_k(t) + \sum_{j=1}^{N} B_{kj}(t, x_{jt}) + f_k(t), \tag{6}$$

$k = 1, \cdots, N$, with $x_k(0) = x_{k0}$. This equation can be viewed as an interconnection of *N isolated subsystems* of the form

$$\dot{z}_k(t) = H_k z_k(t) + f_k(t), \tag{7}$$

$z_k(0) = z_{k0}$, $k = 1, \cdots, N$, with *interconnecting structure* specified by B_{kj}, $k, j = 1, \cdots, N$.

COROLLARY 1. *In* (6) *assume that*

(i) $H_k \in \mathscr{L}$ *with stable resolvent* R_k, $k = 1, \cdots, N$;

(ii) $g(B_{kj}) < \infty$, $k, j = 1, \cdots, N$;

(iii) $f_k \in L_2$, $k = 1, \cdots, N$; *and*

(iv) *the successive principal minors of the test matrix* $M = [m_{ij}]$ *are all positive, where*

$$m_{ij} = \begin{cases} 1 - g(R_i)g(B_{ij}), & i = j \\ -g(R_i)g(B_{ij}), & i \neq j. \end{cases}$$

Then system (6) *is* L_2*-stable and for H regular also asymptotically stable in the sense of Lyapunov.*

Under certain assumptions one can show that if at least one of the isolated systems (7) is L_2-unstable, then the interconnected system (6) is also L_2-unstable.

THEOREM 4. *Assume that for system* (6) *the following conditions hold*:

(i) $H_k \in \mathscr{L}$, $k = 1, \cdots, N$;

(ii) $g(B_{ij}) < \infty$, $i, j = 1, \cdots, N$;

(iii) $f_j \in L_2$, $j = 1, \cdots, N$;

(iv) *for each* k, H_k *satisfies either property F and/or property L and for at least one* k_0 *it satisfies property F only*; *and*

(v) *let* $M = [m_{ij}] = [g(B_{ij})g_c(R_j)]$, *and let* $\rho(M)$ *be the spectral radius of* M.

Assume that either $\rho(M) < 1$ *or* $\rho(M) = 1$ *and* $m_{ij} > 0$ *for all* $i, j = 1, \cdots, N$.

Then system (6) *is* L_2*-unstable.*

Theorem 4 is a corollary of Theorem 5 given below. This result yields sufficient conditions for L_2-instability of interconnected systems described by equations of the form

$$\begin{aligned} e_k &= f_k - \sum_{j=1}^{m} B_{kj}x_j, \\ x_k &= R_k(e_j + v_j), \end{aligned} \tag{8}$$

$k = 1, \cdots, N$, where f_k, $v_k \in L_2^{n_k}$, e_k, $x_k \in L_e^{n_k}$, $R_k \in \mathscr{L}_e^{n_k}$, and $B_{kj}: L_{2e}^{n_j} \to L_{2e}^{n_k}$. Define $X_k = \{x \in L_2^{n_k}: R_k x \in L_2^{n_k}\}$, the stable manifold of R_k.

THEOREM 5. *Suppose* $(X_k)^{\perp} \neq \{0\}$ *for at least one* $k = k_0$ *and* $g(B_{kj}) < \infty$ *for all* k *and* j. *Let* $M = [m_{ij}] = [g(B_{ij})g_c(R_j)]$ *and let* $\rho(M)$ *be the spectral radius of* M. *Assume that either* $\rho(M) < 1$ *or* $\rho(M) = 1$ *and* $m_{ij} > 0$ *for all* i *and* j. *If* $f_k + v_k \in (X_k)^{\perp}$ *for all* k *and* $f_k + v_k \not\equiv 0$ *when* $k = k_0$, *then system* (8) *is* L_2*-unstable, i.e.,* $x_j \notin L_2$ *for at least one value* j.

Remark 4. For BIBO stability results of system (8) refer to Porter and Michel [17], [18] and Lasley and Michel [9]. Theorem 5 may be viewed as the instability counterpart of the results of [9], [17], [18]. Theorem 5 is a generalization of a result reported in [22]. Specifically, in [22] it is required that *all* subsystems be L_2-unstable and all $v_k = 0$.

4. An example. The use of the theory presented above will now be illustrated by studying the stability of a coupled core nuclear reactor. This well known physical model is sufficiently complicated to illustrate both stability and instability results. We wish to give a proof of the physically reasonable assertion that with weak interactions between cores, stability (or instability) of the individual cores implies stability (or instability) of the entire system. Most importantly we can also give a computable method of deciding what we mean by "weak interactions".

For a general discussion of reactor models, refer to [1]. The point kinetics model for a reactor with N cores, as described in [16], is given by

$$\begin{aligned} \dot{p}_j(t) &= \frac{\rho_j - \varepsilon_j - \beta_j}{\Lambda_j} p_j(t) + \frac{\rho_j}{\Lambda_j} + \sum_{i=1}^{6} \frac{\beta_{ij}}{\Lambda_j} c_{ij}(t) \\ &\quad + \frac{1}{\Lambda_j} \sum_{k=1}^{N} \varepsilon_{jk} \frac{P_{k0}}{P_{j0}} \int_0^t h_{kj}(t-s)p_k(s)\, ds, \\ \dot{c}_{ij}(t) &= \lambda_{ij}[p_j(t) - c_{ij}(t)] \end{aligned} \tag{9}$$

for $j=1,\cdots.N$ and $i=1,\cdots,6$. Here $p_j(t)=[P_j(t)-P_{jo}]/P_{j0}$ and $c_{ij}(t)=(\lambda_{ij}\Lambda_j/P_{j0}\beta_{ij})(C_{ij}(t)-C_{ij0})$, where P_j denotes the power in the jth core, C_{ij} represents the effective concentration of the ith precursor in the jth core, β_{ij}, λ_{ij}, ε_j, ε_{kj}, and Λ_j are positive constants, $\beta_j=\sum_{i=1}^{6}\beta_{ij}$, and $h_{kj}(t)$ is the coupling function relating neutron migration from the kth to the jth core. P_{j0} and C_{ij0} are the equilibrium power and precursor concentrations in the jth core while ρ_j is the reactivity in the jth core. Assume that

$$\rho_j(t)=\int_0^t W_j(t-s)p_j(s)\,ds$$

is correct at least to linear terms, where the feedback function W_j is in L_1.

Solving for $c_{ij}(t)$ in terms of $p_j(t)$ one has

$$c_{ij}(t)=c_{ij}(0)\,e^{-\lambda_{ij}t}+\lambda_{ij}(e^{\lambda_{ij}t}*p_j).$$

Substituting $c_{ij}(t)$ into the equation for $\dot{p}_j(t)$ and linearizing, one obtains

$$\begin{aligned}\dot{p}_j(t)=f_j(t)-\frac{(\varepsilon_j+\beta_j)}{\Lambda_j}p_j+\frac{W_j*p_j}{\Lambda_j}+\left(\sum_{i=1}^{6}\frac{\beta_j\lambda_{ij}}{\Lambda_j}e^{-\lambda_{ij}t}\right)*p_j\\ +\frac{1}{\Lambda_j}\sum_{k=1}^{M}\varepsilon_{kj}\frac{P_{k0}}{P_{j0}}h_{kj}*p_k,\end{aligned}\tag{10}$$

with $p_j(0)$, $j=1,\cdots,N$ given.

Corollary 1 is now applied, with $x_j=p_j$ and $n_j=1$ for all j, with

$$\begin{aligned}(H_jx)(t)&=-\left(\frac{\varepsilon_j+\beta_j}{\Lambda_j}\right)x(t)\\ &\quad+\frac{1}{\Lambda_j}\int_0^t\left[W_j(t-s)+\sum_{i=1}^{6}\beta_{ij}\lambda_{ij}\,e^{-\lambda_{ij}(t-s)}\right]x(s)\,ds,\end{aligned}$$

$$f_j(t)=\sum_{i=1}^{6}c_{ij}(0)\,e^{-\lambda_{ij}t},$$

and

$$B_{jk}(t,x_t)=\frac{1}{\Lambda_j}\frac{P_{k0}}{P_{j0}}(h_{kj}*x)(t).$$

The resolvent R_j has Laplace transform $R_j^*(s)=1/D_j(s)$ and is stable if and only if

$$D_j(s)\equiv s+\left(\frac{\varepsilon_j+\beta_j}{\Lambda_j}\right)-\sum_{i=1}^{6}\frac{\beta_{ij}\lambda_{ij}}{\Lambda_j(s+\lambda_{ij})}-\frac{W_j^*(s)}{\Lambda_j}\neq 0$$

in the half plane $\mathrm{Re}\,s\geqq 0$. This condition can be checked graphically by plotting the Fourier transform of $D_j(s)$. *Moreover*, $1/g(R_j)$ is equal to the minimum distance from the graph of $D_j(j\omega)$, $-\infty<\omega<\infty$, to the origin in the complex plane. If the successive principal minors of the test matrix $M=I-[g(R_i)g(B_{ij})]$ are all positive, then for all initial values $p_j(0)$ and $c_{ij}(0)$ the solutions of (9) are in $L_2\cap C_0$. Moreover, they are stable in the sense that given $\varepsilon>0$ there exists $\delta>0$

such that when $|p_j(0)| \leqq \delta$ and $|c_{ij}(0)| \leqq \delta$ for all i and j, then $|p_j(t)| \leqq \varepsilon$ and $|c_{ij}(t)| \leqq \varepsilon$ for all i and j and for all $t \geqq 0$.

If $D_j(j\omega) \neq 0$ for $-\infty < \omega < \infty$ and $j = 1, \cdots, N$ but $D_j(s_0) = 0$ for some s_0, Re $s_0 > 0$ and some j, and if the matrix $M = [g(B_{ij})g_c(R_j)]$ has spectral radius less than one (e.g., weak interconnections between cores), then Theorem 4 implies that (10) is L_2-unstable. Since for this type of linear systems, L_2-stability and asymptotic stability in the sense of Lyapunov are equivalent (see [12]), this implies that (10) is Lyapunov unstable. The results in [14] imply the fact that when a linearized equation of the form (10) is Lyapunov unstable, this instability will carry over to the corresponding unlinearized equations (9). Thus, the solutions of (9) are unstable.

5. Proof of main results. Presently the results of § 3 are proved.

Proof of Theorem 1. The variation of constants formula (V) applies. Thus,

$$x(t) = r(t)x(0) + Rf(t) + RB(t, x_t) \tag{11}$$

on $0 \leqq t < \infty$. For any $T > 0$ define vectors $\|X\|_T = [\|x_{iT}\|_2]'$, $\|X(0)\| = [|x_i(0)|]'$ and $\|F\|_2 = [\|f_i\|_2]'$. Then

$$\|X\|_T \leqq [\|r_{ij}\|_2]\|X(0)\| + g(R)\|F\|_2 + g(R)g(B)\|X\|_T,$$

or

$$[I - g(R)g(B)]\|X\|_T \leqq [\|r_{ij}\|_2]\|X(0)\| + g(R)\|F\|_2.$$

Since the successive principal minors of the test matrix $I - g(R)g(B)$ are all positive, this test matrix has an inverse ρ whose entries are all nonnegative (see [3], [15]). Thus,

$$\|X\|_T \leqq \rho([\|r_{ij}\|_2]\|X(0)\| + g(R)\|F\|_2). \tag{12}$$

Since (12) is true for all $T > 0$, the proof is complete.

Proof of Theorem 2. The variation of constants formula (V′) applies. Thus,

$$x(t) = r(t)x(0) + Rf(t) + R[B_1(t, x(t)) - h_1 B_2(t, x(t))] + \dot{r} * B_2(t, x(t))$$

on R^+. For any $T > 0$,

$$\|X\|_T \leqq [\|r_{ij}\|_2]\|X(0)\| + g(R)\|F\|_2 + g(R)g(B_1 - hB_2)\|X\|_T + g(\dot{r})g(B_2)\|X\|_T$$

or

$$[I - g(R)g(B_1 - h_1 B_2) - g(\dot{r})g(B_2)]\|X\|_T \leqq [\|r_{ij}\|_2]\|x(0)\| + g(R)\|F\|_2.$$

The assumed condition on the test matrix implies (as in the proof of Theorem 1) that it has an inverse ρ whose entries are all nonnegative. Thus,

$$\|X\|_T \leq \rho([\|r_{ij}\|_2]\|X(0)\| + g(R)\|F\|_2)$$

for all $T > 0$. This completes the proof.

Proof of Theorem 3. The proofs under the hypotheses of Theorem 1 and of Theorem 2 are essentially the same, so that only one case will be considered. Assuming the hypotheses of Theorem 1, it will be shown that (5) is asymptotically stable in the sense of Lyapunov.

Note that if $s, \varphi \in L_2$, then the convolution $s * \varphi \in L_\infty$ and indeed by the Schwarz inequality, $|(s * \varphi)(t)| \leqq \|s\|_2 \|\varphi\|_2$. Moreover, $r * \varphi \in C_0$ (see for example Rudin [19, p. 4]). Since R is L_2-stable, the kernel $r \in L_2 \cap C_0$ and (11) implies

$$|x(t)| \leqq [\|r_{ij}\|_\infty] \|X(0)\| + [\|r_{ij}\|_2] \|F\|_2 + [\|r_{ij}\|_2] g(B) \|X\|_2.$$

The last inequality and (12) imply that

$$|x(t)| \leqq [\|r_{ij}\|_\infty] \|x(0)\| + [\|r_{ij}\|_2] \|F\|_2 + [\|r_{ij}\|_2] g(B) \rho([\|r_{ij}\|_2] \|X(0)\| + g(R) \|F\|_2).$$

This proves the Lyapunov stability.

To show that $x \in C_0$, use (11). Since $r \in C_0$ and since $s * \varphi \in C_0$ whenever $s, \varphi \in L_2$, (11) implies that $x \in C_0$ for all $x(0) \in R^m$ and all $f \in L_2$.

Proof of Theorem 4. Equations (6) can be written in the equivalent form

$$e_k(t) = f_k(t) + \sum_{j=1}^{N} B_{kj}(t, x_{jt}), \tag{13}$$

$$\dot{x}_k(t) = H_k x_k(t) + e_k(t), \tag{14}$$

$x_k(0) = x_{k0}$, $k = 1, \cdots, N$. Choosing $x_{k0} = 0$ for all k and integrating (14), one obtains $x_k = R_k e_k$. This equation along with (13) are in the proper form for Theorem 5 to apply.

Proof of Theorem 5. The proof is similar to the proof of Theorem 3 in [22]. For purposes of contradiction, assume that $x_k \in L_2$ for all k. If $\rho(M) < 1$, then one can replace each value $g_c(H_i)$ by $g_c(H_i) + A$ and each value $g(B_{ij})$ by $g(B_{ij}) + A$, where A is small and positive. If A is sufficiently small, the new matrix M will still have $\rho(M) < 1$ but now all entries of $[m_{ij}]$ are positive. Assume that this replacement has been accomplished. Then $m_{ij} > 0$ for all i and j and by the Perron theorem (see, e.g., [4, p. 53]), there exist numbers $\lambda_j > 0$ such that $\Lambda = [\lambda_1, \cdots, \lambda_m]$ is a row eigenvector corresponding to the dominant eigenvalue $\rho(M)$ of (M).

Put $z_i = \sum_{j=1}^{m} B_{ij} x_j = \sum_{j=1}^{m} B_{ij} R_j (e_j + v_j)$. Then $z_i = f_i + v_i - (e_i + v_i)$. Since $(f_i + v_i) \in X_i^\perp$ and $(e_i + v_i) \in X_i$, it follows that $z_i \in L_2$ and

$$\|z_i\|_2^2 = \|f_i + v_i\|_2^2 + \|e_i + v_i\|_2^2 \geqq \|e_i + v_i\|_2^2.$$

Since $\|f_i + v_i\| > 0$ for at least one $i = k_0$ and since all λ_i are positive, it follows that

$$\sum_{i=1}^{m} \lambda_i \|z_i\|_2 > \sum_{i=1}^{m} \lambda_i \|e_i + v_i\|_2. \tag{15}$$

The definition of z_i and the assumptions given yield the estimate

$$\begin{aligned}\|z_i\|_2 &\leqq \sum_{j=1}^{m} \|B_{ij}x_j\|_2 \leqq \sum_{j=1}^{m} g(B_{ij})\|x_j\|_2 \\ &\leqq \sum_{j=1}^{m} g(B_{ij})g_c(R_j)\|e_j + v_j\|_2 = \sum_{j=1}^{m} m_{ij}\|e_j + v_j\|_2.\end{aligned} \tag{16}$$

Since Λ is a row eigenvector of M it follows that

$$\sum_{i=1}^{m} \lambda_i m_{ij} = \rho(M)\lambda_j$$

for all j. Thus, (16) and $\rho(M) \leqq 1$ imply that

$$\begin{aligned}\sum_{j=1}^{m} \lambda_i\|z_i\|_2 &\leqq \sum_{i=1}^{m}\sum_{j=1}^{m} \lambda_i m_{ij}\|e_j + v_j\|_2 \\ &= \sum_{j=1}^{m} \rho(M)\lambda_j\|e_j + v_j\|_2 \leqq \sum_{j=1}^{m} \lambda_j\|e_j + v_j\|_2.\end{aligned}$$

This contradicts (15) and the theorem is proved.

6. Concluding remarks. At this point some additional comments are in order.

1. It is emphasized that the present results yield not only L_2-stability and L_2-instability conditions for systems described by integrodifferential equations, but also asymptotic stability conditions in the sense of Lyapunov.

2. Some (but not all) of the results presented offer the advantage of allowing stability analysis of complex systems in terms of lower order subsystems and in terms of system interconnecting structure.

3. It is possible to analyze large classes of problems by the present results, using frequency domain techniques (with all the practical advantages which such techniques offer).

REFERENCES

[1] Z. AKCASU, G. S. LILLOUCHE AND L. M. SHOTKIN, *Mathematical Methods in Nuclear Reactor Dynamics*, Academic Press, New York, 1971.

[2] C. A. DESOER AND M. VIDYASAGAR, *Feedback Systems: Input–Output Properties*, Academic Press, New York, 1975.

[3] M. FIEDLER AND V. PTAK, *On matrices with non-positive off-diagonal elements and positive principal minors*, Czechoslovak Math. J., 12 (1962), pp. 382–400.

[4] F. R. GANTMACHER, *Matrix Theory*, vol. II, Chelsea, New York, 1959.

[5] S. I. GROSSMAN AND R. K. MILLER, *Perturbation theory for Volterra integrodifferential systems*, J. Differential Equations, 8 (1971), pp. 457–474.

[6] ———, *Nonlinear Volterra integrodifferential systems with L^1-kernels*, Ibid., 13 (1973), pp. 551–566.

[7] G. S. JORDAN AND R. L. WHEELER, *Asymptotic behavior of unbounded solutions of linear Volterra integral equations*, J. Math. Anal. Appl., to appear.

[8] E. L. LASLEY AND A. N. MICHEL, *Input–output stability of large-scale systems*, Proc. Eighth Asilomar Conf. on Circuits, Systems, and Computers, Western Periodicals, North Hollywood, CA, 1975, pp. 472–482.

[9] ———, *Input–output stability of interconnected systems*, Proc. IEEE International Symp. on Circuits and Systems, Boston, 1975, pp. 131–134.

[10] ———, *L_∞- and l_∞-stability of interconnected systems*, IEEE Trans. Circuits Systems 23 (1976), pp. 261–270.

[11] R. K. MILLER, *Nonlinear Volterra Integral Equations*, W. A. Benjamin, Menlo Park, Calif., 1971.

[12] ———, *Asymptotic stability properties of linear Volterra integrodifferential equations*, J. Differential Equations, 10 (1971), pp. 485–506.

[13] ———, *Structure of solutions of unstable linear Volterra integrodifferential equations*, Ibid., 15 (1974), pp. 129–157.

[14] R. K. MILLER AND J. A. NOHEL, *A stable manifold theorem for a system of Volterra integrodifferential equations*, this Journal, 6 (1975), pp. 506–522.

[15] A. OSTROWSKI, *Determinanten mit überwiegender Hauptdiagonale und die absolute Konvergenz von linearen Iterationprozessen*, Comment. Math. Helv., 30 (1956), pp. 175–210.

[16] H. PLAZA AND W. H. KOHLER, *Coupled-reactor kinetics equations*, Nuclear Sci. and Engrg., 22 (1966), pp. 419–422.

[17] D. W. PORTER AND A. N. MICHEL, *Stability of multiple-loop nonlinear time-varing systems*, Rep. ISU-ERI-Ames-73167, Iowa State University, Ames, 1973.

[18] ———, *Input–output stability of time-varing nonlinear multiloop feedback systems*, IEEE Trans. Automatic Control, 19 (1974), pp. 422–427.

[19] W. RUDIN, *Fourier Analysis on Groups*, Interscience, New York, 1962.

[20] I. W. SANDBERG. *On the L_2-boundedness of solutions of nonlinear functional equations*, Bell System Tech. J., 43 (1974), pp. 1581–1599.

[21] ———, *Some results on the theory of physical systems governed by nonlinear functional equations*, Ibid., 44 (1965), pp. 871–898.

[22] M. VIDYASAGAR, *L_2-stability and L_2-instability of interconnected feedback systems*, Proc. IEEE, SIAM J. Control, to appear.

[23] J. C. WILLEMS, *The Analysis of Feedback Systems*, MIT Press, Cambridge, Mass., 1971.

[24] G. ZAMES, *On the input–output stability of time-varying nonlinear feedback systems—part I: Conditions derived using the concepts of loop gain, conicity and positivity*, IEEE Trans. Automatic Control, 11 (1966), pp. 228–238.

[25] ———, *On the input-output stability of time-varying nonlinear feedback systems—part II: Conditions involving circles in the frequency plane and sector non-linearities*, IEEE Trans. Automatic Control, 11 (1966), pp. 465–476.

SIAM J. MATH. ANAL.
Vol. 8, No. 3, May 1977

ON THE PRODUCTS OF SOLUTIONS OF SECOND ORDER DISCONJUGATE DIFFERENTIAL EQUATIONS AND THE WHITTAKER DIFFERENTIAL EQUATION*

PHILIP HARTMAN†

Abstract. Let $x''-q(t)x=0$ be a disconjugate differential equation on $(-\infty\leqq)\alpha<t<\omega(\leqq\infty)$. Let $x=x_0(t)>0$ be a principal solution at $t=\omega$ and $x=y(t)>0$ a linearly independent solution (e.g., a principal solution at $t=\alpha$). The paper deals with an investigation of the monotony and convexity properties of the product $u=x_0y$ under suitable conditions on q. Applications are made to the modified Bessel equation (with $u=tI_\nu(t)K_\nu(t)$) and to the Whittaker differential equation.

1. Introduction. The first part of [3] concerns sufficient conditions on a positive monotone q to assure that the differential equation $x''+q(t)x=0$ has a pair of solutions $x(t)$, $y(t)$ such that $w=[x^2(t)+y^2(t)]^{1/2}>0$ satisfies $w>0$, $w'\leqq 0$, $w''\geqq 0$ or satisfies $w>0$, $w'\geqq 0$, $w''\leqq 0$. We deal here with an analogue of this question in which we consider a differential equation

$$x''-q(t)x=0, \tag{1.1}$$

where $q(t)$ is positive and continuous for $t>0$. In this case, (1.1) has a principal solution $x_0(t)$, unique up to multiplicative constants, determined by the inequalities

$$x_0>0, \quad x_0'<0, \quad x_0''>0 \tag{1.2}$$

(A. Kneser; cf., e.g., [4, p. 357]). Also if a solution $x=y(t)$ satisfies

$$y>0, \quad y'>0, \quad y''>0 \tag{1.3}$$

at some $t=t_0>0$, then (1.3) holds for $t\geqq t_0$. In this situation, it is natural to ask for conditions on $q(t)$ which imply that, for some $y(t)$ (say, a principal solution $x=y(t)$ at $t=0$), the product $u=x_0y$ has specified monotony and convexity properties. This question is suggested by a result in [5] on modified Bessel functions corresponding to the case

$$q(t)=1+\beta/t^2, \qquad \beta=\nu^2-\tfrac{1}{4}, \tag{1.4}$$

which states that if $\nu\geqq\frac{1}{2}$, then (1.1) has the solutions $x_0(t)=t^{1/2}K_\nu(t)$, $y(t)=t^{1/2}I_\nu(t)$ with the properties that $u=u_\nu(t)=tI_\nu(t)K_\nu(t)$ satisfies

$$u>0, \quad u'>0 \quad \text{for } t>0, \quad \nu\geqq\tfrac{1}{2}. \tag{1.5}$$

Here $t^{1/2}K_\nu$, $t^{1/2}I_\nu$ are standard solutions of (1.1), (1.4) and are principal solutions at $t=\infty$, $t=0$; cf. [10, pp. 77–80].

We discuss general monotony properties of products $u=x_0y$ of solutions of (1.1) in § 2 and general convexity properties in § 3. Section 4 deals with the modified Bessel equation (1.1), (1.4), where we show that if $\nu>\frac{1}{2}$, then there exists

* Received by the editors June 26, 1975.

† Mathematics Department, Johns Hopkins University, Baltimore, Maryland 21218. This work was supported by the National Science Foundation under Grant GP-MPS71-03219A03.

$\tau = \tau_\nu > 0$ such that

$$u > 0, \quad u' > 0, \quad u'' < 0 \quad \text{for } t > 0,$$

$$u''' < 0 \quad \text{for } 0 < t < \tau, \quad u''' > 0 \quad \text{for } t > \tau,$$

and if $0 < \nu < \frac{1}{2}$, then there exist numbers τ_1, τ_2 such that $0 < (-\beta)^{1/2} < \tau_1 < \tau_2$,

$$u > 0 \quad \text{for } t > 0, \quad u' > 0 \quad \text{for } 0 < t < \tau_1, \quad u' < 0 \quad \text{for } t > \tau_1,$$

$$u'' < 0 \quad \text{for } 0 < t < \tau_2, \quad u'' > 0 \quad \text{for } t > \tau_2.$$

Section 5 concerns the more general Whittaker differential equation.

2. Monotony of $u = x_0 y$. In what follows, we assume that q in

$$x'' - q(t)x = 0 \tag{2.1}$$

is continuous on $(-\infty \leqq)\alpha < t < \omega (\leqq \infty)$. Often we assume that $q \geqq 0$ is monotone, $q \not\equiv$ const. near ω, and that

$$\int^{\omega} q(t)\, dt = \infty \quad \text{if } \omega < \infty. \tag{2.2}$$

Thus in the case $\omega < \infty$, q is nondecreasing and $q(\omega) = \infty$. Furthermore, (2.2) implies that (2.1) has a solution $x = x_0(t)$, unique up to constant factors, satisfying (1.2); cf. [4, Exercise 6.7, p. 358]. Below $x = y(t)$ is a linearly independent solution which is positive for t near ω, so that (1.3) holds for t near ω. *Sometimes we assume the existence of such a* $y(t) > 0$ *on* (α, ω).

We shall use the notation

$$r_0 = x_0'/x_0 \quad \text{and} \quad r = y'/y, \tag{2.3}$$

when x_0, $y > 0$, so that we have the Riccati equations

$$r_0' = q - r_0^2 \quad \text{and} \quad r' = q - r^2. \tag{2.4}$$

Also, we define

$$h_0 = (x_0')^2 - x_0^2 q = x_0^2(r_0^2 - q), \qquad h = (y')^2 - y^2 q = y^2(r^2 - q), \tag{2.5}$$

so that, when q is monotone,

$$dh_0 = -x_0^2\, dq \quad \text{and} \quad dh = -y^2\, dq, \tag{2.6}$$

and h_0 and h are monotone. For reference, we state

PROPOSITION 2.1. *Let* $q \geqq 0$ *be continuous on* (α, ∞), $q(\infty) = \lim q(t)$ *exists as* $t \to \infty$, $0 \leqq q(\infty) \leqq \infty$, *and* $x = x_0(t)$, $y(t)$ *be solutions of* (2.1) *as specified above. Then* $x_0'/x_0 \to -q^{1/2}(\infty)$ *and* $y'/y \to q^{1/2}(\infty)$, *as* $t \to \infty$. *Also* $u(\infty) = \lim u(t)$ *exists as* $t \to \infty$, *with* $0 \leqq u(\infty) \leqq \infty$; *furthermore,* $u(\infty) = 0$, $0 < u(\infty) < \infty$, *or* $u(\infty) = \infty$ *according as* $q(\infty) = \infty$, $0 < q(\infty) < \infty$, *or* $q(\infty) = 0$.

The first part of this statement, concerning the limits of x_0'/x_0 and y'/y, goes back to Poincaré and Perron. The second part follows from an application of l'Hôpital's rule to

$$x_0(t)y(t) = (\text{const.}) \int_t^{\infty} y^{-2}(s)\, ds / y^{-2}(t),$$

as in [2, p. 573].

PROPOSITION 2.2. *Let* $q \geqq 0$ *be continuous on* (α, ω), *nonincreasing* [*or nondecreasing*], $q \not\equiv$ const. *near* ω *and satisfy* (2.2). *Then*

$$0 > x_0'/x_0 > -q^{1/2} \quad [\text{or } x_0'/x_0 < -q^{1/2} < 0] \tag{2.7}$$

for $\alpha < t < \omega$. *Also, if* (1.3) *and*

$$y'/y > q^{1/2} > 0 \quad [\text{or } 0 < y'/y < q^{1/2}] \tag{2.8}$$

hold for some $t = t_0$, *then they hold for* $t_0 \leqq t < \omega$, *and, in any case,* (1.3) *and* (2.8) *hold near* ω. *Hence if* (1.3) *and* (2.8) *hold at* $t = t_0$, *then for* $t_0 \leqq t < \omega$,

$$u = x_0 y > 0 \quad \textit{and} \quad u' > 0 \quad [\text{or } u' < 0]. \tag{2.9}$$

Note that (2.9) follows from (2.7) and (2.8) since

$$u' = u(r_0 + r), \quad \text{where } r_0 = x_0'/x_0, \quad r = y'/y. \tag{2.10}$$

If $\omega = \infty$, this proposition is essentially contained in [7] (cf. [4, Exercise 3.9, pp. 514 and 579]). If $\omega < \infty$, a similar proof is valid if one notes that when (2.2) holds, then there exists only one solution $x = x_0(t)$, up to multiplicative constants, satisfying (1.2) for t near ω.

Remark. If we omit the assumption "$q \not\equiv$ const. near ω", the strict inequalities in (2.7) and (2.8) should be replaced by the corresponding weak inequalities. Similar comments apply throughout this paper.

PROPOSITION 2.3. *Let* q *be continuous and* (2.1) *disconjugate on* (α, ω). *Let* $x = x_0(t) > 0$ *be a principal solution at* $t = \omega$ *and* $x = y(t)$ *a positive solution on* (α, ω) *linearly independent of* $x_0(t)$.

(i) *Let* q *satisfy the conditions of Proposition* 2.2. *Then* $u = x_0 y > 0$, $u' > 0$ [*or* $u' < 0$] *near* ω, u' *has at most one zero, say* $t = \tau$, *on* (α, ω) *and* u' *changes signs at* $t = \tau$ *if it exists.*

(ii) *Assertion* (i) *remains valid if* $-\infty < \alpha$ *and* (α, ω) *is replaced by* $[\alpha, \omega)$.

(iii) *Let* $\alpha < t_+ < \omega$. *Let* q *satisfy the conditions above with* (α, ω) *replaced by* (t_+, ω) *and* $q < 0$ *on* (α, t_+), *so that* $q \geqq 0$ *is nondecreasing on* (t_+, ω). *Then the conclusions of* (i) *remain valid.*

(iv) *In particular, in* (i)–(iii),

$$\begin{aligned} &u'(t) \leqq 0 \quad \textit{for some } t \textit{ near } \alpha \ (\textit{e.g., } u(\alpha+) = \infty) \\ &\qquad \Rightarrow u' < 0 \textit{ on } (\alpha, \tau) \textit{ and } u' > 0 \textit{ on } (\tau, \omega) \ [\textit{or } u' < 0 \textit{ on } (\alpha, \omega)], \end{aligned} \tag{2.11}$$

$$\begin{aligned} &u'(t) \geqq 0 \quad \textit{for some } t \textit{ near } \alpha \ (\textit{e.g., } u(\alpha+) = 0) \\ &\qquad \Rightarrow u' > 0 \textit{ on } (\alpha, \omega) \ [\textit{or } u' > 0 \textit{ on } (\alpha, \tau) \textit{ and } u' < 0 \textit{ on } (\tau, \omega)]. \end{aligned} \tag{2.12}$$

Proof. (i) We consider only the case of nondecreasing q, as the other case is similar. It follows from Proposition 2.2 that $x_0'/x_0 < -q^{1/2} \leqq 0$ on (α, ω) and that $0 < y'/y < q^{1/2}$ near ω. Thus $h < 0$ near ω and, since $dh \leqq 0$ by (2.6), we have two possibilities: either

$$h < 0 \quad \text{on } (\alpha, \omega) \quad \text{or} \quad h \geqq 0 \quad \text{on } (\alpha, t_1) \quad \text{and} \quad h < 0 \quad \text{on } (t_1, \omega) \tag{2.13}$$

for some $t_1 \in (\alpha, \omega)$. In the first case, $u' < 0$ on (α, ω) by (2.10). In the second case,

$u'<0$ on (t_1, ω) and $(r_0+r)'<0$ on (α, t_1) since $r_0'<0$ and $r'\leqq 0$. Differentiating (2.10) gives

$$(u'/u)'=(r_0+r)'. \tag{2.14}$$

As $u>0$ and the right side is negative on (α, t_1), it follows that u' has at most one zero on (α, t_1).

(ii) This is clear from the arguments above.

(iii) Similar arguments give $x_0'/x_0<-q^{1/2}\leqq 0$ on (t_+, ω) and there exists $t_1\in[t_+, \omega)$ such that $h\geqq 0$ on $[t_+, t_1]$ and $h<0$ on (t_1, ω). Since $q<0$ implies $r_0'<0$ and $r'<0$ on (α, t_+), the conclusions follow from the arguments in case (i).

Proposition 2.2 with $\omega=\infty$ was essentially used in [5] in the proof of (1.5). Actually, this proposition, with both $\omega=\infty$ and $\omega<\infty$, implies (1.5) without using, as in [5], the asymptotic behavior of $I_\nu(t)$, $K_\nu(t)$ at $t=0$ to verify (2.7), (2.8) for small $t>0$. More generally, we have

THEOREM 2.1. *Let $q>0$ be continuous, nonincreasing for $(-\infty\leqq)\alpha<t<\infty$, $q\not\equiv$ const. near α and ∞, and*

$$\int_\alpha q\,dt=\infty \quad \text{if } \alpha>-\infty. \tag{2.15}$$

Then (2.1) has solutions $x=x_0(t)$ and $x=y(t)$, unique up to multiplicative constants, satisfying (1.2) and (1.3) on (α, ∞). They also satisfy $0>x_0'/x_0>-q^{1/2}$, $y'/y>q^{1/2}>0$, $u=x_0y>0$ and $u'>0$ on (α, ∞), and $0<u(\infty)<\infty$ or $u(\infty)=\infty$ according as $q(\infty)>0$ or $q(\infty)=0$.

This is a consequence of Proposition 2.2. In fact, one obtains $x_0(t)$ by a direct application of this proposition with "nonincreasing q", while $y(t)$ is obtained by an application of the case "nondecreasing q" after the change of independent variable $t\to -t$ (and $y(t)=x_0(-t)$ in terms of the corresponding $x_0(t)$).

THEOREM 2.2. *Let $-\infty<t_0<\infty$. Let $q\geqq 0$ be continuous on $(-\infty, \infty)$, nondecreasing [or nonincreasing] on $(-\infty, t_0)$ and nonincreasing [or nondecreasing] on (t_0, ∞), and $q\not\equiv$ const. for large $t>0$ and large $-t>0$. Let $x=x_0(t)>0$, $y(t)>0$ be principal solutions at $t=\infty$, $t=-\infty$. Then there exists a number τ such that $u=x_0y>0$ satisfies $u'<0$ [or $u'>0$] on $(-\infty, \tau)$, $u'>0$ [or $u'<0$] on (τ, ∞).*

Remark. In the first [i.e., unbracketed] case, ∞[or $-\infty$] can be replaced by $\omega<\infty$ [or $\alpha>-\infty$] if (2.2) [or (2.15)] holds and $q\not\equiv$ const. near ω[or α].

Proof. Consider only the first [unbracketed] case, as the other is similar. By Proposition 2.3 (ii), $u'>0$ near ∞ and u' has at most one zero on $[t_0, \infty)$. Replacing t by $-t$ and applying the same proposition, we see that $u'<0$ near $-\infty$ and u' has at most one zero on $(-\infty, t_0]$. This implies the theorem.

THEOREM 2.3. *Let $\alpha<t_-<t_+<\omega$. Let q be continuous and (2.1) disconjugate on (α, ω), $q\geqq 0$ nonincreasing on (α, t_-), $q<0$ on (t_-, t_+), $q\geqq 0$ nondecreasing on (t_+, ω), $q\not\equiv$ const. near α and ω, (2.2) and (2.15) hold. Let $x=x_0(t)>0$, $y(t)>0$ be principal solutions at $t=\omega$, $t=\alpha$. Then there exists $\tau\in(\alpha, \omega)$ such that $u=x_0y>0$ satisfies $u'>0$ on (α, τ) and $u'<0$ on (τ, ω).*

Proof. The proof is similar to that of the last theorem except that we apply Proposition 2.3 (iii) on (α, t_+) and (t_-, ω).

In § 5 on the Whittaker functions, we shall have to deal with the situations in the following two theorems.

PROPOSITION 2.4. *Let q be continuous on (α, ω), (2.1) disconjugate on (α, ω), $q \not\equiv$ const. near ω, and (2.2) holds. Let $\alpha < t_- < t_+ < \omega$. Let $q \geqq 0$ be nonincreasing on (α, t_-), $q < 0$ on (t_-, t_+), $q \geqq 0$ nondecreasing on (t_+, ω). Let $x = x_0(t) > 0$ be a principal solution of (2.1) at $t = \omega$ and $x = y(t) > 0$ a solution linearly independent of $x_0(t)$. Then $u' < 0$ near ω and the set $\{t: u' = 0\}$ consists of 0, 1 or 2 subintervals of (α, ω). These subintervals can reduce to points, and do reduce to points if q is strictly monotone on (α, t_-). If $u' > 0$ for some t near α, then $\{t: u' = 0\}$ consists of exactly one interval (or point).*

Proof. We shall sketch the proof leaving details to the reader. In this proof, we use the fact that $x_0 y' - x_0' y$ is a constant (> 0), so that $r > r_0$.

Applying the proof of Proposition 2.3(iii) to the interval (t_-, ω), we obtain $t_1 \in [t_+, \omega)$ such that $(r_0 + r)' < 0$ on $(t_-, t_1]$ and $u' < 0$ on (t_1, ω). Also, $h_0 > 0$ on (t_-, ω), $h \geqq 0$ on $(t_-, t_1]$.

Since $dh_0 \geqq 0$, $dh \geqq 0$ on $(\alpha, t_-]$, we have the following possibilities: either

$$h_0 \geqq 0 \quad \text{on } (\alpha, t_-) \tag{2.16}$$

or there is a $t_2 \in (\alpha, t_-)$ such that

$$h_0 < 0 \quad \text{on } (\alpha, t_2), \qquad h_0 \geqq 0 \quad \text{on } [t_2, t_1], \tag{2.17}$$

and either

$$h \geqq 0 \quad \text{on } (\alpha, t_-) \tag{2.18}$$

or there is a $t_3 \in (\alpha, t_-]$ such that

$$h < 0 \quad \text{on } (\alpha, t_3), \qquad h \geqq 0 \quad \text{on } [t_3, t_1]. \tag{2.19}$$

Case 1. On (2.16), (2.18). In this case, $(r + r_0)' \leqq 0$ on (α, t_-), hence, on (α, t_1). Thus $\{t: u' = 0\}$ is at most one interval in $(\alpha, t_-]$.

Case 2. On (2.16), (2.19). In this case, $(r + r_0)' \leqq 0$ on $[t_3, t_1]$. On (α, t_3), $-q^{1/2} < r < q^{1/2}$ and $|r_0| \geqq q^{1/2}$. By $r > r_0$, $r_0 \leqq -q^{1/2} < r < q^{1/2}$, so that $r_0 + r < 0$. Hence $u' < 0$ on (α, t_3). Consequently, $u' < 0$ on (α, ω).

Case 3. On (2.17), (2.18). On (α, t_2), $-q^{1/2} < r_0 < q^{1/2} \leqq r$ and so, $u' > 0$. On $[t_2, t_1]$, $(r + r_0)' \leqq 0$. Thus $\{t: u' = 0\}$ consists of exactly one interval in $[t_2, t_-]$.

Case 4. On (2.17), (2.19), $t_3 > t_2$. On $[t_3, t_1]$, $(r + r_0)' \leqq 0$. On $[t_2, t_3)$, $r_0 \leqq -q^{1/2} < r < q^{1/2}$, so that $u' < 0$ on $[t_2, t_3)$. Hence $u' < 0$ on $[t_2, \omega)$. Finally, on (α, t_2), $(r + r_0)' > 0$. Consequently, $u' < 0$ on (α, ω).

Case 5. On (2.17), (2.19), $t_3 \leqq t_2$. On $[t_2, t_1)$, $(r + r_0)' \leqq 0$. On (t_3, t_2), $-q^{1/2} < r_0 < q^{1/2} \leqq r$, so that $u' > 0$ on $(t_3, t_2]$ and $\{t: u' = 0\} \cap (t_3, \omega)$ is a subinterval of $[t_2, t_1]$. On (α, t_3), $(r + r_0)' > 0$ and $\{t: u' = 0\} \cap (\alpha, t_3]$ is either empty or is an interval. This completes the proof.

PROPOSITION 2.5. *Let q be continuous and (2.1) disconjugate on (α, ∞), and $q \not\equiv$ const. near ∞. Let $\alpha < t_+ < t_0 < \infty$. Let $q < 0$ on (α, t_+), $q \geqq 0$ nondecreasing on (t_+, t_0), $q > 0$ nonincreasing on (t_0, ∞). Let $x = x_0(t)$, $y(t)$ be as in Proposition 2.4. Then $u' > 0$ for large t and the set $\{t: u' = 0\}$ consists of 0, 1, or 2 subintervals of (α, ∞). These subintervals can reduce to points, and do reduce to points if q is increasing on (t_+, t_0). If $u' < 0$ for some t near α, then $\{t: u' = 0\}$ consists of exactly one interval (or point).*

Proof. We have $t_1 \in [t_+, t_0]$ such that $0 > r_0 > -q^{1/2}$ on (t_1, ∞), $r_0 \leqq -q^{1/2}$ on $[t_+, t_1]$, $r_0' < 0$ on (α, t_+). For y, we have the possibilities: either

$$r > q^{1/2} \quad \text{on } [t_+, \infty), \quad r' < 0 \quad \text{on } (\alpha, t_+) \tag{2.20}$$

or there exists t_2 and t_3 such that $t_+ \leqq t_2 \leqq t_0 \leqq t_3 < \infty$,

$$\begin{gathered} r > q^{1/2} \quad \text{on } (t_3, \infty), \quad |r| \leqq q^{1/2} \quad \text{on } [t_2, t_3], \\ r' < 0 \quad \text{on } (\alpha, t_2). \end{gathered} \tag{2.21}$$

Case 1. On (2.20). On (t_1, ∞), we have $0 > r_0 > -q^{1/2}$ and $r \geqq q^{1/2}$, so that $u' > 0$. On (α, t_1), $(r + r_0)' < 0$. Thus $u' > 0$ on (α, ∞).

Case 2. On (2.21), $t_1 < t_2$. On (t_3, ∞), we have $u' > 0$. On $[t_2, t_3]$, $(r + r_0)' \leqq 0$. On (t_1, t_2), $-q^{1/2} \leqq r_0 \leqq q^{1/2} < r$, so that $u' > 0$ on (t_1, ∞). On (α, t_1), $(r + r_0)' \leqq 0$. Thus $u' > 0$ on (α, ∞).

Case 3. On (2.21), $t_1 \geqq t_2$. On (t_3, ∞), $u' > 0$. On $[t_1, t_3]$, we have $(r + r_0)' \leqq 0$. On (t_2, t_1), $q^{1/2} \geqq r \geqq -q^{1/2} > r_0$, so that $u' < 0$. Hence $\{t: u' = 0\} \cap (t_2, \infty)$ consists of an interval in $[t_1, t_3]$. On (α, t_2), $(r + r_0)' \leqq 0$ and so $\{t: u' = 0\} \cap (\alpha, t_2)$ is either empty or is an interval. This completes the proof.

Remark. Results of this section can be transferred from equations (2.1) to those of the form

$$x'' + p(t)x' - q(t)x = 0,$$

by transforming the latter into

$$d^2x/ds^2 - q(t) \exp\left(2 \int^t p(r)\, dr\right) x = 0,$$

where $ds = \exp\left(-\int^t p(r)\, dr\right) dt$.

3. Convexity of $u = x_0 y$. We investigate the sign of u'' by the use of the standard Liouville change of variables

$$z = q^{1/4}x \quad \text{and} \quad s = \int_{t_0}^{t} q^{1/2}(r)\, dr \tag{3.1}$$

and Appell's [1] equation

$$u''' = 4qu' + 2q'u = 4q^{1/2}(q^{1/2}u)', \tag{3.2}$$

satisfied by the product $u = xy$ of any pair of solutions $x(t)$, $y(t)$ of (2.1). By (3.1), (2.1) becomes

$$d^2z/ds^2 - Q(t)z = 0, \tag{3.3}$$

$$Q(t) = 1 + q''/4q^2 - (5q')^2/16q^3 = 1 + (q'/q^{5/4})'/4q^{3/4}, \tag{3.4}$$

and $t = t(s)$ is the inverse of $s = s(t)$ in (3.1). Thus the analogue of (2.2) for (3.3) is

$$\int^{\omega} Q(t)q^{1/2}(t)\, dt = \infty \quad \text{if} \int^{\omega} q^{1/2}(t)\, dt < \infty. \tag{3.5}$$

The solutions x_0, y of (2.1) become the solutions $q^{1/4}x_0$, $q^{1/4}y$ of (3.3), and their product is

$$U = q^{1/2}x_0 y = q^{1/2}u \quad \text{and} \quad dU/ds = q^{-1/2}(q^{1/2}u)'. \tag{3.6}$$

We verify that Propositions 2.1–2.3 have the following consequences.

PROPOSITION 3.1. *Let $q \in C^2(\alpha, \omega)$ be positive, satisfy* (2.2) *and* (3.5), *and have the properties that $Q \geqq 0$ is nonincreasing [or nondecreasing] on $[t_0, \omega)$, $\alpha < t_0 < \omega$, and $Q \not\equiv$ const. near ω. Then*

$$0 > x_0'/x_0 + q'/4q > -(Qq)^{1/2} \quad [\text{or } < -(Qq)^{1/2} < 0] \tag{3.7}$$

for $t \geqq t_0$. Also if (1.3) *and*

$$y'/y + q'/4q > (Qq)^{1/2} \quad [\text{or } < (Qq)^{1/2}] \tag{3.8}$$

hold for $t = t_0$, then they hold for $t_0 \leqq t < \omega$; and, in any case, they hold near ω. Hence if $q > 0$, $Q \geqq 0$ are nonincreasing [or nondecreasing] on $[t_o, \omega)$, not constant near ω, and if (1.3) *and*

$$u = x_0 y > 0 \quad \text{and} \quad u' > 0, \quad u''' > 0 \quad [\text{or } u' < 0, \quad u''' < 0] \tag{3.9}$$

hold at $t = t_0$, then they hold for $t_0 \leqq t < \omega$. If, in addition, $\omega = \infty$ and $0 < q(\infty) \leqq \infty$, then for $t_0 < t < \infty$,

$$\begin{gathered} u > 0 \quad \text{and} \quad u' > 0, \quad u'' < 0, \quad u''' > 0 \\ [\text{or } u' < 0, \quad u'' > 0, \quad u''' < 0]. \end{gathered} \tag{3.10}$$

Remark 1. If Q is monotone and $\omega = \infty$, then (3.4) shows that $q(t)$ is monotone for large t, and so $q(\infty) = \lim q(t)$ exists, $t \to \infty$, with $0 \leqq q(\infty) \leqq \infty$.

Remark 2. If $q \in C^3$, then $Q' \geqq 0$ if

$$q > 0, \quad q' \geqq 0, \quad q'' \leqq 0, \quad q''' \geqq 0, \tag{3.11}$$

since (3.4) gives

$$Q' = q'''/4q^2 - 9q'q''/8q^3 + (15q')^3/16q^4. \tag{3.12}$$

Thus if (3.11) holds, then $q > 0$, Q are nondecreasing.

Proof. The assertions concerning (3.7) and (3.8) follow by applying Proposition 2.2 to the equation (3.3), instead of (2.1). If (3.7) and (3.8) hold, then $(q^{1/2}u)' > 0$ [or < 0] and, by (3.2), $u''' > 0$ [or $u''' < 0$]. This, when combined with (2.9) in Proposition 2.2, implies (3.9). Note that the statements concerning the sign of u'' on (3.10), (3.11) follow from the fact that if $u(\infty)$ exists (finite) and $u'u''' > 0$, then $u'(\infty) = 0$ and $u'u'' < 0$.

THEOREM 3.1. *Let $q \in C^2(\alpha, \infty)$ satisfy the conditions of Theorem* 2.1 *and have the property that $Q \geqq 0$ is nonincreasing on (α, ∞), $Q \not\equiv$ const. near α and ω, and*

$$\int_\alpha Q(t)q^{1/2}(t)\,dt = \infty, \qquad \int^\infty q^{1/2}(t)\,dt = \infty.$$

Then the solutions $x = x_0(t)$, $y(t)$ of (2.1) *in Theorem* 2.1 *satisfy*

$$0 > x_0'/x_0 + q'/4q > -(Qq)^{1/2}, \qquad y'/y + q'/4q > (Qq)^{1/2} > 0$$

and $u = x_0 y > 0$, $u' > 0$ *and* $u''' > 0$ *on* (α, ∞); *also* $u'' < 0$ *on* (α, ∞) *if* $q(\infty) > 0$.

This is clear from Theorem 2.1 and Proposition 3.1. For applications to the modified Bessel equation, we need the following variant.

THEOREM 3.2. *Let* $q \in C^2(\alpha, \infty)$ *satisfy the conditions of Theorem* 2.1,

$$\int_\alpha q^{1/2}(t)\,dt = \infty \quad \textit{and} \quad \int^\infty q^{1/2}(t)\,dt = \infty,$$

and have the properties that $Q \geqq 0$ *on* (α, ∞) *and, for some* $t_0 \in (\alpha, \infty)$, Q *is nondecreasing on* $(\alpha, t_0]$ *and nonincreasing on* $[t_0, \infty)$, $Q \not\equiv$ const. *near* α *and* ω. *Then the solutions* $x = x_0(t)$, $y(t)$ *of* (2.1) *in Theorem* 2.1 *satisfy*

$$\begin{gathered} u = x_0 y > 0, \quad u' > 0 \quad \textit{on } (\alpha, \infty), \\ u''' < 0 \quad \textit{on } (\alpha, \tau) \quad \textit{and} \quad u''' > 0 \quad \textit{on } (\tau, \infty) \end{gathered} \tag{3.13}$$

for some $\tau \in (\alpha, \infty)$. *Furthermore,*

$$q(\infty) > 0 \Rightarrow u(\infty) < \infty \Rightarrow u'' < 0 \quad \textit{on } [\tau, \infty), \tag{3.14}$$

$$u'' \leqq 0 \textit{ for some } t \textit{ near } \alpha \textit{ (e.g., } u'(\alpha) = \infty) \Rightarrow u'' < 0 \quad \textit{on } (\alpha, \tau], \tag{3.15}$$

$$\alpha = -\infty \Rightarrow u'' < 0 \quad \textit{on } (\alpha, \tau]. \tag{3.16}$$

Proof. Theorem 2.2 is applicable to the differential equation (3.3) on $-\infty < s < \infty$ and gives the existence of $\tau \in (\alpha, \infty)$ such that $U = q^{1/2}u$ satisfies $U > 0$ on (α, ∞), $U' < 0$ on (α, τ) and $U' > 0$ on (τ, ∞). This gives (3.13); cf. (3.2). The first implication in (3.14) follows from Proposition 2.1. The other implications in (3.14), (3.15) are clear. Finally, (3.16) follows from $u' > 0$ and (3.13), (3.15).

4. On the modified Bessel functions. The differential equation

$$x'' - (1 + \beta/t^2)x = 0, \quad \text{where } \beta = \nu^2 - \tfrac{1}{4}, \tag{4.1}$$

has the positive solutions $x_0(t) = t^{1/2}K_\nu(t)$, $y(t) = t^{1/2}I_\nu(t)$ for $t > 0$. These solutions satisfy (1.2), (1.3) for $t > 0$ if $\nu \geqq \frac{1}{2}$, and for large t if $0 \leqq \nu < \frac{1}{2}$. Corresponding to (4.1), we have

$$q = 1 + \beta/t^2 \quad \text{for } t > 0 \quad \text{and} \quad Q = 1 + \beta(6t^2 + \beta)/4(t^2 + \beta)^3, \tag{4.2}$$

$$Q' = 3\beta t(\beta - 4t^2)/2(t^2 + \beta)^4, \tag{4.3}$$

where $q > 0$ and Q is defined for $t > 0$ if $\beta \geqq 0$ ($\nu \geqq \frac{1}{2}$) and for $t > (-\beta)^{1/2}$ if $\beta < 0$ ($0 \leqq \nu < \frac{1}{2}$). For $t > 0$, let

$$u = u_\nu(t) = tI_\nu(t)K_\nu(t). \tag{4.4}$$

THEOREM 4.1. *Let* $\nu > \frac{1}{2}$ (*i.e.*, $\beta > 0$). *Then*

$$u > 0, \quad u' > 0, \quad u'' < 0 \quad \textit{for } t > 0, \tag{4.5}$$

$$u''' < 0 \quad \textit{for } 0 < t < \tau, \quad u''' > 0 \quad \textit{for } t > \tau \tag{4.6}$$

for some $\tau = \tau_\nu > 0$, *where* $u''' = 4q^{1/2}(q^{1/2}u)'$.

Proof. In view of (4.2) and (4.3), Theorem 3.2 is applicable with $\alpha = 0$ and $t_0 = \beta^{1/2}/2$. Standard power series expansions (cf. [10, pp. 77–78]) show that if ν is

not an integer, then

$$u''<0 \quad \text{for small } t>0; \tag{4.7}$$

cf. Proposition 4.1 below. Hence $u''<0$ follows from (3.14) and (3.15) if ν is not an integer. If ν is an integer, we use (3.14), (3.15) and continuity considerations.

THEOREM 4.2. *Let* $0\leqq\nu<\frac{1}{2}$ (*i.e.*, $0>\beta\geqq-\frac{1}{4}$). *Then there exist numbers* τ_1, τ_2 *such that* $0<(-\beta)^{1/2}<\tau_1<\tau_2$,

$$u>0 \quad \text{for } t>0, \quad u'>0 \quad \text{on } (0,\tau_1) \quad \text{and} \quad u'<0 \quad \text{on } (\tau_1,\infty), \tag{4.8}$$

$$u''<0 \quad \text{on } (0,\tau_2) \quad \text{and} \quad u''>0 \quad \text{on } (\tau_2,\infty). \tag{4.9}$$

Proof. Since $u(0)=0$, it follows from Proposition 2.3(iii) where $t_+=(-\beta)^{1/2}>0$, that there exists $\tau_1>0$ satisfying (4.8). Also, Proposition 2.3(iii) applied to the equation (3.3) on the s-interval corresponding to $t_+<t<\infty$ gives the existence of $t_3\in(t_+,\infty)$ such that $q^{1/2}u>0$ on (t_+,∞), $(q^{1/2}u)'>0$ on (t_+,τ_3) and $(q^{1/2}u)'<0$ on (τ_3,∞). It is clear from $u(\infty)<\infty$ that $u'(\infty)=0$ and that there exists $\tau_2\in(\tau_1,\tau_3)$ such that $u''>0$ on $t>\tau_2$, $u''(\tau_2)=0$.

If it is verified that $\tau_1>(-\beta)^{1/2}$, then it follows that $u''<0$ on $((-\beta)^{1/2},\tau_2)$ since $u'''>0$ on this interval. Hence Theorem 4.2 is a consequence of the following:

PROPOSITION 4.1. *Let* $\nu>-\frac{1}{2}$. *Then* $u'>0$ *for* $0<t\leqq[2(1+5^{1/2})(8\nu^2+1)]^{1/2}$ *and* $u''<0$ *for* $0<t\leqq4(8\nu^2+1)^{1/2}$, *i.e.*, $\tau_1>[2(1+5^{1/2})(8\nu^2+1)]^{1/2}\geqq[2(1+5^{1/2})]^{1/2}$ *and* $\tau_2>4(8\nu^2+1)^{1/2}\geqq4$.

In the proof of this proposition, we shall use

LEMMA 4.1. *Let* $R(r)$ *be continuous on* $0<r<\omega$ *and* $v(r)$ *a solution of*

$$d^2v/dr^2+R(r)v=0 \tag{4.10}$$

which is positive for small $r>0$. *Let* $0<h(r)\in C^2(0,\omega)$ *be such that* $h(r)v(r)\to0$ *as* $r\to0$, *and*

$$R_1(r)\equiv h^{-4}(r)\{R(r)-h(r)[1/h(r)]''\} \quad \text{is nondecreasing} \tag{4.11}$$

on an r-*interval* (r_0,ω) *while* $R_1(r)\leqq0$ *on* $(0,r_0)$, $0\leqq r_0\leqq\omega$. *Then*

$$\infty\geqq\int_0^X h^3(r)v(r)\,dr\geqq0 \quad \text{for } 0<X<\omega. \tag{4.12}$$

If, in addition, v *has at least two positive zeros and* $r_1<r_2$ *are the smallest, then*

$$\infty\geqq\int_0^X h^3v\,dr\geqq\int_0^{r_2} h^3v\,dr\geqq0 \quad \text{for } r_2<X<\omega. \tag{4.13}$$

The inequality "$\geqq0$" *in* (4.12) [*or* (4.13)] *can be replaced by* ">0" *if* $R_1(r)$ *is not a constant on* $0<r<X$ [*or* $0<r<r_2$].

Proof of Lemma 4.1. The change of variables $w=h(r)v$ and $ds=h^2(r)\,dr$ reduces (4.10) to

$$d^2w/ds^2+R_1(r)w=0, \quad \text{where } r=r(s).$$

If the interval $0<r<\omega$ goes over into $(-\infty \leqq)\alpha_1<s<\omega_1$, then the second Sturm comparison theorem and the alternating series argument, as applied in Hartman and Wintner [6] (cf. [4, Exercise 3.5, p. 513 and p. 578]), imply that

$$\infty \geqq \int_0^X h^3 v\,dr = \int_{\alpha_1}^S hv\,ds = \int_{\alpha_1}^S w\,ds \geqq 0,$$

where $S=s(X)$. The arguments in [6] also give the last part of Lemma 4.1.

COROLLARY 4.1. *Let* $0<h(r)\in C^2(0,\infty)$ *satisfy* $h(r)r^{\mu+1/2}\to 0$ *as* $r\to 0$, *and*

$$R_1(r)=h^{-4}(r)\{1-\gamma/r^2-h(r)[1/h(r)]''\},\qquad \gamma=\mu^2-\tfrac{1}{4}, \tag{4.14}$$

satisfy the conditions of Lemma 4.1 *with* $\omega=\infty$. *Then*

$$\infty \geqq \int_0^X h^3(r)r^{1/2}J_\mu(r)\,dr>0 \quad \textit{for } X>0. \tag{4.15}$$

The integral exceeds a positive constant if $X\geqq$ const. >0.

This follows from the fact that $v=r^{1/2}J_\mu(r)$ is a solution of (4.10) with $R(r)=1-\gamma/r^2$.

Proof of Proposition 4.1. This proof will depend on the formula

$$2u(t/2)=t\int_0^\infty (r^2+t^2)^{-1/2}J_{2\nu}(r)\,dr \quad \text{for } \nu>-\tfrac{1}{2},\quad t>0;$$

[10, p. 435]. It is readily verified that formal differentiations are valid. Two differentiations give

$$u'(t/2)=\int_0^\infty (r^2+t^2)^{-3/2}r^2J_{2\nu}(r)\,dr,$$

$$u''(t/2)=-6t\int_0^\infty (r^2+t^2)^{-5/2}r^2J_{2\nu}(r)\,dr.$$

Thus in view of Corollary 4.1, it is sufficient to show that (4.14) satisfies the conditions of Lemma 4.1 when $\mu=2\nu$ and $h(r)$, for fixed t, is either of the functions

$$h(r)=r^{1/2}(r^2+t^2)^{-1/2} \quad \text{and} \quad 0<t^2\leqq(1+5^{1/2})(8\nu^2+1)/2, \tag{4.16}$$

$$h(r)=r^{1/2}(r^2+t^2)^{-5/6} \quad \text{and} \quad 0<t^2\leqq 4(8\nu^2+1). \tag{4.17}$$

In the case of (4.16), a straightforward calculation gives

$$r^5R_1'(r)=2r^6-2t^4r^2+(4\gamma+3)t^4+(4\gamma+3)t^2r^2.$$

The coefficient of t^4 is $4\gamma+3-2r^2$, so that $R_1'>0$ for $r^2\leqq(4\gamma+3)/2$. Also

$$r^5R_1'=2r^2\{r^4+(4\gamma+3)^2/16-[t^2-(4\gamma+3)/4]^2\}+(4\gamma+3)t^4,$$

so that $R_1'>0$ if $r^2>(4\gamma+3)/2$ and

$$t^2\leqq(1+5^{1/2})(4\gamma+3)/4=(1+5^{1/2})(8\nu^2+1)/2,$$

since $\gamma = \mu^2 - \frac{1}{4}$, $\mu = 2\nu$.

In the case (4.17), we have $R_1(r) = (r^2+t^2)^{4/3} R_2(r)$,

$$R_2(r) = t^4/2 - (4\gamma+3)t^4/4r^4 - (4\gamma+3)t^2/2r^2 + r^2 + 2t^2 - \gamma - \tfrac{7}{36},$$

$$r_1^5 R_2' = -\lambda t^4 + (4\gamma+3)r^2t^2 + 2r^6 \quad \text{if } \lambda = 2r^2 - (4\gamma+3).$$

If $\lambda \leqq 0$, i.e., $r^2 \leqq (4\gamma+3)/2$, then $R_2' > 0$. If $\lambda > 0$, we write

$$r_1^5 R_2' = \lambda(2r^4/\lambda - t^2)(t^2+r^2).$$

Hence $R_2' \geqq 0$ if $\lambda > 0$ and

$$t^2 \leqq 4r^4/2\lambda = 4\gamma + 3 + [\lambda + (4\gamma+3)^2/\lambda]/2,$$

where $2r^2 = \lambda + 4\gamma + 3$. But

$$\lambda + (4\gamma+3)^2/\lambda \geqq 2(4\gamma+3) \quad \text{for } \lambda > 0,$$

so that $R_2' \geqq 0$ if $\lambda > 0$ and $t^2 \leqq 2(4\gamma+3) = 4(8\nu^2+1)$. This completes the proof.

5. On Whittaker's differential equation. This equation is

$$x'' - (\tfrac{1}{4} - \kappa/t + \beta/t^2)x = 0, \quad \text{where } \beta = \nu^2 - \tfrac{1}{4}, \tag{5.1}$$

and κ, ν are real constants; cf. [9, pp. 88–91]. This is the special case of (2.1) with

$$q(t) = \tfrac{1}{4} - \kappa/t + \beta/t^2 = (t^2 - 4\kappa t + 4\beta)/4t^2 \equiv P(t)/4t^2. \tag{5.2}$$

We shall assume that

$$\kappa \neq 0, \tag{5.3}$$

for otherwise the change of variables $t \to 2t$ reduces (5.1) to (4.1). We shall also assume κ, ν are chosen so that

$$\text{(5.1) is disconjugate on } t > 0. \tag{5.4}$$

Note that

$$q(t) = 0 \Leftrightarrow t = t_\pm = 2\kappa \pm 2(\kappa^2 - \beta)^{1/2}. \tag{5.5}$$

Thus it follows that

$$q > 0 \text{ for } t > 0 \Leftrightarrow \text{either } \beta > \kappa^2 \text{ or } \beta \geqq 0,\ \kappa \leqq 0. \tag{5.6}$$

In particular, these conditions are sufficient for (5.4). They are not necessary; e.g., (5.4) holds if $\kappa < \frac{1}{2}$ (cf. [8, p. 89]). See also [8, pp. 182–183].

In the case of (5.2), the function (3.4) is

$$Q = 1 - 4[2\kappa t^3 - 3(\kappa^2 + 2\beta)t^2 + 12\beta\kappa t - 4\beta^2]/P^3(t). \tag{5.7}$$

Simple calculations give

$$q' = (\kappa t - 2\beta)/t^3, \tag{5.8}$$

$$Q' = 24tH(t)/P^4(t), \tag{5.9}$$

$$H = \kappa t^3 - 2(\kappa^2 + 2\beta)t^2 + 2\kappa(\kappa^2 + 5\beta)t + 4\beta(\beta - 3\kappa^2), \tag{5.10}$$

$$H' = 3\kappa t^2 - 4(\kappa^2 + 2\beta)t + 2\kappa(\kappa^2 + 5\beta), \tag{5.11}$$

(5.12) $H'(t)=0 \Leftrightarrow t=T_{\pm}=\{2(\kappa^2+2\beta)\pm[2(8\beta+\kappa^2)(\beta-\kappa^2)]^{1/2}\}/3\kappa.$

Thus if $\kappa \neq 0$, then

(5.13) $H' \neq 0$ on $(-\infty, \infty) \Leftrightarrow -\kappa^2/8 < \beta < \kappa^2.$

But if $\beta \geqq 0$ and $\beta - \kappa^2 \geqq 0$, then

(5.14) $T_{\pm} < 0$ and $H' < 0$ for $t > 0$ if $\kappa < 0$,

(5.15) $T_{\pm} > 0$ if $\kappa > 0$.

When (5.4) holds, the standard solutions $x = N_{\kappa\nu}(t)$ and $x = W_{\kappa\nu}(t)$ of (5.1) are positive principal solutions at $t=0$ and $t=\infty$; cf. [9, pp. 88–91]. We shall consider monotony and convexity properties of

(5.16) $u = u_{\kappa\nu}(t) = N_{\kappa\nu}(t) W_{\kappa\nu}(t).$

THEOREM 5.1. *Let $\kappa < 0$ and $\beta \geqq 0$ (i.e., $\nu \geqq \frac{1}{2}$). Then $q > 0$, $q' < 0$ and $Q > 0$ for $t > 0$.* (i) *If $0 \leqq \beta \leqq 3\kappa^2$, then $Q' < 0$ for $t > 0$ and*

(5.17) $u > 0, \quad u' > 0, \quad u'' < 0, \quad u''' > 0 \quad$ *for* $t > 0.$

(ii) *If $\beta > 3\kappa^2$, then there exist positive t_0, τ such that $Q' > 0$ on $(0, t_0)$ and $Q' < 0$ on (t_0, ∞), and*

(5.18) $u > 0, \quad u' > 0, \quad u'' < 0 \quad$ *for* $t > 0,$

(5.19) $u''' < 0 \quad$ *on* $(0, \tau), \quad u''' > 0 \quad$ *on* $(\tau, \infty).$

Proof. Let $\kappa < 0$ and $\beta \geqq 0$. Then the inequalities $q > 0$, $q' < 0$, $Q > 0$ for $t > 0$ follow trivially.

(i) If $0 \leqq \beta < \kappa^2$, then $H' < 0$ for all t, by (5.13) and (5.10), and if $\kappa^2 \leqq \beta \leqq 3\kappa^2$, then $H' < 0$ for $t > 0$ by (5.14). Since $H(0) = 4\beta(\beta - 3\kappa^2) \leqq 0$, it follows that $H < 0$ for $t > 0$, and so $Q' < 0$ for $t > 0$ by (5.9). Thus (5.17) follows from Theorem 3.1.

(ii) If $\kappa < 0$ and $\beta > 3\kappa^2$; then $H' < 0$ for $t > 0$ and $H(0) > 0$, and so the existence of t_0 is clear from (5.9). Hence (5.18), (5.19) follow from Theorem 3.2 (in particular, from (3.14) and (3.16); cf. [9, p. 88]).

THEOREM 5.2. *Let $\kappa > 0$ and let* (5.4) *hold. Then there exists a $\tau > 0$ such that $u > 0$ on $(0, \infty)$, $u' > 0$ on $(0, \tau)$ and $u' < 0$ on $t > \tau$.*

Proof. Case 1. $\beta \geqq \kappa^2$. In this case, $q(t) > 0$ for $t > 0$, except that $q(2\kappa) = 0$ if $\beta = \kappa$, $q' < 0$ on $(0, 2\beta/\kappa)$ and $q' > 0$ on $(2\beta/\kappa, \infty)$, by (5.6) and (5.8). The existence of τ follows from Theorem 2.2 and the Remark following it.

Case 2. $0 < \beta < \kappa^2$. In this case, $q < 0$ if and only if $(0 <) t_- < t < t_+$, $q' < 0$ on $(0, 2\beta/\kappa)$ and $q' > 0$ on $(2\beta/\kappa, \infty)$. The existence of τ follows from Theorem 2.3.

Case 3. $\beta \leqq 0$. In this case with $t > 0$, $q < 0$ if and only if $0 < t < t_+$, while $q' > 0$ on $(0, \infty)$. The assertion follows from Proposition 2.3(iii) in the case (2.12), since $u(+0) = 0$. This completes the proof.

Remark. In Theorem 5.2, let H (hence Q') be positive for $t > t_+ > 0$. [This is the case, e.g., if $-\kappa^2/8 \leqq \beta \leqq \kappa^2$ (so that $H' \geqq 0$ for all real t) and if $H(t_+) \geqq 0$.] Then Proposition 2.3(iii) applied to the differential equation (3.3) on the s-interval corresponding to $t > t_+ > 0$ implies the existence of $\tau_1 \in (t_+, \infty) \cap (\tau, \infty)$ such that $u''' > 0$ on (t_+, τ_1) and $u''' < 0$ on (τ_1, ∞). Here we have used the analogue

of (2.12), i.e., $U=q^{1/2}u\to 0$ as $t\to t_+$. We illustrate this in the case $\beta=\kappa^2>0$, so that $t_+=2\kappa$,

$$q=(t-2\kappa)^2/4t^2, \qquad q'=\kappa(t-2\kappa)/t^3,$$
$$Q=1-4\kappa(2t-\kappa)/(t-2\kappa)^4, \qquad Q'=24\kappa t/(t-2\kappa)^5.$$

THEOREM 5.3. *Let* $-\frac{1}{4}\leqq\beta<0$ (*i.e.*, $0\leqq\nu<\frac{1}{2}$) *and* $\kappa<0$. *Then there exist* τ_1 *and* τ_2 *such that* $0<\tau_1<\tau_2$, $u'>0$ *on* $(0,\tau_1)$, $u'<0$ *on* (τ_1,τ_2) *and* $u'>0$ *on* (τ_2,∞).

Proof. Note that $t_\pm$ in (5.5) satisfies $t_-<0<t_+$, so that $q<0$ on $(0,t_+)$ and, by (5.8), $q'>0$ on $(0,t_0)$, $q'<0$ on (t_0,∞) with $t_0=2\beta/\kappa>t_+>0$. Proposition 2.5 implies therefore that either $u'>0$ on $(0,\infty)$ or that τ_1, τ_2 exist, as asserted. Let β be fixed, $-\frac{1}{4}\leqq\beta<0$. The proof will depend on the use of continuity arguments for varying $\kappa\leqq 0$.

If $\kappa=0$, there exists a $\tau_0>0$ such that $u'>0$ on $(0,\tau_0)$ and $u'<0$ on (τ_0,∞). This assertion is contained in Theorem 4.2 (after the change of variable $t\to 2t$). In particular, $u'(\tau_0+1)<0$. By continuity considerations, $u'(\tau_0+1)<0$, for small $-\kappa>0$. Hence τ_1, τ_2 exist for small $-\kappa>0$.

This argument makes it clear that the set of $\kappa<0$ for which τ_1, τ_2 exist is open. It will be shown that it is also closed on $\kappa<0$. Since $q(t)$ is strictly monotone on $(0,2\beta/\kappa)$ and on $(2\beta/\kappa,\infty)$, the proof of Proposition 2.5 shows that $u''(s)\neq 0$ if $u'(s)=0$. Thus by the implicit function theorem, $\tau_1(\kappa)$ and $\tau_2(\kappa)$ are continuous on the κ-set (in $\kappa<0$) on which they exist.

Suppose τ_1 and τ_2 exist on $0>\kappa>\kappa_0$. The behavior of the factors $N_{\kappa\nu}(t)$, $W_{\kappa\nu}(t)$ of u near $t=0$ show that there is a constant $c(\kappa_0)$ such that $0<c(\kappa_0)\leqq\tau_1(\kappa)<\tau_2(\kappa)$ for small $\kappa-\kappa_0>0$.

In the proof of Proposition 2.5, it is clear that $t_2=t_2(\kappa)$, $t_3=t_3(\kappa)$ exist for $0>\kappa>\kappa_0$ and that $t_+(\kappa)\leqq t_2(\kappa)<t_0<t_3(\kappa)$. By continuity considerations, $t_2(\kappa_0)$, $t_3(\kappa_0)$ exist and $t_+(\kappa_0)\leqq t_2(\kappa_0)\leqq t_0(\kappa_0)\leqq t_3(\kappa_0)$. Since $\tau_2(\kappa)<t_3(\kappa)$ by the proof of Proposition 2.5, we have that $t_2(\kappa)<t_3(\kappa)\leqq t_3(\kappa_0)+1$ for small $\kappa-\kappa_0>0$. Thus there exists a sequence $\kappa_1,\kappa_2,\cdots$ such that $0<\kappa_n<\kappa_0$, $\kappa_0=\lim\kappa_n$ and $\tau_{10}=\lim\tau_1(\kappa_n)$, $\tau_{20}=\lim\tau_2(\kappa_n)$ exist as $n\to\infty$. We have $c(\kappa_0)\leqq\tau_{10}\leqq\tau_{20}\leqq t_3(\kappa_0)+1$. The case $\tau_{10}=\tau_{20}$ is impossible for otherwise $u'\geqq 0$ on $t>0$ but $u'(\tau_{10})=0$ when $\kappa=\kappa_0$. Thus τ_1, τ_2 exist for $\kappa=\kappa_0$ (and $\tau_1(\kappa_0)=\tau_{10}$, $\tau_2(\kappa_0)=\tau_{20}$). This completes the proof.

REFERENCES

[1] P. APPELL, *Sur les transformations des équations différentielles lineaires*, C.R. Acad. Sci. Paris, 91 (1880), pp. 211–214.

[2] P. HARTMAN, *Unrestricted solution fields of almost separable differential equations*, Trans. Amer. Math. Soc., 63 (1948), pp. 560–580.

[3] ———, *On differential equations and the function* $J_\mu^2+Y_\mu^2$, Amer. J. Math., 83 (1961), pp. 154–188.

[4] ———, *Ordinary Differential Equations*, S. M. Hartman, Baltimore, 1973.

[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] P. HARTMAN AND A. WINTNER, *On non-conservative linear oscillators of low frequency*, Amer. J. Math., 70 (1948), pp. 529–539.

[7] ———, *On non-oscillatory linear differential equations with monotone coefficients*, Ibid., 76 (1954), pp. 207–219.

[8] H. BUCHHOLZ, *The Confluent Hypergeometric Function*, Springer Tracts in Natural Philosophy, vol. 15, Springer-Verlag, Berlin, 1969.

[9] W. MAGNUS AND C. OBERHETTINGER, *Formulas and Theorems for the Special Functions of Mathematical Physics*, Chelsea, New York, 1949.

[10] G. N. WATSON, *A Treatise on the Theory of Bessel Functions*, 2nd ed., Cambridge University Press, Cambridge, 1958.

SIAM J. MATH. ANAL.
Vol. 8, No. 4, August 1977

FILTER STABILITY IN DYNAMICAL SYSTEMS*

JOSEPH AUSLANDER†

Abstract. Lyapunov stability of a closed (in general noncompact) set in a dynamical system is studied by means of a pair of neighborhood filters of the set. In this way, several distinct stability notions arise (which coincide if the set is compact). Stability of the trivial solution of nonautonomous systems of differential equations is included as a special case. Necessary and sufficient conditions for filter stability are given in terms of families of Lyapunov functions.

Introduction. In this paper we develop the elements of a theory of stability for closed noncompact invariant sets in locally compact dynamical systems. The results are, in some respects, analogous to those for compact invariant sets [4]. There are, however, some substantial differences. In particular, there is no obvious way to define Lyapunov stability of a noncompact set. The "neighborhood" and "$\varepsilon-\delta$" definitions, which obviously coincide for compact sets, are different in the noncompact case. Rather than choose one of these as our definition, we define "filter stability" which includes both as special cases. For each pair of neighborhood filters $\mathscr{E}$ and $\mathscr{F}$ of the closed invariant set M, we have a notion of "$(\mathscr{E}, \mathscr{F})$ stability."

In the study of stability of compact invariant sets, certain set functions, "prolongations," and real valued functions, "Lyapunov functions" have proved to be useful. We introduce "$\mathscr{F}$ prolongations" and "$(\mathscr{E}, \mathscr{F})$ Lyapunov functions" which play similar roles. In general, the existence of an $(\mathscr{E}, \mathscr{F})$ Lyapunov function does not imply stability. Thus we are led to the notion of an "$(\mathscr{E}, \mathscr{F})$ Lyapunov family"; the existence of such a family is equivalent to $(\mathscr{E}, \mathscr{F})$ stability. We also briefly consider filter attraction and asymptotic stability.

Other discussions of stability of noncompact sets occur in [3], [4], and [8].

1. Definitions. Let X be a locally compact metric space, with metric d. If $M \subset X$, and $r > 0$, we write $S(M, r)$ for the set of points whose distance from M is less than r, int M and ∂M for the interior and boundary of M respectively, and $\mathscr{N}_M$ (or just $\mathscr{N}$) for the neighborhood system of M.

By a *dynamical system* or *flow* on X, we mean a continuous action of the additive group of real numbers $\mathbb{R}$ on X. If $x \in X$, $t \in \mathbb{R}$, we denote the action of t on x by xt. If $x \in X$, the *orbit*, *positive semi-orbit*, and *omega limit set* of x, are, respectively, the set $\gamma(x) = [xt/t \in R]$, $\gamma^+(x) = [xt/t \geqq 0]$, and $\omega(x) = \bigcap_{t \geqq 0} \gamma^+(xt)$. A subset M of X is said to be *invariant* (*positively invariant*) if, whenever $x \in X$, $\gamma(x) \subset M$ ($\gamma^+(x) \subset M$).

At this point, we introduce the idea of filter stability. Recall that a filter on a set X is a nonempty collection $\mathscr{F}$ of nonempty subsets of X, which is closed under finite intersections, and which contains every superset of each of its members. Suppose, as above, a dynamical system is given on the locally compact metric space X, and let M be a closed invariant subset of X. A *neighborhood filter* of M is a filter $\mathscr{F} \subset \mathscr{N}_M$. Note that $\mathscr{N}_M$ is itself a neighborhood filter, as is $\mathscr{U}_M = \mathscr{U}$, the

* Received by the editors April 23, 1975.

† Department of Mathematics, University of Maryland, College Park, Maryland 20742.

collection of metric neighborhoods of M ($U \in \mathcal{U}$ if and only if $U \supset S(M, \varepsilon)$, for some $\varepsilon > 0$).

If $\mathcal{F}$ is a neighborhood filter of M, we denote by $\mathcal{F}^+$ the positively invariant members of $\mathcal{F}$; $\mathcal{F}^+ = [\gamma^+(F)/F \in \mathcal{F}]$. The collection $\mathcal{F}^+$ is a filter base (the intersection of any two members contains a third) and we write $\{\mathcal{F}^+\}$ for the filter generated by $\mathcal{F}^+$. Thus $\{\mathcal{F}^+\}$ consists of the supersets of positively invariant members of $\mathcal{F}$.

Now suppose that $\mathcal{E}$ and $\mathcal{F}$ are two neighborhood filters of M. Then we say that M is $(\mathcal{E}, \mathcal{F})$ *stable* provided that $\mathcal{F} \subset \{\mathcal{E}^+\}$ (so certainly $\mathcal{F} \subset \mathcal{E}$). Then M is $(\mathcal{E}, \mathcal{F})$ stable if and only if, whenever $F \in \mathcal{F}$, there is an $E \in \mathcal{E}$ such that $\gamma^+(E) \subset F$. We say that M is $\mathcal{F}$ *stable* if it is $(\mathcal{F}, \mathcal{F})$ stable; clearly this is the case if and only if $\{\mathcal{F}^+\} = \mathcal{F}$. Examples of this are *topological* stability ($\mathcal{F} = \mathcal{N}_M$) and *metric* (or "$\varepsilon - \delta$") stability ($\mathcal{F} = \mathcal{U}_M$).

If M is $(\mathcal{E}, \mathcal{F})$ stable, and $\mathcal{E}'$ and $\mathcal{F}'$ are neighborhood filters with $\mathcal{E} \subset \mathcal{E}'$ and $\mathcal{F}' \subset \mathcal{F}$, obviously M is $(\mathcal{E}', \mathcal{F}')$ stable. However, there is, in general, no relation between $\mathcal{E}$ and $\mathcal{F}$ stability, even when one of the filters is contained in the other. For example, an invariant line in a parallel flow in the plane is metrically stable but not topologically stable. On the other hand, consider a dynamical system on the set $[x/x \geqq 1]$ with singular points at all $x \in \cup_{n=1,2,\ldots} [n, n + 1/2^n]$ and such that, if $n + 1/2^n < x < n + 1$, then $xt \to n + 1$ as $t \to \infty$. Let $M = \{1, 2, 3, \cdots\}$. Then M is topologically stable (for example, by Theorem 3); however, for no $\delta > 0$ is it the case that $\gamma^+(S(M, \delta)) \subset S(M, \frac{1}{4})$, so M is not metrically stable.

Stability notions of the trivial solution of nonautonomous systems of differential equations can be formulated as special cases of $(\mathcal{E}, \mathcal{F})$ stability. For, consider the n-dimensional system (NA): $\dot{x} = f(t, x)$, $f(t, 0) = 0$, $(t \in \mathbb{R})$; we assume that f is sufficiently smooth on $\mathbb{R} \times \mathbb{R}^n$ so that, given any $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}^n$, there exists a unique solution $\phi(t, t_0, x_0)$ defined for all $t \in \mathbb{R}$, which depends continuously upon (t_0, x_0) and which equals x_0 for $t = t_0$. The trivial solution $x = 0$ is *stable* if, given any $\varepsilon > 0$ and any $t_0 \in \mathbb{R}$, there exists $\delta = \delta(\varepsilon, t_0) > 0$ such that $\|x_0\| < \delta$ implies $\|\phi(t, t_0, x_0)\| < \varepsilon$ for $t \geqq t_0$, and is *uniformly stable* if δ depends only on ε, [1]. Now, by a familiar device, we may form the associated $(n+1)$-dimensional autonomous system (A): $\dot{y} = g(y)$, where $y = (t, x)$, $g = (1, f)$. The solutions of (A) define a dynamical system in $\mathbb{R}^{n+1}$ for which $M = [(t, 0)/t \in \mathbb{R}]$ is a closed invariant set. It follows easily that the solution $x = 0$ of (NA) is stable if and only if M is $(\mathcal{N}, \mathcal{U})$ stable, and is uniformly stable if and only if M is $\mathcal{U}$ stable. (Note that the set $[(t, x)/\|x\| < \delta(\varepsilon, t), t \in \mathbb{R}]$ is a neighborhood of M—this follows from continuous dependence on initial conditions.)

It is natural to inquire whether a closed invariant set M is $\mathcal{F}$ stable for some neighborhood filter $\mathcal{F}$. If the question is posed in this way, it has a trivial affirmative answer. In fact, if $\mathcal{F}$ is any neighborhood filter of M, then obviously M is $\{\mathcal{F}^+\}$ stable. For this reason, we impose an additional condition, "sufficiency" on those neighborhood filters which we consider: a neighborhood filter $\mathcal{F}$ will be called *sufficient* if its adherence $\cap [\bar{F}/F \in \mathcal{F}]$ is equal to M. Hence $\mathcal{F}$ is sufficient if and only if whenever $x \in X \setminus M$, there is an $F \in \mathcal{F}$ such that $x \notin \bar{F}$.

The consideration of $\mathcal{F}$ stability for various neighborhood filters $\mathcal{F}$ is only of interest when the closed invariant set M is not compact. The reason for this is that if M is compact, and $\mathcal{F}$ is a sufficient neighborhood filter which contains a compact

set then $\mathcal{F}$ coincides with $\mathcal{N}_M$, the complete neighborhood filter. Indeed, suppose M is compact, and let $\mathcal{F}$ be such a neighborhood filter. Let W be a compact member of $\mathcal{F}$ and let $U \in \mathcal{N}_M$, with $U \subset W$. If $z \in W \setminus U$, let $V_z \in \mathcal{F}$ with $z \notin \bar{V}_z$, and let N_z be an open neighborhood of z with $\bar{N}_z \cap \bar{V}_z = \varnothing$ and $N_z \cap M = \varnothing$. Now $W \setminus U$ is compact, so there are $z_1, \cdots, z_n \in W \setminus U$ such that $W - U \subset \bigcup_{j=1}^{n} N_{z_j}$. Let $V = \bigcap_{j=1}^{n} V_{z_j} \cap W$. Then $V \in \mathcal{F}$ and $V \subset U$. It follows that $U \in \mathcal{F}$, and therefore $\mathcal{F} = \mathcal{N}_M$.

We recall the definition of that *prolongation* of a point [9]. If $x \in X$, the (first) prolongation of x, $D_1(x) = \bigcap_{U \in \mathcal{N}_x} \overline{\gamma^+(U)}$. Note that $y \in D_1(x)$ if and only if there are sequences $\{x_n\}, \{y_n\}$ in X, with $x_n \to x$, $y_n \in \gamma^+(x_n)$, and $y_n \to y$. A compact invariant set M is Lyapunov stable if and only if $D_1(M) = M$ (where $D_1(M) = \bigcup_{x \in M} D_1(x)$) (see [9]).

We generalize this notion by defining the prolongation of a set relative to a neighborhood filter. If $A \subset X$, and $\mathcal{F}$ is a neighborhood filter of A, the $\mathcal{F}$-prolongation of A is the set $D_{\mathcal{F}}(A) = \bigcap_{F \in \mathcal{F}} \overline{\gamma^+(F)}$. If $\mathcal{F}$ and $\mathcal{G}$ are neighborhood filters of A with $\mathcal{F} \subset \mathcal{G}$, then obviously $D_{\mathcal{G}}(A) \subset D_{\mathcal{F}}(A)$.

2. Results.

LEMMA 1. *Let A be a closed subset of X. Then $D_{\mathcal{N}}(A) = \operatorname{cl}(\bigcup_{x \in A} D_1(x))$.*

Proof. Clearly the right side is contained in the left side. Now, let $y \in D_{\mathcal{N}}(A)$. We show there is a sequence $x_n \in A$, with $y_n \in D_1(x_n)$ and $y_n \to y$. If not, there is an $r > 0$ such that $S(y, r)$ is compact and such that $D_1(x) \cap S(y, r) = \varnothing$, for all $x \in A$. Then every $x \in A$ has a neighborhood U_x such that $\gamma^+(U_x) \cap S(y, r) = \varnothing$. Let $U = \bigcup_{x \in A} U_x$. U is a neighborhood of A, and $\gamma^+(U) \cap S(y, r) = \varnothing$, so $\overline{\gamma^+(U)} \cap S(u, r/2) = \varnothing$, and $y \in D_{\mathcal{N}}(A)$. This contradiction completes the proof.

THEOREM 1. *Let M be a closed invariant set in X. Then the following are equivalent*:

(i) $D_{\mathcal{N}}(M) = M$.

(ii) $D_{\mathcal{F}}(M) = M$, *for some neighborhood filter* $\mathcal{F}$.

(iii) $D_1(x) \subset M$, *for every* $x \in M$.

(iv) M *is* $\mathcal{G}$ *stable, for some sufficient neighborhood filter* $\mathcal{G}$.

(v) M *is* $(\mathcal{E}, \mathcal{F})$ *stable for some neighborhood filters* $\mathcal{E}$ *and* $\mathcal{F}$, *with* $\mathcal{F}$ *sufficient.*

Proof. Obviously (i) implies (ii). If (ii) holds, and $x \in M$, $D_1(x) \subset D_{\mathcal{F}}(M) = M$, so (iii) holds. If (iii) holds, $D_{\mathcal{N}}(M) = \operatorname{cl} \bigcup_{x \in M} D_1(x) = M$ (using Lemma 1), and (iii) is proved.

Now, (ii) obviously implies that $\{\mathcal{F}^+\}$ is sufficient. Since, as we have observed, M is always $\{\mathcal{F}^+\}$ stable, (iv) holds with $\mathcal{G} = \{\mathcal{F}^+\}$. Finally, suppose M is $\mathcal{G}$ stable, where $\mathcal{G}$ is sufficient. Then if $y \notin M$, there is a $G \in \mathcal{G}$ such that $y \notin \bar{G}$. Let $G' \in \mathcal{G}$ such that $\gamma^+(G') \subset G$, so $y \notin \gamma^+(G')$. Hence $D_{\mathcal{G}}(M) = M$, so (iv) implies (ii). Obviously (iv) implies (v). If (v) holds, $\mathcal{F} \subset \{\mathcal{E}^+\}$, so $\{\mathcal{E}^+\}$ is sufficient, and (iv) holds with $\mathcal{G} = \{\mathcal{E}^+\}$. This completes the proof.

A closed invariant set M satisfying any one (hence all) of the conditions in Theorem 1 will be called *conditionally stable.* In this case, let

$$\Phi_0 = \left[\mathcal{F} \Big/ \begin{array}{l} \mathcal{F} \text{ is a sufficient neighborhood} \\ \text{filter and } M \text{ is } \mathcal{F} \text{ stable.} \end{array} \right]$$

Let $\mathcal{F}^* = \vee[\mathcal{F}/\mathcal{F} \in \Phi_0]$, the filter generated by finite intersection of elements of

$\cup[\mathcal{F}/\mathcal{F} \in \Phi_0]$. Then M is $\mathcal{F}^*$ stable. For, let $W_0 \in \mathcal{F}^*$. Then $W_0 \supset W_1 \cap \cdots \cap W_n$ ($W_i \in \mathcal{F}_i \in \Phi_0$), and there are $U_i \subset W_i$ ($i = 1, 2, \cdots, n$) such that $U_i \in \mathcal{F}$ and $\gamma^+(U_i) \subset W_i$, so $\gamma^+(U_1 \cap \cdots \cap U_n) \subset W_1 \cap \cdots \cap W_n \subset W_0$. We have proved:

THEOREM 2. *If M is conditionally stable, there is a unique maximal sufficient neighborhood filter $\mathcal{F}^*$ for which M is $\mathcal{F}^*$ stable.*

An example of conditional stability is *equi-stability*, as defined by Bhatia and Szegö ([7, p. 84])—M is equi-stable if, for any $x \in M$, there is a $\delta = \delta_x > 0$ such that $x \notin \gamma^+(S(M, \delta))$. This is just the assertion that $\{\mathcal{U}^+\}$ is sufficient.

Our next result concerns topological stability, which, as we will show, is closely related to compactness of prolongations (see also [8]). First we require a lemma.

LEMMA 2. *Let $x \in X$. Then $D_1(x)$ is compact if and only if, whenever $\{x_n\}, \{t_n\}$ are sequences with $x_n \to x$, $t_n > 0$, then $\{x_n t_n\}$ has a convergent subsequence.*

Proof. If $D_1(x)$ is not compact, then there is a sequence $\{z_n\}$, with no convergent subsequence, and $z_n \in D_1(x)$. Then one easily obtains sequences $x_n \to x$, $t_n > 0$ such that $d(x_n t_n, z_n) \to 0$, so $\{x_n t_n\}$ has no convergent subsequence. Now suppose $D_1(x)$ is compact. Let K be a compact subset of X with $D_1(x) \subset \operatorname{int} K$. If $x_n \to x$, $t_n > 0$ and $\{x_n t_n\}$ has no convergent subsequence, then, for $n \geqq n_0$, $x_n \in \operatorname{int} K$, $x_n t_n \notin K$. Let $0 < \tau_n < t_n$ such that $x_n \tau_n \in \partial K$, and a (subsequence of) $x_n \tau_n \to x^* \in \partial K$, so $x^* \in D_1(x)$ and $x^* \notin K$, a contradiction.

In our next theorem we assume that the set M has no interior. This is not an essential restriction, for if $\operatorname{int} M \neq \varnothing$, we may consider the stability of the (invariant) set $M^* = M \setminus \operatorname{int} M$ with respect to the subflow on $X^* = X \setminus \operatorname{int} M$. Obviously M^* is stable in X^* if and only if M is stable in X. It is also possible to consider the stability of M directly by defining a "relative" prolongation [2] but the statement of the theorem in this context is rather cumbersome.

THEOREM 3. *Let M be a closed invariant subset of X, with* $\operatorname{int} M = \varnothing$. *Then the following are equivalent*:

(i) *M is topologically stable.*

(ii) *If $x \in M$, $D_1(x)$ is compact subset of M.*

(iii) *$D_1(x)$ is compact, for every $x \in M$, and M is $\mathcal{F}$ stable, for some sufficient neighborhood filter $\mathcal{F}$.*

Proof. (i) $\Rightarrow$ (ii). Since $\mathcal{N}$ is sufficient, $D_1(x) \subset M$, for every $x \in M$, by Theorem 1. Suppose $D_1(x)$ is not compact, for some $x \in M$. Then there are sequences $x_n \to x$, $t_n > 0$ such that $\{x_n t_n\}$ has no convergent subsequence (Lemma 1). Since $M = \partial M$, we may suppose $x_n \notin M$. Let $W = X \setminus \cup_{n=1,2,\ldots} \{x_n t_n\}$. Then $W \in \mathcal{N}$, and, if $U \in \mathcal{N}$, $x_n \in U$, for $n \geqq n_0$, $x_n t_n \in \gamma^+(U) \setminus W$ so $\gamma^+(U) \not\subset W$. This contradicts topological stability.

(ii) $\Rightarrow$ (i). Let $W \in \mathcal{N}$, and let $W^* = [x \in W / \gamma^+(x) \cap \partial W = \varnothing]$. Now $\gamma^+(W^*) \subset W$, so, to prove topological stability, it is sufficient to show that $W^* \in \mathcal{N}$. If not, there is an $x \in M$ such that W^* is not a neighborhood of x. Then there is a sequence $z_n \to x$ such that $z_n \notin W^*$. Now $z_n \in W$, so there are $t_n > 0$ such that $z_n t_n \in \partial W$. Since $D_1(x)$ is compact, (a subsequence of) $z_n t_n \to z \in \partial W$. Thus $z \in D_1(x) \setminus M$, which is a contradiction.

Obviously (i) and (ii) imply (iii), and (iii) $\Rightarrow$ (ii), by Theorem 1.

We now turn to the role of real valued functions in the study of filter stability. In the compact case, the stability of M is equivalent to the existence of a

"Lyapunov" function for M [4]. If M is not compact, the existence of a single Lyapunov function is not sufficient for filter stability of M. This is because a neighborhood filter of a noncompact set does not, in general, have a countable base. In order to obtain a satisfactory analogue of the compact result, it is necessary to consider families of Lyapunov functions.

We now give the precise definitions. Let M be a closed invariant subset of X and let $\mathscr{E}$ be a neighborhood filter of M. A real valued function v defined on some positively invariant $E_0 \in \mathscr{E}$ will be called a *weak $\mathscr{E}$ Lyapunov function* for M provided:

(i) $v \geqq 0$ and $v^{-1}(0) = M$.

(ii) If $x \in E_0$, $t > 0$, then $v(xt) \leqq v(x)$.

(iii) If $\varepsilon > 0$, $v^{-1}([0, \varepsilon)) \in \mathscr{E}$.

Note that a weak $\mathscr{E}$ Lyapunov function need not be continuous.

A family of real valued functions Φ on X will be called an *$(\mathscr{E}, \mathscr{F})$ Lyapunov family* (for M), provided:

(i) Each $v \in \Phi$ is a weak $\mathscr{E}$ Lyapunov function.

(ii) If $F \in \mathscr{F}$, there is a $v \in \Phi$ and $\beta > 0$ such that $v^{-1}([0, \beta)) \subset F$ (equivalently—for every $F \in \mathscr{F}$, there is a $v \in \Phi$ such that $\inf [v(x)/x \in F] > 0$).

If $\{x_n\}$ is a net in X and $\mathscr{E}$ a neighborhood filter of M, we write $x_n \overset{\mathscr{E}}{\to} M$ provided $\{x_n\}$ is eventually in every $E \in \mathscr{E}$. If $x_n \overset{\mathscr{E}}{\to} M$ and v is a weak $\mathscr{E}$ Lyapunov family for M, then $v(x_n) \to 0$. In fact, if Φ is an $(\mathscr{E}, \mathscr{F})$ Lyapunov family for M, then $x_n \overset{\mathscr{E}}{\to} M$ implies that $v(x_n) \to 0$, for all $v \in \Phi$, which in turn implies that $x_n \overset{\mathscr{F}}{\to} M$. Therefore, if Φ is an $\mathscr{F}$ Lyapunov family for M (meaning an $(\mathscr{F}, \mathscr{F})$ Lyapunov family), and $\{x_n\}$ a net in X, it follows that $x_n \overset{\mathscr{F}}{\to} M$ if and only if $v(x_n) \to 0$, for all $v \in \Phi$.

THEOREM 4. *Let Φ be an $(\mathscr{E}, \mathscr{F})$ Lyapunov family for M. Then M is $(\mathscr{E}, \mathscr{F})$ stable.*

Proof. Let $F \in \mathscr{F}$ and let $v \in \Phi$, $\beta > 0$ such that $v^{-1}([0, \beta)) \subset F$. Then $E = v^{-1}([0, \beta)) \in \mathscr{E}$. If $x \in E$, $t > 0$, $v(xt) \leqq v(x) < \beta$, so $xt \in E$. Then $\gamma^+(E) \subset E \subset F$ and M is $(\mathscr{E}, \mathscr{F})$ stable.

Our next result establishes necessary and sufficient conditions for $(\mathscr{E}, \mathscr{F})$ stability in term of Lyapunov families.

THEOREM 5. *Suppose X is second countable and $\mathscr{E}$ and $\mathscr{F}$ are neighborhood filters of M, with $\mathscr{F}$ sufficient. Then M is $(\mathscr{E}, \mathscr{F})$ stable if and only if there is an $(\mathscr{E}, \mathscr{F})$ Lyapunov family for M.*

Proof. Sufficiency has been proved above (Theorem 4). We prove necessity. Suppose M is $(\mathscr{E}, \mathscr{F})$ stable, and let $F \in \mathscr{F}$. Let $x \in F - M$ and choose $E_x \in \mathscr{E}$ such that $x \in \overline{\gamma^+(E_x)}$ and $\gamma^+(E_x) \subset F$. (The existence of such an E_x is guaranteed by $(\mathscr{E}, \mathscr{F})$ stability and sufficiency of $\mathscr{F}$.) Let W_x be an open neighborhood of x such that $W_x \subset F$ and $W_x \cap \overline{\gamma^+(E_x)} = \varnothing$. Then $F - M = \bigcup_{x \in F - M} W_x$. By the Lindelöf covering theorem, there is a countable subfamily $\{W_{x_j}\}$ of $\{W_x\}$ such that $F - M = \bigcup_{j=1,2,\ldots} W_{x_j}$. Let $B_j \in \mathscr{E}$ be defined inductively by $B_1 = \gamma^+(E_{x_1}), \cdots, B_j = B_1 \cap \cdots \cap B_{j-1} \cap \overline{\gamma^+(E_{x_j})}$. Then $B_1 \supset B_2 \supset \cdots$ and $\bigcap_j \bar{B}_j = M$. (For, if $y \in F - M$, $y \in W_{x_i}$, some i, so $y \notin \gamma^+(E_{x_i})$, so $y \notin \bar{B}_i$.)

Now define v on X by $v(x) = \inf [1/k \mid \gamma^+(x) \in B_k]$. (If $\gamma^+(x)$ is contained in no B_k put $v(x) = 1$.) Clearly $v \geqq 0$, $v^{-1}(0) = M$, and $v(x, t) = v(x)$ $(x \in X, t > 0)$. Note that each B_k is positively invariant, so, if $x \in B_k$, $v(x) \leqq 1/k$. If $\varepsilon > 0$, let

$1/k < \varepsilon$, so $v^{-1}([0, \varepsilon)) \supset v^{-1}([0, 1/k)) \supset B_k \in \mathscr{E}$. Hence $v^{-1}([0, \varepsilon)) \in \mathscr{E}$, and v is a weak $\mathscr{E}$ Lyapunov function. If $x \in F$, $x \in B_k$ $(k = 1, 2, \cdots)$, so $v(x) = 1$. Given $F \in \mathscr{F}$, we have constructed a weak $\mathscr{E}$ Lyapunov function which is bounded away from 0 on the complement of F. This completes the proof.

Our final result on this topic shows that either a single Lyapunov function suffices or uncountably many are necessary.

THEOREM 6. *Suppose M is $\mathscr{F}$ stable. If $\mathscr{F}$ has a countable base, then there is an $\mathscr{F}$ Lyapunov function v such that the singleton $\{v\}$ is an $\mathscr{F}$ Lyapunov family. On the other hand, if $\mathscr{F}$ does not have a countable base, any $\mathscr{F}$ Lyapunov family must be uncountable.*

Proof. Suppose $\{F_1, F_2, \cdots\}$ is a countable base for $\mathscr{F}$. We may suppose that $F_1 \supset F_2 \supset \cdots$ and that each F_i is positively invariant. Define v by $v(x) = 1/k$ if $x \in F_k \backslash F_{k+1}$. It is clear that $\{v\}$ is an $\mathscr{F}$ Lyapunov family. If $\{v_1, v_2, \cdots\}$ is an $\mathscr{F}$ Lyapunov family, then the collection $\{v_i^{-1}([0, 1/k))/i, k = 1, 2, \cdots\}$ is a countable base for $\mathscr{F}$.

Now we turn to attracting and asymptotic properties. If $\mathscr{E}$ and $\mathscr{F}$ are neighborhood filters of M, we say that M is an $(\mathscr{E}, \mathscr{F})$ *attractor* provided there is an $E_0 \in \mathscr{E}^+$ such that, if $x \in E_0$, $F \in \mathscr{F}$, there is a $T = T_{x,F} > 0$ such that $xt \in F$, for $t \geqq T$. The maximal such $E_0 \in \mathscr{E}^+$ will be called the *domain of* $(\mathscr{E}, \mathscr{F})$ *attraction.*

If the T in the above definition depends only on F, M will be called a *uniform* $(\mathscr{E}, \mathscr{F})$ *attractor.* If M is both an $(\mathscr{E}, \mathscr{F})$ attractor and $(\mathscr{E}, \mathscr{F})$ stable, we say that it is $(\mathscr{E}, \mathscr{F})$ *asymptotically stable.*

Note that if M is an $(\mathscr{E}, \mathscr{F})$ attractor, with $\mathscr{F}$ sufficient, then, if $x \in E_0$, its omega limit set $\omega(x)$ is a (possibly empty) subset of M. Also, if M is a uniform $(\mathscr{E}, \mathscr{F})$ attractor, and if $E \in \mathscr{E}$ implies $Et \in \mathscr{E}$ $(t \geqq 0)$ then M is $(\mathscr{E}, \mathscr{F})$ asymptotically stable.

Consider again the system $\dot{x} = f(t, x)$, $f(t, 0)$ and the closed invariant set $M = [(t, 0)/t \geqq 0]$ in $\mathbb{R}^{n+1}$. A number of "asymptotic stability" notions may be phrased in these terms. We omit the definitions (see [1]), and just remark that the trivial solution $x = 0$ is:

(i) *quasi-asymptotically stable* if and only if M is an $(\mathscr{N}, \mathscr{U})$ attractor.

(ii) *asymptotically stable* if and only if M is $(\mathscr{N}, \mathscr{U})$ asymptotically stable.

(iii) *quasi-uniform-asymptotically stable* if M is a $\mathscr{U}$ attractor.

(iv) *uniform-asymptotically stable* if M is $\mathscr{U}$ asymptotically stable.

THEOREM 7. *Let M be $\mathscr{F}$ stable, and let Φ be an $\mathscr{F}$ Lyapunov family for M. Then M is $\mathscr{F}$ asymptotically stable if and only if there is an $F_0 \in \mathscr{F}^+$ such that $\lim_{t\to\infty} v(xt) = 0$ for $v \in \Phi$ and $x \in F_0$, and is uniformly asymptotically stable, if, for each $v \in \Phi$, $\lim_{t\to\infty} v(xt) = 0$ uniformly on F_0.*

Proof. This is a consequence of the equivalence of $x_n \xrightarrow{\mathscr{F}} M$ and $v(x_n) \to 0$ $(v \in \Phi)$ for nets $\{x_n\}$ in X.

THEOREM 8. *Let M be a closed invariant set. Then the following are equivalent*:

(i) *M is a topological attractor, (i.e., an $\mathscr{N}$ attractor).*

(ii) *There is a neighborhood W of M such that, for every $x \in W \setminus M$, $\overline{\gamma^+(x)}$ is compact, and $\omega(x) \subset M$.*

(iii) *M is an $\mathscr{F}$ attractor, for some sufficient neighborhood filter $\mathscr{F}$, and $\overline{\gamma^+(x)}$ is compact for all x in some neighborhood W of M.*

Proof. Part (ii) obviously implies (i), and using the remark in the previous paragraph, it is easy to see that (iii) implies (ii).

We prove (i) implies (ii). If (ii) fails, there is a sequence $\{x_j\}$ with $x_j \in M$, $x_j \to x \in M$ such that either (α) $\gamma^+(x_j)$ is not compact, or (β) there is an $x_j' \in \omega(x_j) \setminus M$. In either case, we may suppose $x_j \in W$. In case (α), there is, for each j, a sequence $\{t_{n_j}\}$ such that $\{x_j t_{n_j}\}$ $(n = 1, 2, \cdots)$ has no convergent subsequence. Let $W_j = X \setminus \cup_j \{x_j t_{n_j}\}$. Then each W_j is a neighborhood of M which violates the attractor property. In case (β), we proceed similarly. That is, choose sequences $\{t_{n_j}\}$ such that $t_{n_j} \to \infty$ (as $n_j \to \infty$) and $x_j t_{n_j} \to x_j' \in M$. Again take $W_j = X \setminus \cup_j \{x_j t_{n_j}\}$ to violate the attractor property.

To complete the proof, we note that (i) and (ii) clearly imply (iii).

Our final result shows that topological stability, together with a rather weak attraction property imply topological attraction.

THEOREM 9. *Let M be topologically stable. Suppose there is a $W \in \mathcal{N}_M$ such that if $x' \in W$, $\omega(x') \cap M \neq \emptyset$. Then M is a topological attractor.*

Proof. We first observe that if $x, z \in X$ with $z \in \omega(x)$, then $\omega(x) \subset D_1(z)$. For, let $y \in \omega(x)$. Then there are sequences $s_n, t_n \to \infty$ such that $xs_n \to z$, $xt_n \to y$. We may suppose $t_n > s_n$, so $t_n = s_n + \tau_n$, with $\tau_n > 0$. Now $(xs_n)\tau_n = x(s_n + \tau_n) = xt_n \to y$, and $y \in D_1(z)$. (For a related result, compare [3, Lemma 4].)

Now, let $x' \in W$ and suppose $z \in \omega(x') \cap M$. We show $\overline{\gamma^+(x')}$ is compact. If not, $\{x't_n\}$ has no convergent subsequence, for some $t_n \to \infty$. Then, if U is a relatively compact neighborhood of $\{x'\} \cup D_1(z)$, there is a sequence $s_n \to \infty$ such that $\{x's_n\} \in \partial U$. Then (a subsequence of) $x's_n \to x'' \in \omega(x')$, so $x'' \in D_1(z)$, by the above observation, which is a contradiction. Hence $\overline{\gamma^+(x')}$ is compact. Since $\omega(x') \subset D_1(z)$, $\omega(x')$ is compact. By Theorem 7, M is a topological attractor.

Acknowledgment. I would like to thank Peter Seibert for useful conversations on the topics of this paper.

REFERENCES

[1] H. A. ANTOSIEWICZ, *A survey of Lyapunov's second method*, Contributions to the Theory of Nonlinear Oscillations, Vol. 4, Ann. Math. Studies, no. 41, Princeton University Press, Princeton, N.J., 1958, pp. 141–166.

[2] J. AUSLANDER, *On stability of closed sets in dynamical systems*, Lecture Notes in Mathematics, No. 144, Springer-Verlag, New York, 1970, pp. 1–4.

[3] J. AUSLANDER, N. P. BHATIA AND P. SEIBERT, *Attractors in dynamical systems*, Bol. Soc. Mat. Mexicana, 9 (1964), pp. 55–66.

[4] J. AUSLANDER AND P. SEIBERT, *Prolongations and stability in dynamical systems*, Ann. Inst. Fourier (Grenoble), 14 (1964), pp. 237–267.

[5] N. P. BHATIA, *Stability and Lyapunov functions in dynamical systems*, Contributions to Differential Equations, 3 (1964), pp. 175–188.

[6] N. P. BHATIA, *Characteristic properties of stable sets and attractors in dynamical systems*, Instituto Nazionale di Alta Matematico, Symposia Mathematica, 6 (1971), pp. 155–155.

[7] N. P. BHATIA AND G. P. SZEGÖ, *Stability Theory of Dynamical Systems*, Springer-Verlag, New York, Band 161, 1970.

[8] O. HAJEK, *Compactness and asymptotic stability*, Math. Systems Theory, 4 (1970), pp. 154–156.

[9] T. URA, *Sur le courant exterieur à une région invariante*, Funkcial. Ekvac., 2 (1959), pp. 143–200.

SIAM J. MATH. ANAL.
Vol. 8, No. 4, August 1977

THE POINCARÉ–BERTRAND FORMULA: A DERIVATION AND A GENERALIZATION*

L. C. BAIRD, S. SANCAKTAR AND P. F. ZWEIFEL†

Abstract. The Poincaré–Bertrand formula is derived without recourse to arguments involving the complex plane. The derivation is also modified to yield a generalized formula which is seen to simplify some calculations which arise in transport theory.

1. Introduction. The Poincaré–Bertrand formula (PBF) concerns changing the order of integration in a certain class of singular integrals:

$$(1)\qquad \int_{-1}^{1}\int_{-1}^{1}\frac{g(x,y)}{(x-y)(y-z)}\,dx\,dy=-\pi^2 g(z,z)+\int_{-1}^{1}\int_{-1}^{1}\frac{g(x,y)}{(x-y)(y-z)}\,dy\,dx$$

where the singularities are integrated as Cauchy principal values. The PBF holds whenever $z\in(-1,1)$ provided $g(x,y)$ satisfies the Hölder condition discussed in Appendix A. The most noteworthy feature of the PBF is the residual term, $-\pi^2 g(z,z)$. Fubini's theorem, which concerns less singular integrals, might lead one to expect a residual term of zero.

It turns out that PBF plays an important role in describing the diffusion of neutrons through a nuclear reactor [1, p. 69]. Several years ago, the nuclear engineering community became involved in a controversy regarding the physical significance of the residual term [4], [5], [6], [7], [8], [9]. It soon became apparent that there was a degree of confusion regarding the mathematical origins of this term. The confusion was understandable, for derivations of the residual term are traditionally shrouded in esoteric discussions of residues and boundary values of analytic functions [3], [10, p. 61], [12], [13, p. 242]. The controversy was eventually resolved in a mathematically acceptable fashion [6]; but the ultimate conclusions were not altogether satisfying from an intuitive point of view.

It is in this context that the authors have felt compelled to work out a new approach for the derivation of the PBF. We present here a derivation which is *conceptually* simpler than any now in the literature. In particular, we avoid the complex plane by a judicious use of integral tables. Our basic line of reasoning is elementary, and the origin of the residual term is manifestly obvious.

Our quest for a *conceptually* simpler PBF derivation has not necessarily led to a savings in computational effort. In a sense, we do not even avoid the complex plane, for residue analysis was certainly (though not unavoidably) used to construct the integral tables which we must ultimately use. Nevertheless, we feel that our derivation fills the need for a *conceptually* simpler approach to PBF. That approach is essentially spelled out in the opening paragraph of § 2. The remainder of our derivation is simply a filling in of details.

* Received by the editors May 28, 1974, and in final revised form March 29, 1976.

† Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061. The first author is now with the Department of Medical Physics, University of Wisconsin—Madison, Madison, Wisconsin 53706. This work was supported in part by the National Science Foundation under Grant GK-35903.

We also introduce a *generalized* PBF, which is useful in certain practical applications. The generalization consists of adopting a modified principal value convention which can simplify (or even eliminate) the residual term. In Appendix C the generalization is applied to an integral arising in neutron transport theory. The modified principal value turns out to be a natural choice for the problem in question, and the resulting residual term quickly gives an answer otherwise requiring a significant amount of additional computation.

2. The derivation. Our derivation makes a direct appeal to the definition of the Cauchy principal value. Letting I_1 and I_2 denote the double integrals on the left and right sides of (1), respectively, we have (see Appendix B)

$$I_1 = \lim_{\alpha\to 0}\lim_{\beta\to 0}\int_{\Delta_1}\frac{g(x,y)}{(x-y)(y-z)}\,dA, \tag{2a}$$

$$I_2 = \lim_{\gamma\to 0}\lim_{\alpha\to 0}\lim_{\beta\to 0}\int_{\Delta_2}\frac{g(x,y)}{(x-y)(y-z)}\,dA \tag{2b}$$

where dA denotes integration over the regions, Δ_1 and Δ_2, defined by Figs. 1 and 2. A cursory examination might have led one to believe that I_1 and I_2 differ only in

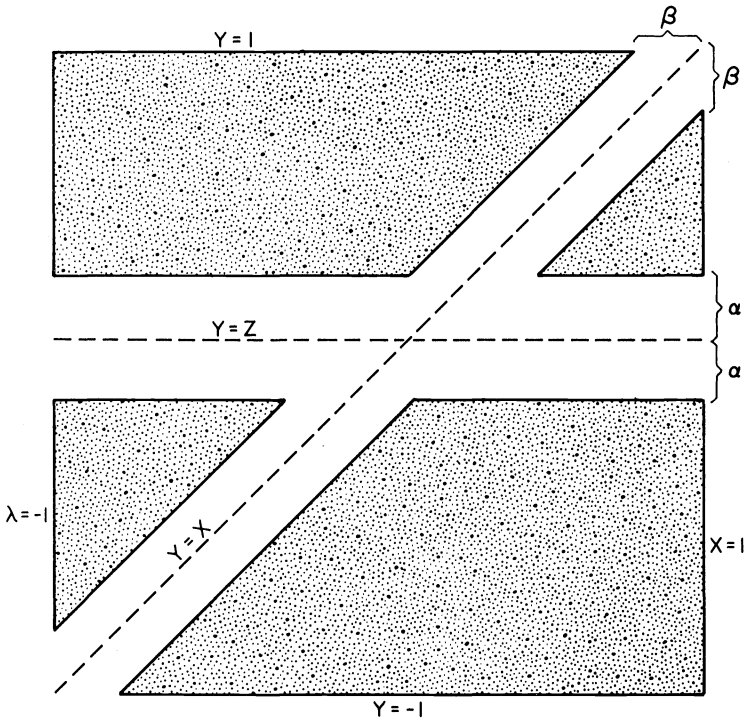


FIG. 1. *The region,* Δ_1, *which is the domain of integration for* I_1

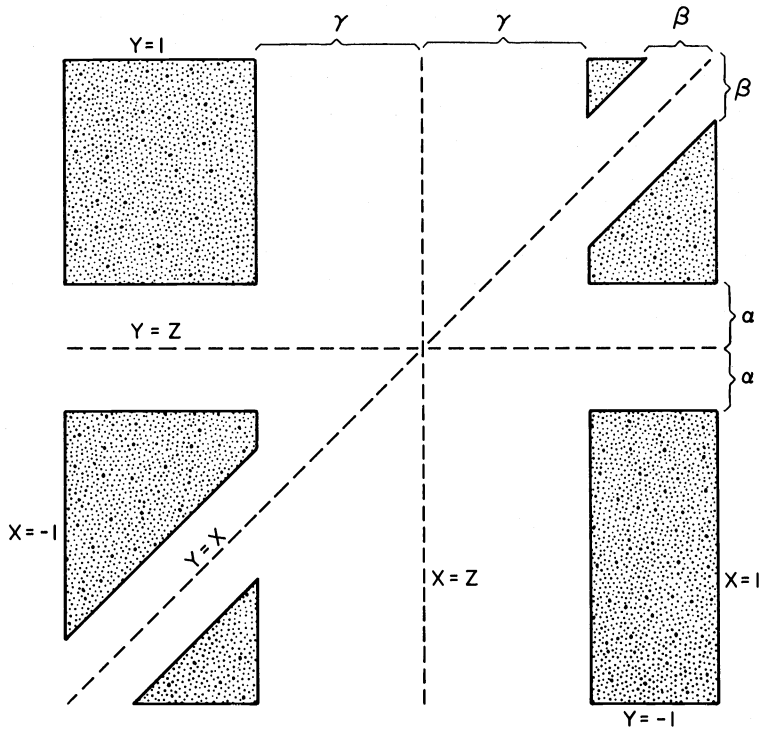


FIG. 2. *The region,* Δ_2, *which is the domain of integration for* I_2

a reversed order of integration. However, we now see that they also differ with respect to the domains of integration, Δ_1 and Δ_2. This difference in the domains is the crucial feature which ultimately leads to the residual term

$$I_1 - I_2 = -\pi^2 g(z, z). \tag{3}$$

To obtain this result we take the difference of equations (2):

$$I_1 - I_2 = \lim_{\gamma\to 0}\lim_{\alpha\to 0}\lim_{\beta\to 0}\int_{\Delta_1-\Delta_2}\frac{g(x, y)}{(x-y)(y-z)}\,dx\,dy. \tag{4}$$

We now proceed to evaluate the right-hand side of (4). To this end it is convenient to partition $(\Delta_1 - \Delta_2)$ into the twelve regions $\{R_1, \cdots, R_{12}\}$ shown in Fig. 3. The integral over the nth region will be denoted by J_n. Thus

$$\begin{aligned} I_1 - I_2 &= \lim_{\gamma\to 0}\lim_{\alpha\to 0}\lim_{\beta\to 0}\sum_{n=1}^{12} J_n \\ &= \lim_{\delta\to 0}\lim_{\gamma\to 0}\lim_{\alpha\to 0}\lim_{\beta\to 0}\sum_{n=1}^{12} J_n \qquad (5) \\ &\equiv \lim \sum_{n=1}^{12} J_n \end{aligned}$$

where the quadruple limit is simply denoted by "lim". The remainder of our derivation consists of the mundane task of evaluating the limit of each J_n integral. In each case the limits of integration are obtained from the geometry of Fig. 3.

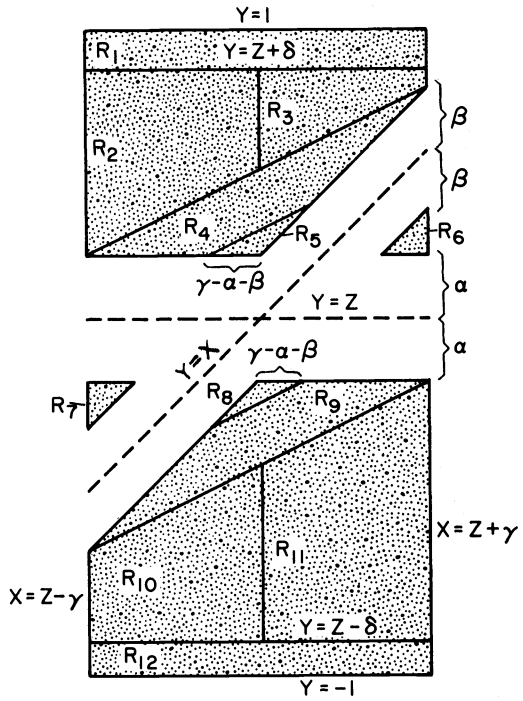


FIG. 3. *The* 12 *regions of* $(\Delta_1 - \Delta_2)$. *The parameter* δ *is a small constant. Diagonal lines bounding the shaded regions are parallel to* $y = x$. *The remaining diagonal lines are parallel to* $2\gamma y = (z + \gamma + \beta - \alpha)x$.

The Hölder condition on $g(x, y)$ assures us that the integrand is bounded on any closed set excluding the lines of singularity. Thus

$$|\lim J_1| \leqq \left(\lim \int_{R_1} dA\right) \sup_{(x,y)\in R_1} |g(x, y)| - 0. \tag{6}$$

Therefore

$$\lim J_1 = 0. \tag{7}$$

Similarly

$$\lim J_{12} = 0. \tag{8}$$

A brief study of Fig. 3 shows that if $(x, y) \in R_6$ then $(2y - x, y) \in R_5$, $(2z - x, 2z - y) \in R_7$, and $(2z - 2y + x, 2z - y) \in R_8$. Furthermore, these four points are located symmetrically with respect to the lines of singularity. We also observe that $(x - y)(y - z)$ only varies with respect to algebraic sign at the four

points in question. In view of these comments, the following equation is easily deduced:

$$J_5+J_6+J_7+J_8 = \int_{R_6} \frac{g(x,y)-g(2y-x,y)+g(2z-x,2z-y)-g(2z-2y+x,2z-y)}{(x-y)(y-z)}. \tag{9}$$

Applying the Lemma of Appendix A we obtain

$$|J_5+J_6+J_7+J_8| \leqq 2^{2+p}A \int_{R_6} (x-y)^{p/2-1}\, dA. \tag{10}$$

It follows that

$$\lim (J_5+J_6+J_7+J_8) = 0. \tag{11}$$

Noting the continuity of $g(x, y)$ we have

$$\lim J_4 = g(z,z) \lim J_4' \tag{12}$$

where

$$\begin{aligned} J_4' &= \int_{R_4} \frac{1}{(x-y)(y-z)}\, dA \\ &= \int_{z+\alpha}^{z+\gamma-\beta} \int_{(2\gamma y - z(\gamma-\beta+\alpha)-\gamma(\alpha+\beta+\gamma))/(\gamma+\beta-\alpha)}^{2y-z-\gamma} \frac{1}{(x-y)(y-z)}\, dx\, dy \\ &+ \int_{z+\gamma-\beta}^{z+\gamma+\beta} \int_{(2\gamma y - z(\gamma-\beta+\alpha)-\gamma(\alpha+\beta+\gamma))/(\gamma+\beta-\alpha)}^{y-\beta} \frac{1}{(x-y)(y-z)}\, dx\, dy \\ &\xrightarrow[\beta\to 0]{} \left(\ln \frac{\gamma-\alpha}{\gamma+\alpha}\right)\left(\ln \frac{\gamma}{\alpha}\right) \\ &\xrightarrow[\alpha\to 0]{} 0. \end{aligned} \tag{13}$$

Equations (12) and (13) imply

$$\lim J_4 = 0. \tag{14}$$

Similarly

$$\lim J_9 = 0. \tag{15}$$

We use the transformation

$$x - z = -\gamma e^{-u} \tag{16}$$

and again invoke the continuity of $g(x, y)$ to obtain

$$\lim J_2 = g(z,z) \lim J_2' \tag{17}$$

where

$$J_2' = \int_{R_2} \frac{1}{(x-y)(y-z)}\,dA$$

$$= \int_0^\infty \int_{((\alpha+\gamma+\beta+2z)-(\gamma+\beta-\alpha)e^{-u})/2}^{\delta+z} \left(\frac{1}{y-z-\gamma e^{-u}} - \frac{1}{y-z}\right) dy\,du$$

(18)
$$\xrightarrow[\alpha,\beta\to 0]{} \int_0^\infty \ln\frac{\delta+\gamma e^{-u}}{\delta}\,\frac{1-e^{-u}}{1+e^{-u}}\,du$$

$$\xrightarrow[\gamma\to 0]{} -\int_0^\infty \ln\frac{e^u+1}{e^u-1}\,du$$

$$= -\pi^2/4.$$

The last step has been accomplished by appeal to a table of definite integrals [11, p. 70]. Thus

(19a)
$$\lim J_2 = -\frac{\pi^2}{4} g(z, z).$$

The remaining integrals are evaluated similarly to obtain

(19b)
$$\lim J_3 = -\frac{\pi^2}{4} g(z, z),$$

(19c)
$$\lim J_{10} = -\frac{\pi^2}{4} g(z, z),$$

(19d)
$$\lim J_{11} = -\frac{\pi^2}{4} g(z, z).$$

The PBF (equation (3)) is obtained by combining (5), (7), (8), (11), (14), (15), (19).

3. The generalized PBF. The usual PBF (equation (1)) employs a principal value convention which deletes a symmetric neighborhood about each line of singularity: $(x = y)$, $(y = z)$, $(z = x)$. One can modify this convention by letting the deleted neighborhoods become *asymmetric* strips along the lines of singularity. We shall denote these asymmetric strips as being the regions between each of the following pairs of lines:

(20a) $y = z + a\alpha,$

(20b) $y = z - \alpha,$

(21a) $y = x + b\beta,$

(21b) $y = x - \beta,$

(22a) $x = z + c\gamma,$

(22b) $x = z - \gamma,$

where $a, b, c > 0$. We define the *generalized* Poincaré–Bertrand formula (GPBF) to be that formula obtained if I_1 and I_2 are evaluated via the asymmetric principal values. To obtain the GPBF we repeat the arguments of § 2. The details of the computation are not particularly interesting and will be omitted. The final result is the GPBF

$$\int_{-1}^{1}\int_{-1}^{1}\frac{g(x,y)}{(x-y)(y-z)}\,dx\,dy = g(z,z)(-\pi^2+\ln b\,\ln c-\ln c\,\ln a-\ln a\,\ln b) + \int_{-1}^{1}\int_{-1}^{1}\frac{g(x,y)}{(x-y)(y-z)}\,dx\,dy \tag{23}$$

where the singularities are integrated as asymmetric principal values à la equations (20), (21), (22). A doubly generalized PBF could have been generated by choosing one set of asymmetry indices $\{a, b, c\}$ for I_1 and a different set for I_2.

We see that the asymmetry indices have rendered (1) and (23) formally dissimilar. An analogous effect occurs when an asymmetry index is used with a one-dimensional principal value:

$$\begin{aligned}\int_{-1}^{1}\frac{f(x)}{x}\,dx\Big|_{\text{symmetric}} &- \int_{-1}^{1}\frac{f(x)}{x}\,dx\Big|_{\text{asymmetric}}\\ &= \lim_{\alpha\to 0}\left(\int_{-1}^{-\alpha}+\int_{\alpha}^{1}\right)\frac{f(x)}{x}\,dx - \lim_{\alpha\to 0}\left(\int_{-1}^{-\alpha}+\int_{a\alpha}^{1}\right)\frac{f(x)}{x}\,dx\\ &= \lim_{\alpha\to 0}\int_{\alpha}^{a\alpha} f(x)x^{-1}\,dx\\ &= f(0)\ln a.\end{aligned} \tag{24}$$

This result is just another form of a well-known theorem [2, p. 11] which says that different regularizations of nonlocally-integrable functions differ only by multiples of the delta distribution.

It is interesting to note that a suitable choice of asymmetry indices completely eliminates the residual term in (23). Thus the generalized PBF is sometimes formally equivalent to Fubini's theorem.

An application of the GPBF appears in Appendix C.

Appendix A. A variety of conditions on $g(x, y)$ could be stipulated under which the PBF would be valid. We shall adopt the Hölder condition used by Muskhelishvili [10, p. 11]. That is, we require that there exist constants $A, B, P, Q > 0$ such that for any two points in the domain of $g(x, y)$ we have

$$|g(x_2, y_2)-g(x_1, y_1)| \leqq A|x_2-x_1|^P + B|y_2-y_1|^Q. \tag{A.1}$$

Clearly there is no loss of generality in assuming that $A = B$ and $P = Q$. Thus (A.1) reduces to

$$|g(x_2, y_2)-g(x_1, y_1)| < A|x_2-x_1|^P + A|y_2-y_1|^P. \tag{A.2}$$

We note the fact that the Hölder condition implies that $g(x, y)$ is continuous.

The following lemma is needed in the evaluation of (9).

LEMMA.

(A.3)

where

$$G(x, y) = |g(x, y)| = |g(2y - x, y) + g(2z - x, 2z - y) - g(2z - 2y + x, 2z - y)|. \tag{A.4}$$

Proof. We first establish the result

$$\begin{aligned} G(x, y) &\leqq |g(x, y) - g(2y - x, y)| \\ &\quad + |g(2z - x, 2z - y) - g(2z - 2y + z, 2z - y)| \\ &\leqq A|2x - 2y|^P + A|-2x + 2y|^P \\ &= 2^{1+P} A|x - y|^P \\ &\leqq 2^{2+P} A|x - y|^P. \end{aligned} \tag{A.5}$$

On the other hand,

$$\begin{aligned} G(x, y) &\leqq |g(x, y) - g(2z - 2y + x, 2z - y)| \\ &\quad + |g(2z - x, 2z - y) - g(2y - x, y)| \\ &\leqq A|2z - 2y|^P + A|2y - 2z|^P + A|2z - 2y|^P + A|2z - 2y|^P. \end{aligned} \tag{A.6}$$

Thus

$$\begin{aligned} G(x, y) &\leqq 2^{2+P} A \min\{|x - y|^P, |y - z|^P\} \\ &= 2^{2+P} A \min\{|x - y|^{P/2}, |y - z|^{P/2}\} \min\{|x - y|^{P/2}, |y - z|^{P/2}\} \\ &\leqq 2^{2+P} A \min\{|x - y|^{P/2}, |y - z|^{P/2}\} \max\{|x - y|^{P/2}, |y - z|^{P/2}\} \\ &= 2^{2+P} A|x - y|^{P/2}|y - z|^{P/2}. \end{aligned} \tag{A.7}$$

Appendix B. It will be convenient to extend the domain of $g(x, y)$ to include the region $(|x| \geqq 1 \geqq |y|)$. Any Hölder continuous extension of $g(x, y)$ will be acceptable; for example,

$$g(x, y) = g(\pm 1, y) \qquad \text{if } \pm x \geqq 1 \geqq |y|. \tag{B.1}$$

We will now proceed to show that equations (2) are valid representations of I_1 and I_2. We first consider I_1, which was originally defined to be the left side of (1):

$$I_1 = \lim_{\alpha \to 0} \int_{-1}^{1} \lim_{\beta \to 0} \int_{-1}^{1} \frac{g(x, y)\gamma(\alpha, |y - z|)\gamma(\beta, |x - y|)}{(x - y)(y - z)}\, dx\, dy \tag{B.2}$$

where

$$\gamma(a, b) = \begin{cases} 1 & \text{if } a \leqq b, \\ 0 & \text{otherwise.} \end{cases} \tag{B.3}$$

The question to be answered is whether or not (B.2) is invariant under an interchange of the β limit and the dy integration:

$$
\begin{aligned}
0 \stackrel{?}{=} & \int_{-1}^{1} \lim_{\beta\to 0} \int_{-1}^{1} \frac{g(x,y)\gamma(\alpha,|y-z|)\gamma(\beta,|x-y|)}{(x-y)(y-z)}\,dx\,dy \\
& - \lim_{\beta\to 0} \int_{-1}^{1}\int_{-1}^{1} \frac{g(x,y)\gamma(\alpha,|y-z|)\gamma(\beta,|x-y|)}{(x-y)(y-z)}\,dx\,dy \\
= & \lim_{\beta\to 0} \int_{-1}^{1}\int_{-1}^{1} \lim_{\nu\to 0} \int_{-1}^{1} \frac{g(x,y)\gamma(\alpha,|y-z|)[\gamma(\nu,|x-y|)-\gamma(\beta,|x-y|)]}{(x-y)(y-z)}\,dx\,dy \\
= & \lim_{\beta\to 0}\left\{\int_{-1}^{z-\alpha}+\int_{z+\alpha}^{1}\right\}\lim_{\nu\to 0}\left\{\int_{y-\beta}^{y-\nu}\gamma(-1,x)+\int_{y+\nu}^{y+\beta}\gamma(x,1)\right\}\frac{g(x,y)}{(x-y)(y-z)}\,dx\,dy \\
= & \lim_{\beta\to 0}\left\{\int_{-1}^{z-\alpha}+\int_{z+\alpha}^{1}\right\}\lim_{\nu\to 0}\left\{-\int_{y-\beta}^{y-\nu}\gamma(x,-1)+\int_{y-\beta}^{y-\nu}\right. \\
& \left.+\int_{y}^{y+\beta}-\int_{y+\nu}^{y+\beta}\gamma(1,x)\right\}\frac{g(x,y)}{(x-y)(y-z)}\,dx\,dy \qquad \text{(B.4)} \\
= & \lim_{\beta\to 0}\left\{\int_{-1}^{z-\alpha}+\int_{z+\alpha}^{1}\right\}\lim_{\nu\to 0}\left\{\int_{y-\beta}^{y-\nu}+\int_{y+\nu}^{y+\beta}\right\}\frac{g(x,y)}{(x-y)(y-z)}\,dx\,dy \\
& - \lim_{\beta\to 0}\int_{-1}^{\beta-1}\lim_{\nu\to 0}\int_{y-\beta}^{\min\{-1,y-\nu\}}\frac{g(x,y)}{(x-y)(y-z)}\,dx\,dy \\
& - \lim_{\beta\to 0}\int_{1-\beta}^{1}\lim_{\nu\to 0}\int_{\max\{1,y+\nu\}}^{y+\beta}\frac{g(x,y)}{(x-y)(y-z)}\,dx\,dy.
\end{aligned}
$$

Equation (B.4) is seen to involve points outside the original domain of $g(x, y)$. This situation is justified in the opening paragraph of Appendix B.

The three terms of (B.4) will be handled individually. We first restrict our attention to the term

$$
\begin{aligned}
T_1 = & \lim_{\beta\to 0}\left\{\int_{-1}^{z-\alpha}+\int_{z+\alpha}^{1}\right\}\lim_{\nu\to 0}\left\{\int_{y-\beta}^{y-\nu}+\int_{y+\nu}^{y+\beta}\right\}\frac{g(x,y)}{(x-y)(y-z)}\,dx\,dy \\
= & \lim_{\beta\to 0}\left\{\int_{-1}^{z-\alpha}+\int_{z+\alpha}^{1}\right\}\lim_{\nu\to 0}\int_{y+\nu}^{y+\beta}\frac{g(x,y)-g(2y-x,y)}{(x-y)(y-z)}\,dx\,dy \qquad \text{(B.5)} \\
= & \lim_{\beta\to 0}\left\{\int_{-1}^{z-\alpha}+\int_{z+\alpha}^{1}\right\}\int_{y}^{y+\beta}\frac{g(x,y)-g(2y-x,y)}{(x-y)(y-z)}\,dx\,dy.
\end{aligned}
$$

We now invoke the Hölder condition to obtain

$$
\begin{aligned}
|T_1| \leqq & \lim_{\beta\to 0}\left\{\int_{-1}^{z-\alpha}+\int_{z+\alpha}^{1}\right\}\int_{y}^{y+\beta}\frac{2^P A(x-y)^{P-1}}{|y-x|}\,dx\,dy \qquad \text{(B.6)} \\
= & \lim_{\beta\to 0}\frac{1}{P}2^P\beta^P A\ln\frac{1-z^2}{\alpha^2}=0.
\end{aligned}
$$

Having disposed of T_1, we turn to the term

$$T_2 = \lim_{\beta\to 0}\int_{-1}^{\beta-1}\lim_{\nu\to 0}\int_{y-\beta}^{\min\{-1,y-\nu\}}\frac{g(x,y)}{(x-y)(y-z)}\,dx\,dy$$
$$= \lim_{\beta\to 0}\int_{-1}^{\beta-1}\int_{y-\beta}^{-1}\frac{g(x,y)}{(x-y)(y-z)}\,dx\,dy. \tag{B.7}$$

Letting

$$S = \sup_{\substack{y-\beta\leqq x\leqq -1\\ -1\leqq y\leqq \beta-1}}\frac{|g(x,y)}{|y-z|} \tag{B.8}$$

we easily obtain

$$|T_2| \leqq \lim_{\beta\to 0} S\int_{-1}^{\beta-1}\int_{y-\beta}^{-1}|x-y|^{-1}\,dx\,dy$$
$$= \lim_{\beta\to 0}\beta S = 0. \tag{B.9}$$

Similarly

$$T_3 = \lim_{\beta\to 0}\int_{1-\beta}^{1}\lim_{\nu\to 0}\int_{\max\{1,y+\nu\}}^{y+\beta}\frac{g(x,y)}{(x-y)(y-z)}\,dx\,dy = 0. \tag{B.10}$$

In view of (B.6), (B.9), (B.10) we conclude that the zero does indeed belong on the left side of (B.4). That is, (2a) is valid. The validity of (2b) is established by applying the preceding argument to I_2 instead of I_1.

Appendix C: An application of GPBF. In many areas of linear transport theory the Boltzman equation is solved in terms of so-called "singular eigenfunctions" which, for one speed neutron transport theory, take the simple form [1, Chap. 4]

$$\varphi_\nu(\mu) = \frac{c\nu}{2}\mathcal{P}\frac{1}{\nu-\mu} + \lambda(\nu)\delta(\nu-\mu). \tag{C.1}$$

Here ν is the "eigenvalue", $\mu = v_x/v$ is the normalized x-component of the neutron velocity and c is a nonnegative parameter. The usual (symmetric) Cauchy principal value is denoted by $\mathcal{P}$.

In performing eigenfunction expansions in order to solve boundary value problems in transport theory, it is necessary in certain cases to integrate "products" of these eigenfunctions [6]. By redefining the principal value in terms of an asymmetric $\mathcal{P}$ we can eliminate the term involving $\delta(\nu-\mu)$ [2, p. 11] specifically

$$\varphi_\nu(\mu) = \frac{c\nu}{2}\mathcal{P}_a\frac{1}{\nu-\mu} \tag{C.2}$$

where $\mathcal{P}_a$ means that integrals involving $1/(\nu-\mu)$ are to be performed on the interval

$$\lim_{\alpha\to 0}[-1, \nu-\alpha]\cup[\nu+a\alpha, 1]$$

where

$$a = e^{(2\lambda)/(c\nu)}.$$

Clearly it is easier to multiply expressions like (C.2) than those like (C.1).

When the order of integration is reversed with such a product one must take care to use the correct asymmetric PBF. This is no increase in labor over the usual case [1, p. 70], in which the symmetric PBF has to be used anyway. In particular, a typical integral which occurs in transport theory is obtained from applying "orthogonality" to an eigenfunction expansion of the form

$$\psi(\mu) = \int_{-1}^{1} A(\nu)\varphi_\nu(\mu)\, d\nu \tag{C.3}$$

where

$$\int_{-1}^{1} \mu\psi(\mu)\varphi_{\nu'}(\mu)\, d\mu = \int_{-1}^{1} \mu\varphi_{\nu'}(\mu) \int_{-1}^{1} A(\nu)\varphi_\nu(\mu)\, d\nu\, d\mu. \tag{C.4}$$

On the right-hand side of (C.4) if the order of integration can be reversed, the integration over μ can be carried out. Thus, according to the GPBF (equation (23) with $a = b = c = e^{(2\lambda)/(c\nu)}$) we must substitute a term $\nu(\pi^2+(4\lambda^2)/(c^2\nu^2))$. This immediately gives the Case normalization coefficient $N(\nu)$ [1, p. 71] with no further work.

When products of more singular eigenfunctions are piled up, even greater simplifications occur. One such application occurs in a specific derivation of the family of spectral projections for the transport operator [14], but the details are similar to the case derived above.

REFERENCES

[1] K. M. CASE AND P. F. ZWEIFEL, *Linear Transport Theory*, Addison-Wesley, Reading, Mass., 1967.

[2] I. M. GELFAND AND G. E. SHILOV, *Generalized Functions: Vol. I. Properties and Operations*, Academic Press, New York, 1964.

[3] G. M. HARDY, *The theory of Cauchy's principal values*, Proc. London Math. Soc., 7 (1909), pp. 181–208.

[4] A. M. JACOBS AND J. J. MCINERNY, *On the Green's function of monoenergetic transport theory*, Nuclear Sci. Engrg., 22 (1965), pp. 119–120.

[5] A. M. JACOBS AND R. D. MOYER, *Further comments on the use of generalized functions in neutron transport theory*, Ibid., 29 (1966), p. 426.

[6] H. G. KAPER, *Note on the use of generalized functions and the Poincaré–Bertrand formula in neutron transport theory*, Ibid., 24 (1966), pp. 423–425.

[7] I. KUSCER AND N. J. MCCORMICK, *Comment to the preceding letter by Kaper*, Ibid., 24 (1966), pp. 425–426.

[8] ———, *On the use of the Poincaré–Bertrand formula in neutron transport theory*, Ibid., 23 (1965), p. 404.
[9] JOSEPH J. MCINERNY, *A solution of the space-energy-angle-dependent neutron slowing-down problem*, Ibid., 22 (1965), pp. 215–234.
[10] N. I. MUSKHELISHVILI, *Singular Integral Equations*, 2nd ed., P. Noordhoff, Groningen, the Netherlands, 1953.
[11] B. O. PIERCE AND R. M. FOSTER, *A Short Table of Integrals*, 4th ed., Ginn, New York, 1957.
[12] H. POINCARÉ, *Lecons de Mecanique Celeste*, t. III, 1910.
[13] BERNARD W. ROSS, *Analytic Functional Distributions in Physics and Engineering*, John Wiley, New York, 1969.
[14] S. SANCAKTAR, unpublished manuscript.
[15] F. G. TRICOMI, *Mem. Accad. Lincei Roma*, no. 5, 14 (1923), pp. 133–267; Engl. transl., Brown University Press, Providence, R.I., 1948.
[16] ———, *On the finite Hilbert transformation*, Quart. J. Math. Oxford, Ser., 2 (1951), pp. 199–211.

SIAM J. MATH. ANAL.
Vol. 8, No. 4, August 1977

AN INVARIANCE PROPERTY OF SOLUTIONS TO SECOND ORDER DIFFERENTIAL INEQUALITIES IN ORDERED BANACH SPACES*

RUSSELL C. THOMPSON†

Abstract. Let B be a Banach space, K be a cone in B and $I=[a, b]$. Conditions are imposed on $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')$, $\mathbf{f}: I\times B\times B\to B$ such that the following invariance property holds: $\mathbf{u}''+\mathbf{f}(x, \mathbf{u}, \mathbf{u}')\in -K$, $\mathbf{u}(x_1)\in K, \mathbf{u}(x_2)\in K, a\leqq x_1<x_2\leqq b$ implies $\mathbf{u}(x)\in K$ for all $x\in[x_1, x_2]$. From this property, comparison theorems for solutions to differential inequalities are obtained. These comparison results give sufficient conditions for C^2 functions to be subfunctions with respect to two point boundary value problems. These results extend earlier results in R^n by Heimes to more general types of partial orderings in both infinite and finite dimensional spaces. The approach taken in this paper relaxes certain coupling restrictions present in earlier results of this type.

Introduction. In this paper, we investigate an invariance property of solutions to second order differential inequalities in a Banach space B in which the inequality relation is induced by a cone K in B. The invariance property can be briefly described as follows: conditions are imposed on the function $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')$ such that if $\mathbf{u}(x)$ is twice continuously differentiable in $[a, b]$ and satisfies $\mathbf{u}''+\mathbf{f}(x. \mathbf{u}, \mathbf{u}')\in -K$, $\mathbf{u}(x_1)\in K$, and $\mathbf{u}(x_2)\in K$, then $\mathbf{u}(x)\in K$ for all $x\in[x_1, x_2]$, where $x_1, x_2\in[a, b]$, $x_1<x_2$, $\mathbf{f}: [a, b]\times B\times B\to B$, and $-K=\{-\mathbf{u}: \mathbf{u}\in K\}$. This property is a generalization of the minimum principle for one dimensional, second order, ordinary differential equations (cf. [6]). Using this invariance property, we develop comparison theorems for solutions to differential inequalities. Such results play an important role in the existence theory for two point boundary value problems based on Perron's method (cf. [2], [3]). The comparison theorems give sufficient conditions for twice continuously differentiable functions to be subfunctions with respect to two point boundary value problems and also give conditions under which solutions to boundary value problems are unique when they exist. They have also recently been used to obtain error estimates for the numerical method of lines in the approximate solution of elliptic boundary value problems (cf. [12]).

Our conditions on $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')$ are extensions to general partial orderings of some conditions used by Heimes in [2]. By making this extension, we are able to obtain comparison results for a larger class of equations. These results permit coupling to occur in the $\mathbf{u}'$ variable and also more general types of dependence on $\mathbf{u}$. The work in the present paper makes use of some techniques developed by W. Walter [17] and Volkmann [14], [15] in studies of first order equations using quasi-monotone functions. Their results on quasi-monotone functions have been extended to results on flow invariant sets for initial value problems (cf. [7], [16]). The approach used in these results on flow invariant sets has also recently been applied to show existence of solutions to boundary value problems for second order differential equations in a Banach space (cf. [8]). Some general studies of the role of differential inequalities in obtaining upper and lower bounds on solutions to boundary value problems have been made recently by J. Schröder; see [9], [10], [11] for further references in this area.

* Received by the editors December 11, 1975, and in revised form May 10, 1976.

† Department of Mathematical Sciences, Northern Illinois University, De Kalb, Illinois 60115.

The remainder of this paper is arranged in the following way. In § 1, some basic properties of cones in a Banach space are reviewed. These properties are used in the development of the invariance property. A detailed development of these cone properties can be found in [4]. Section 2 contains the proof of the invariance property in the special case where $\mathbf{u}''+\mathbf{f}(x, \mathbf{u}, \mathbf{u}') \in \operatorname{int}(-K)$. This restriction is removed in § 3 in the case where $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')$ satisfies a Lipschitz condition in $\mathbf{u}'$. In § 4, we use the results in § 2 and § 3 to derive comparison theorems for solutions to differential inqualities. A theorem on upper and lower bounds for solutions to boundary value problems in terms of solutions to differential inequalities is obtained and the uniqueness of solutions to boundary value problems is discussed. The final section contains some examples which illustrate the results in the paper.

1. Definitions and some fundamental properties of cones in a Banach space. Let B be a real Banach space with norm $|\cdot|$ and let B^* denote the set of continuous linear functionals on B. The norm of an element ϕ of B^* will be denoted by $\|\phi\|$. A cone in B is defined to be a nonempty, closed, convex subset K of B with the properties: $\alpha \mathbf{x} \in K$ for each $\mathbf{x} \in K$, $\alpha \geqq 0$; and at least one of $\mathbf{x}$ and $-\mathbf{x}$ is not an element of K. We denote by int K, the interior of K. If int $K \neq \varnothing$, then K is called a solid cone. If $K \subseteq B$ is a cone, then it is possible to define a partial ordering $\leqq$ in B in the following way: for $\mathbf{x}, \mathbf{y} \in B$, $\mathbf{x} \leqq \mathbf{y}$ if and only if $\mathbf{y}-\mathbf{x} \in K$. This partial ordering has the usual properties of the sign of inequality (cf. [4], [5]). Further, if K is a solid cone and $\mathbf{y}-\mathbf{x} \in \operatorname{int} K$, then we write $\mathbf{x}<\mathbf{y}$. In terms of the partial ordering $\leqq$, the statement $\mathbf{x} \in K$ is equivalent to $\mathbf{x} \geqq \mathbf{0}$.

A linear functional $\phi \in B^*$ is called a positive linear functional if $\phi(\mathbf{x}) \geqq 0$ whenever $\mathbf{x} \geqq 0$. If $\mathbf{x}_0 \in K$ and $\mathbf{x}_0 \notin \operatorname{int} K$, then there exists a positive linear functional ϕ such that $\phi(\mathbf{x}_0)=0$. If ϕ is a positive linear functional, $\phi \not\equiv 0$, and $\mathbf{x}_0 \in \operatorname{int} K$, then $\phi(\mathbf{x}_0)>0$. The set of positive linear functionals on K is a closed subset of B^*. Since $\phi(\mathbf{x}) \geqq 0$ for all $\mathbf{x} \in K$, it follows that K is contained in the closed halfspace $C_\phi=\{\mathbf{x} \in B | \phi(\mathbf{x}) \geqq 0\}$. Thus the positive linear functionals are support functionals for the convex set K. If ϕ is a positive linear functional, then so is $\alpha\phi$, $\alpha>0$, and furthermore $\{\mathbf{x} \in B | \phi(\mathbf{x}) \geqq 0\}=\{\mathbf{x} \in B | \alpha\phi(\mathbf{x}) \geqq 0\}$. If K is a cone in B, then K is the intersection of all the closed half-spaces which support it. We will denote by U^* the closed unit ball in B^*, $U^*=\{\phi \in B^* | \|\phi\| \leqq 1\}$; and K^* will denote the collection of all positive linear functionals in B^*. If $\mathscr{K} \subseteq K^*$ and $K=\cap\{C_\phi | \phi \in \mathscr{K}\}$, then we will say that $\mathscr{K}$ generates K. So, for example, if $B=R^n$, then $\mathscr{K}=\{\phi_1, \cdots, \phi_n\}$ where $\phi_j(\mathbf{x})=x_j$ generates the cone $K=\{x \in R^n | x_j \geqq 0, j=1, \cdots, n\}$. For $x \in R^n$ we take $|\mathbf{x}|=(\sum_{j=1}^n x_j^2)^{1/2}$.

2. The basic invariance condition. In this section we will assume that K is a solid cone in B and that $\mathscr{K} \subseteq U^*$ generates K. $\bar{\mathscr{K}}$ denotes the closure of $\mathscr{K}$ and $I=[a, b]$.

THEOREM 2.1. *Let* $\mathbf{f}$: $I \times B \times B \to B$ *satisfy the following condition*:

$$(2.1) \quad \phi \in \bar{\mathscr{K}}, \quad \phi(u)=\inf\{\psi(\mathbf{u}) | \psi \in \mathscr{K}\} \leqq 0, \quad \phi(\mathbf{u}')=0 \quad \Rightarrow \quad \phi(\mathbf{f}(x, \mathbf{u}, \mathbf{u}')) \geqq 0.$$

If $\mathbf{u} \in C^2(I, B)$ *satisfies*

$$(2.2) \qquad \mathbf{u}''(\mathrm{x})+\mathbf{f}(x, \mathbf{u}(x), \mathbf{u}'(x))<\mathbf{0}, \qquad x \in I,$$

and for $x_1 < x_2$, $x_1, x_2 \in I$, $\mathbf{u}(x_1) \geqq 0$ *and* $\mathbf{u}(x_2) \geqq 0$, *then* $\mathbf{u}(x) \geqq 0$ *for* $x \in [x_1, x_2]$.

The proof of this theorem makes use of the following lemmas.

LEMMA 2.2. *Let* $S \subseteq U^*$. *If* $\mathbf{u} \in C(I, B)$, *then*

$$\Phi(x) = \inf \{\phi(\mathbf{u}(x)) | \phi \in S\} \tag{2.3}$$

defines a continuous real valued function on I.

Proof. For each $\phi \in S$, we have $|\phi(\mathbf{u}(x))| \leqq \|\phi\| \, |\mathbf{u}(x)| \leqq |\mathbf{u}(x)|$. Therefore $\Phi(x)$ is defined for all $x \in I$.

Let $x \in I$ and let $\varepsilon > 0$ be given. We will show that $\Phi(x)$ is continuous at x by considering two cases. Let s be a point in I.

Case 1. $\Phi(x) - \Phi(s) \leqq 0$. In this case, $|\Phi(x) - \Phi(s)| = \Phi(s) - \Phi(x)$. From (2.3) there exists a $\psi \in S$ such that $\psi(\mathbf{u}(x)) - \varepsilon/2 < \Phi(x)$. Hence it follows that

$$\begin{aligned} |\Phi(x) - \Phi(s)| &= \Phi(s) - \Phi(x) \\ &\leqq \psi(\mathbf{u}(s)) - \psi(\mathbf{u}(x)) + \frac{\varepsilon}{2} \\ &\leqq \|\psi\| \, |\mathbf{u}(s) - \mathbf{u}(x)| + \frac{\varepsilon}{2} \\ &\leqq |\mathbf{u}(x) - \mathbf{u}(s)| + \frac{\varepsilon}{2}. \end{aligned} \tag{2.4}$$

Case 2. $\Phi(x) - \Phi(s) \geqq 0$. In this case, a similar argument shows $|\Phi(x) - \Phi(s)| \leqq |\mathbf{u}(x) - \mathbf{u}(s)| + \varepsilon/2$.

In view of the result of these cases, if $\delta > 0$ is chosen so that $|\mathbf{u}(x) - \mathbf{u}(s)| < \varepsilon/2$ whenever $|x - s| < \delta$, then $|\Phi(x) - \Phi(s)| < \varepsilon$ for $|x - s| < \delta$. Thus Φ is continuous at x. But since x was chosen arbitrarily, it follows that Φ is continuous on I. □

LEMMA 2.3. *Let* $S \subseteq U^*$ *and let* $\mathbf{u} \in B$. If $d = \inf \{\phi(\mathbf{u}) | \phi \in S\}$ *then there is a* $\psi \in \bar{S}$ *such that* $\psi(\mathbf{u}) = d$.

Proof. Since U^* is compact in the weak* topology, $\bar{S}$ is compact. For $\mathbf{u} \in B$, $\mathbf{u}$ fixed, the map $\phi \to \phi(\mathbf{u})$ is weak* continuous. Hence $\bar{S}(\mathbf{u}) = \{\phi(u) : \phi \in \bar{S}\}$ is compact and contains d. □

Proof of Theorem 2.1. Let $\Phi(x)$ be defined by the formula (2.3) relative to the function $\mathbf{u}(x)$. By the result of Lemma 2.2, $\Phi(x)$ is continuous. The inequalities $\mathbf{u}(x_1) \geqq 0$ and $\mathbf{u}(x_2) \geqq 0$ imply that $\Phi(x_1) \geqq 0$ and $\Phi(x_2) \geqq 0$ respectively. The conclusion of the theorem is equivalent to the inequality $\Phi(x) \geqq 0$ for all $x \in [x_1, x_2]$. Suppose, contrary to this conclusion, that $\Phi(x) < 0$ for some $x \in [x_1, x_2]$. By continuity, there exists an $x_0 \in (x_1, x_2)$ such that $\Phi(x_0) = \min \{\Phi(x) | x \in [x_1, x_2]\} < 0$.

Applying Lemma 2.3 at $\mathbf{u} = \mathbf{u}(x_0)$, we have that there exists a positive linear functional $\psi \in \bar{\mathcal{K}}$ such that $\psi(\mathbf{u}(x_0)) = \Phi(x_0)$. Define $h(x) = \psi(\mathbf{u}(x))$. Since $\psi \in \bar{\mathcal{K}}$, it follows that $h(x_1) \geqq 0$, $h(x_2) \geqq 0$ and $h(x_0) = \min \{h(x) : x \in [x_1, x_2]\} = \Phi(x_0)$, that is, $h(x)$ has a negative minimum at x_0. But then

$$h'(x_0) = \frac{d}{dx} \psi(\mathbf{u}(x))|_{x = x_0} = \psi(\mathbf{u}'(x_0)) = 0;$$

and then (2.1) and (2.2) imply that

$$h''(x_0)=\psi(\mathbf{u}''(x_0))<\psi(-\mathbf{f}(x_0,\mathbf{u}(x_0),\mathbf{u}'(x_0)))\leqq 0. \tag{2.5}$$

This last inequality contradicts the assumption that $h(x)$ has a negative minimum at x_0. From this we conclude that $\Phi(x)\geqq 0$, i.e., $\mathbf{u}(x)\geqq 0$. □

Remark. In order to carry out the above proof, it is sufficient to have the inequality

$$\phi(\mathbf{u}''+\mathbf{f}(x,\mathbf{u},\mathbf{u}'))<0, \qquad x\in \operatorname{int} I$$

hold for all $\phi\in\bar{\mathcal{K}}$.

3. Admissibility of the $\leqq$ sign. In this section, some conditions are given on the function $\mathbf{f}(x,\mathbf{u},\mathbf{u}')$, which allow us to replace the strict inequality in (2.2) with $\leqq$. As in § 2, we will assume that K is a solid cone in B and that $\mathcal{K}\subseteq U^*$ generates K. We suppose further that there is a $\mathbf{u}_0\in\operatorname{int} K$ such that $\inf\{\phi(\mathbf{u}_0)|\phi\in\mathcal{K}\}=\delta>0$, for some constant δ. The set of positive linear functionals $\mathcal{K}_{\mathbf{u}_0}$ is defined as follows:

$$\mathcal{K}_{\mathbf{u}_0}=\left\{\phi\in B^*|\phi(\mathbf{u})=\frac{\delta}{\psi(\mathbf{u}_0)}\psi(\mathbf{u});\ \psi\in\mathcal{K},\ \mathbf{u}\in B\right\}. \tag{3.1}$$

It is immediate that $\mathcal{K}_{\mathbf{u}_0}\subseteq U^*$, $\mathcal{K}_{\mathbf{u}_0}$ generates K, and $\phi(\mathbf{u}_0)=\phi$ for every $\phi\in\mathcal{K}_{\mathbf{u}_0}$.

THEOREM 3.1. *Let* $\mathbf{f}(x,\mathbf{u},\mathbf{u}')$ *satisfy the conditions: for* $(x,\mathbf{u},\mathbf{u}')$, $(x,\mathbf{v},\mathbf{v}')\in I\times B\times B$,

$$\begin{aligned}\phi\in\bar{\mathcal{K}}_{\mathbf{u}_0},\quad \phi(\mathbf{u}-\mathbf{v})=\inf\{\psi(\mathbf{u}-\mathbf{v})|\psi\in\mathcal{K}_{\mathbf{u}_0}\}\leqq 0,\quad \phi(\mathbf{u}'-\mathbf{v}')=0\\ \Rightarrow \phi(\mathbf{f}(x,\mathbf{u},\mathbf{u}'))\geqq\phi(\mathbf{f}(x,\mathbf{v},\mathbf{v}));\end{aligned} \tag{3.2}$$

and

$$\mathbf{f}(x,\mathbf{0},\mathbf{0})\geqq\mathbf{0}. \tag{3.3}$$

Assume further that $\mathbf{f}(x,\mathbf{u},\mathbf{u}')$ *satisfies a Lipschitz condition in* $\mathbf{u}'$ *on closed, bounded subsets of* $I\times B^2$. *If* $\mathbf{u}(x)\in C^2(I,B)$ *satisfies the inequality*

$$\mathbf{u}''+\mathbf{f}(x,\mathbf{u},\mathbf{u}')\leqq\mathbf{0}, \tag{3.4}$$

and for $x_1, x_2\in I$, $x_1<x_2$, $\mathbf{u}(x_1)\geqq\mathbf{0}$ *and* $\mathbf{u}(x_2)\geqq\mathbf{0}$, *then* $\mathbf{u}(x)\geqq\mathbf{0}$ for all $x\in[x_1,x_2]$.

Proof. Suppose that $\mathbf{u}(x)\ngeqq\mathbf{0}$ for all $x\in[x_1,x_2]$. Then

$$\Phi(x)=\inf\{\phi(\mathbf{u}(x))|\phi\in\mathcal{K}_{\mathbf{u}_0}\}$$

satisfies $\Phi(x_1)\geqq 0$, $\Phi(x_2)\leqq 0$, and $\Phi(x)<0$ for some $x\in(x_1,x_2)$. By the result of Lemma 2.1, Φ is continuous; hence there is a point $x_0\in(x_1,x_2)$ at which Φ assumes a negative minimum. Let $\varepsilon>0$ be a number such that $\Phi(x_0)=\min\{\Phi(x)|x\in[x_1,x_2]\}=-4\varepsilon$. We define $\mathbf{w}(x)$, $\mathbf{w}: I\to B$, by $\mathbf{w}(x)=\mathbf{u}(x)+\rho(x)\mathbf{u}_0$; where $\rho(x)$ is a solution of the scalar equation $\rho''=(L+1)\rho'$ satisfying $0\leqq\rho<\varepsilon/|\mathbf{u}_0|$, and $-1\leqq\rho'\leqq-\gamma<0$ for some $0<\gamma<1$ and L is a Lipschitz constant to be specified below.

If we take $\mathbf{v}=\mathbf{v}'=0$ in (3.2) and use (3.3), then we obtain for each $\phi\in\bar{\mathcal{K}}_{\mathbf{u}_0}$

$$\phi(\mathbf{u})=\inf\{\psi(\mathbf{u})|\psi\in\mathcal{K}_{\mathbf{u}_0}\}\leqq 0,\quad \phi(\mathbf{u}')=0 \quad\Rightarrow\quad \phi(\mathbf{f}(x,\mathbf{u},\mathbf{u}'))\geqq\phi(\mathbf{f}(x,\mathbf{0},\mathbf{0}))\geqq 0.$$

Thus $\mathbf{f}(x,\mathbf{u},\mathbf{u}')$ satisfies (2.1).

Let L be the Lipschitz constant for $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')$ corresponding to the closed bounded set $\{(x, \mathbf{y}, \mathbf{z}) \in I \times B \times B | x \in I,\ |\mathbf{y}| \leqq \|\mathbf{u}\|_1 + \varepsilon/|\mathbf{u}_0|,\ |\mathbf{z}| \leqq \|\mathbf{u}'\|_1 + 1\}$, where $\|\mathbf{u}\|_1 = \max\{|\mathbf{u}(x)|\ | x \in I\}$. Then, it is easy to show that

$$(3.5)\quad |\phi(\mathbf{f}(x, \mathbf{u}, \mathbf{u}' + \rho'\mathbf{u}_0) - \mathbf{f}(x, \mathbf{u}, \mathbf{u}'))| \leqq |\mathbf{f}(x, \mathbf{u}, \mathbf{u}' + \rho'\mathbf{u}_0) - \mathbf{f}(x, \mathbf{u}, \mathbf{u}')| \leqq L|\rho'|\,|\mathbf{u}_0|$$

holds for every $\phi \in \mathcal{K}_{\mathbf{u}_0}$. Moreover, $\phi(\mathbf{u} - \mathbf{w}) = -\rho\delta$ for every $\phi \in \mathcal{K}_{\mathbf{u}_0}$; therefore (3.2) implies that

$$(3.6)\qquad \phi(\mathbf{f}(x, \mathbf{w}, \mathbf{w}') - \mathbf{f}(x, \mathbf{u}, \mathbf{w}')) \leqq 0$$

for every $\phi \in \mathcal{K}_{\mathbf{u}_0}$.

Using (3.5) and (3.6), we obtain for $\phi \in \mathcal{K}_{\mathbf{u}_0}$,

$$\begin{aligned}
\phi(\mathbf{w}'' + \mathbf{f}(x, \mathbf{w}, \mathbf{w}')) &= \phi(\mathbf{u}'' + \rho''\mathbf{u}_0 + \mathbf{f}(x, \mathbf{w}, \mathbf{w}')) \\
&\leqq \phi(\mathbf{f}(x, \mathbf{w}, \mathbf{w}') - \mathbf{f}(x, \mathbf{u}, \mathbf{u}')) + \rho''\delta \\
&= \phi(\mathbf{f}(x, \mathbf{w}, \mathbf{w}') - \mathbf{f}(x, \mathbf{u}, \mathbf{w}')) + \phi(\mathbf{f}(x, \mathbf{u}, \mathbf{w}') - \mathbf{f}(x, \mathbf{u}, \mathbf{u}')) + \rho''\delta \\
&\leqq L|\rho'|\,|\mathbf{u}_0| + (L+1)\rho'|\mathbf{u}_0| \\
&= |\mathbf{u}_0|\rho' \leqq -\gamma|\mathbf{u}_0| < 0.
\end{aligned}$$

Therefore $\phi(\mathbf{w}'' + \mathbf{f}(x, \mathbf{w}, \mathbf{w}')) < 0$ for every $\phi \in \mathcal{K}_{\mathbf{u}_0}$.

It follows from the remark following the proof of Theorem 2.1 that $\mathbf{w}(x) \geqq 0$ for all $x \in I$, since $\mathbf{w}(x_1) \geqq 0$ and $\mathbf{w}(x_2) \geqq 0$. However, at the point x_0,

$$\begin{aligned}
\inf\{\phi(\mathbf{w}(x_0)) | \phi \in \mathcal{K}_{\mathbf{u}_0}\} &\leqq \boldsymbol{\Phi}(x_0) + \inf\{\phi(\rho(x_0)\mathbf{u}_0) | \phi \in \mathcal{K}_{\mathbf{u}_0}\} \\
&= -4\varepsilon + \frac{\varepsilon}{|\mathbf{u}_0|}\delta \leqq -3\varepsilon < 0
\end{aligned}$$

which contradicts the inequality $\mathbf{w}(x) \geqq 0$. □

Remark. In Theorem 3.1 it is sufficient to require that (3.2) hold on the set $\{(x, \mathbf{y}, \mathbf{z}) \in I \times B \times B | x \in I,\ |\mathbf{y}| \leqq \|\mathbf{u}\|_1 + \|\mathbf{v}\|_1 + \varepsilon/|\mathbf{u}_0|,\ |\mathbf{z}| \leqq \|\mathbf{u}'\|_1 + \|\mathbf{v}'\|_1 + 1\}$.

In many cases of interest, the assumption that int $K \neq \varnothing$ is quite restrictive. For example the usual cones in the L_p and l^p spaces $1 \leqq p < \infty$ have empty interiors. In these cases the proof of the theorem above cannot be carried out because of its dependence on the point $\mathbf{u}_0$. By modifying the condition (2.1) we are able to obtain a result for cones whose interior is empty.

THEOREM 3.2. *Let K be a cone in B (where possibly* int $K = \varnothing$*) generated by a set of positive linear functionals $\mathcal{K}$ and let* $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')$ *satisfy*

$$(3.7)\qquad \phi \in \bar{\mathcal{K}},\quad \phi(\mathbf{u}) = \inf\{\psi(\mathbf{u}) | \psi \in \mathcal{K}\} < 0,\quad \psi(\mathbf{u}') = 0 \quad \Rightarrow \quad \phi(\mathbf{f}(x, \mathbf{u}, \mathbf{u}')) > 0.$$

If $\mathbf{u} \in C^2(I, B)$ *satisfies* (3.4) *and if* $\mathbf{u}(x_1) \geqq 0$ *and* $\mathbf{u}(x_2) \geqq \mathbf{0}$ *for* $x_1 < x_2$; $x_1, x_2 \in I$, *then* $\mathbf{u}(x) \geqq \mathbf{0}$ *for all* $x \in [x_1, x_2]$.

Proof. The proof of this theorem is identical to that of Theorem 2.1 up to inequality (2.5). Instead of (2.5), we have

$$h''(x_0) = \psi(\mathbf{u}''(x_0)) \leqq \psi(-\mathbf{f}(x_0, \mathbf{u}(x_0), \mathbf{u}'(x_0))) < 0$$

because $\psi(\mathbf{u}(x_0)) > 0$. This is again a contradiction to the assumption that Φ has a negative minimum at x_0. □

4. Comparison theorems for solutions of differential inequalities. As an application of the invariance principle developed in the preceding sections, we give some comparison results for solutions of differential inequalities in a Banach space B, in which there is an inequality relation generated by a cone K. In the result presented here we will use Theorem 3.1 to determine the conditions on $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')$ which are sufficient to provide a comparison of solutions. Similar results can also be easily obtained using Theorem 2.1 or Theorem 3.2.

Let us assume that K is a solid cone, $\mathbf{u}_0 \in \operatorname{int} K$, and that $\mathscr{H}_{\mathbf{u}_0}$ is the collection of positive linear functionals defined by (3.1).

THEOREM 4.1. *Let* $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')$ *satisfy* (3.2) *relative to* $\mathscr{H}_{\mathbf{u}_0}$ *and satisfy a Lipschitz condition in* $\mathbf{u}'$ *on closed bounded subsets of* $I \times B \times B$. *If* $\mathbf{u}, \mathbf{v} \in C^2(I, B)$ *and satisfy*

$$\mathbf{u}'' + \mathbf{f}(x, \mathbf{u}, \mathbf{u}') \geqq \mathbf{v}'' + \mathbf{f}(x, \mathbf{v}, \mathbf{v}'), \qquad x \in I, \tag{4.1}$$

$$\mathbf{u}(x_1) \leqq \mathbf{v}(x_1), \qquad \mathbf{u}(x_2) \leqq \mathbf{v}(x_2) \tag{4.2}$$

for some $x_1, x_2 \in I$, $x_1 < x_2$, *then* $\mathbf{u}(x) \leqq \mathbf{v}(x)$ *for all* $x \in [x_1, x_2]$.

Proof. Define a function $\mathbf{f}_1(x, \mathbf{w}, \mathbf{w}')$ by the equality

$$\mathbf{f}_1(x, \mathbf{w}, \mathbf{w}') = \mathbf{u}''(x) - \mathbf{v}''(x) + \mathbf{f}(x, \mathbf{u}(x), \mathbf{u}'(x)) - \mathbf{f}(x, \mathbf{v}(x) - \mathbf{w}, \mathbf{v}'(x) - \mathbf{w}'). \tag{4.3}$$

We will verify that $\mathbf{f}_1(x, \mathbf{w}, \mathbf{w}')$ satisfies the hypotheses of Theorem 3.1.

Suppose $\phi \in \mathscr{H}_{\mathbf{u}_0}$, $\phi(\mathbf{w} - \mathbf{z}) = \inf\{\psi(\mathbf{w} - \mathbf{z}) | \psi \in \mathscr{H}_{\mathbf{u}_0}\} \leqq 0$ and $\phi(\mathbf{w}' - \mathbf{z}') = 0$. Then for each $x \in I$,

$$\phi([\mathbf{v}(x) - \mathbf{z}] - [\mathbf{v}(x) - \mathbf{w}]) = \inf\{\psi([\mathbf{v}(x) - \mathbf{z}] - [\mathbf{v}(x) - \mathbf{w}]) | \psi \in \mathscr{H}_{\mathbf{u}_0}\} \leqq 0$$

and further $\phi([\mathbf{v}'(x) - \mathbf{z}'] - [\mathbf{v}'(x) - \mathbf{w}']) = 0$. Using hypothesis (3.2) we obtain that

$$\phi(\mathbf{f}(x, \mathbf{v}(x) - \mathbf{z}, \mathbf{v}'(x) - \mathbf{z}') - \mathbf{f}(x, \mathbf{v}(x) - \mathbf{w}, \mathbf{v}'(x) - \mathbf{w}')) \geqq 0.$$

However,

$$\mathbf{f}(x, \mathbf{v}(x) - \mathbf{z}, \mathbf{v}'(x) - \mathbf{z}') - \mathbf{f}(x, \mathbf{v}(x) - \mathbf{w}, \mathbf{v}'(x) - \mathbf{w}') = \mathbf{f}_1(x, \mathbf{w}, \mathbf{w}') - \mathbf{f}_1(x, \mathbf{z}, \mathbf{z}');$$

therefore $\phi(\mathbf{f}_1(x, \mathbf{w}, \mathbf{w}')) \geqq \phi(\mathbf{f}_1(x, \mathbf{z}, \mathbf{z}'))$ so that $\mathbf{f}_1(x, \mathbf{w}, \mathbf{w}')$ satisfies (3.2). Further,

$$\mathbf{f}_1(x, \mathbf{0}, \mathbf{0}) = \mathbf{u}''(x) - \mathbf{v}''(x) + \mathbf{f}(x, \mathbf{u}(x), \mathbf{u}'(x)) - \mathbf{f}(x, \mathbf{v}(x), \mathbf{v}'(x)) \geqq 0$$

so that $\mathbf{f}_1(x, \mathbf{w}, \mathbf{w}')$ satisfies (3.3).

Finally, let D be a bounded subset of $I \times B \times B$ and let L be the Lipschitz constant for $f(x, \mathbf{u}, \mathbf{u}')$ with respect to the bounded set $D' = \{(x, \mathbf{u}(x) - \mathbf{y}, \mathbf{u}'(x) - \mathbf{z}) | (x, \mathbf{y}, \mathbf{z}) \in D\}$. Then, for $(x, \mathbf{y}, \mathbf{z}), (x, \mathbf{y}, \bar{\mathbf{z}}) \in D$,

$$|\mathbf{f}_1(x, \mathbf{y}, \mathbf{z}) - \mathbf{f}_1(x, \mathbf{y}, \bar{\mathbf{z}})| \leqq L|\mathbf{z} - \bar{\mathbf{z}}|,$$

which shows that $\mathbf{f}_1(x, \mathbf{w}, \mathbf{w}')$ satisfies a Lipschitz condition in $\mathbf{w}'$ on bounded subsets of $I \times B \times B$. We have now shown that $\mathbf{f}_1(x, \mathbf{w}, \mathbf{w}')$ satisfies the hypotheses of Theorem 3.1.

Let $x_1, x_2 \in I$, with $x_1 < x_2$, and suppose that $\mathbf{u}(x)$ and $\mathbf{v}(x)$ satisfy (4.2) at x_1 and x_2. Define $\mathbf{w}(x) = \mathbf{v}(x) - \mathbf{u}(x)$. Inequalities (4.2) imply that $\mathbf{w}(x_1) \geqq \mathbf{0}$ and $\mathbf{w}(x_2) \geqq \mathbf{0}$. Furthermore,

$$\mathbf{w}''(x) - \mathbf{f}_1(x, \mathbf{w}(x), \mathbf{w}'(x)) = \mathbf{v}''(x) - \mathbf{u}''(x) + [\mathbf{u}''(x) - \mathbf{v}''(x) + \mathbf{f}(x, \mathbf{u}(x), \mathbf{u}'(x)) - \mathbf{f}(x, \mathbf{u}(x), \mathbf{u}'(x))] = 0.$$

Therefore $\mathbf{w}$ satisfies the hypotheses of Theorem 3.1 and it follows that $\mathbf{w}(x) \geqq 0$, i.e., $\mathbf{v}(x) \geqq \mathbf{u}(x)$, for all $x \in [x_1, x_2]$. □

As an easy consequence of this comparison theorem, it is possible to derive upper and lower bounds on solutions to boundary value problems in B using solutions to differential inequalities in terms of the partial ordering induced by K.

COROLLARY 4.2. *Let K and $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')$ satisfy the hypotheses of Theorem* 4.1 *and suppose $\mathbf{v}_1, \mathbf{v}_2 \in C^2(I, B)$ and satisfy $\mathbf{v}_1'' + \mathbf{f}(x, \mathbf{v}_1, \mathbf{v}_1') \geqq \mathbf{0}$, $\mathbf{v}_2'' + \mathbf{f}(x, \mathbf{v}_2, \mathbf{v}_2') \leqq \mathbf{0}$, $\mathbf{v}_1(a) \leqq \mathbf{a}_0 \leqq \mathbf{v}_2(a)$, $\mathbf{v}_1(b) \leqq \mathbf{a}_1 \leqq \mathbf{v}_2(b)$ where $\mathbf{u}_0, \mathbf{u}_1 \in B$. If $\mathbf{u}$ is a solution to the boundary value problem*

$$\mathbf{u}'' + \mathbf{f}(x, \mathbf{u}, \mathbf{u}') = \mathbf{0}, \tag{4.4}$$

$$\mathbf{u}(a) = \mathbf{a}_0, \qquad \mathbf{u}(b) = \mathbf{a}_1, \tag{4.5}$$

then the upper and lower bounds $\mathbf{v}_1(x) \leqq \mathbf{u}(x) \leqq \mathbf{v}_2(x)$ hold for all $x \in I$; in particular, solutions to (4.4), (4.5) *are unique when they exist.*

Proof. We proceed by showing first that $\mathbf{u}(x) \leqq \mathbf{v}_2(x)$ and then $\mathbf{v}_1(x) \leqq \mathbf{u}(x)$. Both inequalities are immediate consequences of Theorem 4.1.

The uniqueness assertion follows from the part of the definition of a cone in a Banach space which states that not both $\mathbf{x}$ and $-\mathbf{x}$ are elements of K unless $\mathbf{x} = \mathbf{0}$. From this it follows that if $\mathbf{x}$ and $\mathbf{y}$ in B satisfy the string of inequalities $\mathbf{y} \leqq \mathbf{x} \leqq \mathbf{y}$, then $\mathbf{x} = \mathbf{y}$. Let $\mathbf{u}_1(x)$ and $\mathbf{u}_2(x)$ be two solutions of (4.4), (4.5). Setting $\mathbf{u}_1(x) = \mathbf{u}(x)$ and $\mathbf{u}_2(x) = \mathbf{v}_1(x) = \mathbf{v}_2(x)$, we obtain from the first part of this corollary that $\mathbf{u}_2(x) \leqq \mathbf{u}_1(x) \leqq \mathbf{u}_2(x)$. □

5. Remarks and examples.

Example 5.1. Let $B = R^n$ and let $\mathbf{u} \in R^n$ be denoted by $\mathbf{u} = (u_1, \cdots, u_n)^T$. Let $\mathbf{f}: I \times R^n \times R^n \to R^n$ satisfy for $x \in I$,

$$u_k - v_k = \min\{u_j - v_j : j = 1, \cdots, n\} \leqq 0, \quad u_j' = v_j' \Rightarrow f_k(x, u, u') \geqq f_k(x, v, v'), \tag{5.1}$$

and further let $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')$ satisfy a Lipschitz condition in $\mathbf{u}'$ on bounded subsets of $I \times R^n \times R^n$. If we take K to be the cone $K = \{\mathbf{u} \in R^n | u_k \geqq 0, k = 1, \cdots, n\}$, then $\mathbf{f}$ satisfies the hypotheses of Theorem 4.1 with respect to the cone K.

In order to see this, let $\mathscr{K} = \{\phi_1, \cdots, \phi_n\}$ be the family of linear functionals, defined for $\mathbf{u} \in R^n$ by $\phi_k(\mathbf{u}) = u_k$, $k = 1, \cdots, n$. $\mathscr{K}$ generates the cone K. Furthermore, if $\mathbf{u}^0 \in R^n$ is defined by $\mathbf{u}_k^0 = 1$, $k = 1, \cdots, n$, then $\mathscr{K}_{\mathbf{u}_0} = \mathscr{K}$. It now follows easily from the conditions placed on $\mathbf{f}$ that the hypotheses of Theorem 4.1 are satisfied.

This example appears in the case $n = 1$ in the work of Jackson (cf. [3]) and in the case $n \geqq 1$ in the work of Heimes (cf. [2]). In these papers, the result plays an important role in studies of existence theory for boundary value problems based on subfunction techniques. Condition (5.1) is due to Heimes and it was this monotonicity condition on $\mathbf{f}$ which suggested the one in the current paper.

For linear expressions, $\mathbf{f}(x, \mathbf{u}, \mathbf{u}') = \mathscr{A}\mathbf{u}' \to \mathscr{B}\mathbf{u}$, where $\mathscr{A}$ and $\mathscr{B}$ are $n \times n$-matrix functions, (5.1) implies that $\mathscr{A}$ is a diagonal matrix and that $\mathscr{B}$ has nonnegative nondiagonal elements and nonpositive row sums. Heimes has shown further, by constructing an example, that if either $\mathscr{A}$ has nonzero nondiagonal elements or $\mathscr{B}$ has negative nondiagonal entries at some point $x_0 \in (a, b)$, then there exist points $x_1, x_2 \in [a, b]$, with $x_1 < x_0 < x_2$, and a function $\mathbf{u}(x)$ such that

$\mathbf{u}''+A\mathbf{u}'+B\mathbf{u}\leqq\mathbf{0}$, $\mathbf{u}(x_1)\geqq\mathbf{0}$, $\mathbf{u}(x_2)\geqq\mathbf{0}$, but $\mathbf{u}(x)\ngeqq\mathbf{0}$ for all $x\in(x_1, x_2)$; i.e., the conclusion of Theorem 4.1 is false for such equations.

Using the same cone K as above, but changing the set of functionals $\mathscr{K}$ used to generate K, it is possible to extend this result to a larger class of equations. We illustrate this in the following example:

Example 5.2. Let B and K be the same as in Example 5.1 and let α_j, $j=1, \cdots, n$, be positive numbers. Define the set of positive linear functionals $\mathscr{K}$ by $\mathscr{K}=\{\phi_k\colon \phi_k(\mathbf{u})=\alpha_k u_k, k=1, \cdots, 2\}$. If we set $\mathbf{u}^0=(\alpha_1^{-1}, \cdots, \alpha_n^{-1})$, then $\mathscr{K}$ generates K and $\mathscr{K}_{\mathbf{u}_0}=\mathscr{K}$. In order to get a condition on $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')$ which will satisfy the hypotheses of Theorem 4.1, we modify (5.1) as follows:

$$\begin{aligned}&\alpha_k(u_k-v_k)=\min\{\alpha_j(u_j-v_j)\colon j=1, \cdots, n\}\leqq 0, \qquad u_k'=v_k'\\ &\Rightarrow f_k(x, \mathbf{u}, \mathbf{u}')\geqq f_k(x, \mathbf{v}, \mathbf{v}'), \qquad x\in I.\end{aligned} \tag{5.2}$$

If $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')$ satisfies (5.2) and also satisfies a Lipschitz condition in u' on closed and bounded subsets of $I\times R^n\times R^n$, then $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')$ satisfies the hypotheses of Theorem 4.1 with respect to K.

This example can be compared with Example 5.1 by considering the linear case $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')=\mathscr{A}\mathbf{u}'+\mathscr{B}\mathbf{u}$ where $\mathscr{A}$ and $\mathscr{B}$ are $n\times n$ matrices. Condition (5.2) implies that A is again a diagonal matrix and that B is a matrix with nonnegative nondiagonal elements for which the weighted row sums $\sum_{k=1}^n \alpha_k^{-1}\tilde{b}_{jk}$ are nonpositive.

If the partial ordering is generated by a different cone, then it is possible tc obtained comparison results for equations not covered under the conditions in Examples 5.1 or 5.2. We will illustrate this by considering two simple examples, the first in R^n and the second in R^3.

Example 5.3. Let $B=R^n$ and let a partial ordering be induced by a nonsingular $n\times n$ *matrix* C in the following way: $\mathbf{u}, \mathbf{v}\in R^n$, $\mathbf{u}\leqq\mathbf{v}\Leftrightarrow(C\mathbf{u})_j\leqq(C\mathbf{v})_j$ for each $j=1, \cdots, n$, where $(Cw)_j$ denotes the jth component of the vector $C\mathbf{w}$ (see Werner [18]). Let $\mathbf{c}^j$ be a vector whose components are the entries in the jth row of the matrix C. Then the cone K defined by this partial ordering is generated by the set of functionals $\mathscr{K}=\{\phi_j|\phi_j(u)=\mathbf{c}^j\circ u, u\in R^n, j=1, \cdots, n\}$ where $\mathbf{v}\circ\mathbf{w}$ denotes the dot product in R^n. If $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')$ satisfies (3.2) with respect to K, then there is a change of basis $y\in R^n\to\bar{\mathbf{y}}\in R^n$ such that in terms of this new basis $\bar{\mathbf{f}}(x, \bar{\mathbf{u}}, \bar{\mathbf{u}}')$ satisfies (5.1). To see this suppose that $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')$ satisfies (3.2). For $\mathbf{u}^0$ we may take the vector $\mathbf{u}^0=C^{-1}\mathbf{e}$ where $e_j=1, j=1, \cdots, n$. Let $\mathbf{p}^j=C^{-1}\mathbf{e}^j$ where $e_i^j=\delta_{ij}$; then $\mathbf{y}\in R^n$ can be written in the form $\bar{\mathbf{y}}=\sum_{j=1}^n(\mathbf{c}^j\circ\mathbf{y})p^j$ and $\bar{y}_j=(\mathbf{c}^j\circ\mathbf{y})$. That $\bar{\mathbf{f}}(x, \bar{\mathbf{u}}, \bar{\mathbf{u}}')$ satisfies (5.1) now follows immediately.

In the case where $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')=\mathscr{A}\mathbf{u}'+\mathscr{B}\mathbf{u}$, (3.2) is equivalent to having the two matrices $C\mathscr{A}C^{-1}$ and $C\mathscr{B}C^{-1}$ satisfy the hypothesis on $\mathscr{A}$ and $\mathscr{B}$ mentioned in Example 5.1.

As an illustration of this property consider the linear expression

$$\mathbf{f}(x, \mathbf{u}, \mathbf{u}')=\begin{bmatrix}1 & -2\\ 0 & -1\end{bmatrix}\mathbf{u}'+\begin{bmatrix}-1 & -1\\ 2 & -4\end{bmatrix}\mathbf{u}. \tag{5.3}$$

If C is the matrix

$$C=\begin{bmatrix}0 & 1\\ 1 & -1\end{bmatrix}$$

and K is the cone generated by the rows of C, then $\mathbf{f}$ satisfies (3.2) with respect to K.

Example 5.4 (Volkmann [15]). Let $B=R^3$, $K=\{\mathbf{u}\in R^3: \sqrt{u_1^2+u_2^2}\leqq u_3\}$, and $\mathbf{u}^0=(0, 0, 2^{1/2})^T$. Let $S=\{\boldsymbol{\sigma}\in R^2: \sigma_1^2+\sigma_2^2=1\}$ and define $\phi_{\boldsymbol{\sigma}}\in B^*$ by the equation $\phi_{\boldsymbol{\sigma}}(\mathbf{u})=2^{-1/2}(-\sigma_1u_1-\sigma_2u_2+u_3)$. The set of functionals $\mathscr{K}=\{\phi_{\boldsymbol{\sigma}}: \boldsymbol{\sigma}\in S\}$ generates K, $\mathscr{K}_{\mathbf{u}^0}=\mathscr{K}$, and $\bar{\mathscr{K}}=\mathscr{K}$. Consider the linear function

$$\mathscr{L}(\mathbf{u})=\begin{bmatrix}0 & -1 & 0\\ 1 & 0 & 0\\ 0 & 0 & -1\end{bmatrix}\mathbf{u}=\begin{bmatrix}-u_2\\ u_1\\ -u_3\end{bmatrix}.$$

We will show that $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')=\mathscr{L}(\mathbf{u})$ satisfies (4.2) with respect to K. Suppose $\phi_{\boldsymbol{\tau}}\in\mathscr{K}$, $\phi_{\boldsymbol{\tau}}(\mathbf{w})=\inf\{\phi_{\boldsymbol{\sigma}}(w): \boldsymbol{\sigma}\in S\}\leqq 0$ where $\mathbf{w}=\mathbf{u}-\mathbf{v}$; then

$$\phi_{\boldsymbol{\tau}}(\mathscr{L}(\mathbf{w}))=2^{-1/2}(\tau_1w_2-\tau_2w_1-w_3)=-\phi_{\boldsymbol{\tau}'}(\mathbf{w})$$

where $\boldsymbol{\tau}'=(-\tau_2, \tau_1)$. But $\boldsymbol{\tau}'\in S$; hence

$$\phi_{\boldsymbol{\tau}}(\mathscr{L}(\mathbf{w}))=-\phi_{\boldsymbol{\tau}'}(\mathbf{w})\geqq-\phi_{\boldsymbol{\tau}}(\mathbf{w})\geqq 0,$$

that is, (4.2) is satisfied.

The last two examples we consider deal with differential inequalities in infinite dimensional spaces.

Example 5.5. Let $B=l^\infty(Z, R)$, where Z is the set of integers, and let K be the cone $K=\{\mathbf{u}\in B|u_i\geqq 0, i\in Z\}$. K is generated by the set of functionals $\mathscr{K}=\{\phi_i|\phi_i(\mathbf{u})=u_i, i\in Z\}$. Setting $\mathbf{u}^0$ to be the element of B with $u_i^0=1, i\in Z$, we obtain $\mathbf{u}^0\in\operatorname{int}K$ and $\mathscr{K}_{\mathbf{u}^0}=\mathscr{K}$.

Suppose $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')$ satisfies the following conditions:

$$\text{(5.4)}\qquad u_k-v_k=\inf\{u_j-v_j: j\in Z\}\leqq 0,\qquad u_k'=v_k'$$
$$\Rightarrow f_k(x, \mathbf{u}, \mathbf{u}')\geqq f_k(x, \mathbf{v}, \mathbf{v}'),\qquad x\in I;$$

for every closed bounded subset $M\subseteq I\times B\times B$ there exists a continuous function $d: R\to[0, \infty)$, with $d(s)$ increasing for $s\geqq 0$, $d(0)=0$, and $d(-s)=d(s)$, such that for every $k\in Z$,

$$\text{(5.5)}\qquad f_k(x, \mathbf{u}, \mathbf{u}')-f_k(x, \mathbf{u}-s\mathbf{e}^k, \mathbf{u}')\geqq -d(s),\qquad (x, \mathbf{u}, \mathbf{u})\in M,\qquad s\geqq 0,$$

where $\mathbf{e}^k\in B$ is defined by $e_i^k=\delta_{ik}$, $i\in Z$; and $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')$ satisfies a Lipschitz condition in $\mathbf{u}'$ on closed and bounded subsets of $I\times B\times B$.

We will show that $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')$ satisfies the hypotheses of Theorem 4.2. Suppose $\phi\in\bar{\mathscr{K}}_{\mathbf{u}^0}$ and $\phi(\mathbf{u}-\mathbf{v})=\inf\{\phi_k(\mathbf{u}-\mathbf{v}): k\in Z\}\leqq 0$, $\phi(\mathbf{u}'-\mathbf{v}')=0$.

$\phi(\mathbf{f}(x, \mathbf{u}, \mathbf{u}') - \mathbf{f}(x, \mathbf{v}, \mathbf{v}'))$ can be rewritten

$$
\begin{aligned}
\phi(\mathbf{f}(x, \mathbf{u}, \mathbf{u}') - \mathbf{f}(x, \mathbf{v}, \mathbf{v}')) &= (\phi - \phi_k)(\mathbf{f}(x, \mathbf{u}, \mathbf{u}') - \mathbf{f}(x, \mathbf{v}, \mathbf{v}')) \\
&\quad + \{f_k(x, \mathbf{u}, \mathbf{u}') - f_k(x, \mathbf{u}, \mathbf{u}' + t\mathbf{e}^k)\} \\
&\quad + \{f_k(x, \mathbf{u}, \mathbf{u}' + t\mathbf{e}^k) - f_k(x, \mathbf{u} - s\mathbf{e}^k, \mathbf{u}' + t\mathbf{e}^k)\} \\
&\quad + \{f(x, \mathbf{u} - s\mathbf{e}^k, \mathbf{u}' + t\mathbf{e}^k) - f_k(x, \mathbf{u}, \mathbf{u}')\}.
\end{aligned} \tag{5.6}
$$

Let $\varepsilon > 0$ be given, let L be the Lipschitz constant for $\mathbf{f}(x, \mathbf{u}, \mathbf{u}')$ relative to the closed, bounded set $\{(x, \mathbf{y}, \mathbf{z}) | x \in I,\ |\mathbf{y}| \leqq |\mathbf{u}|,\ |\mathbf{z}| \leqq |\mathbf{u}'| + 1\}$, and let $d(s)$ be the function in (5.5) relative to the set $M = \{(x, \mathbf{y}, \mathbf{z}) | x \in I,\ |\mathbf{y}| \leqq |\mathbf{u}| + 1,\ |\mathbf{z}| \leqq |\mathbf{u}'| + 1\}$. Choose s, $0 < s < 1$, so small that $d(s) < \varepsilon/3$. Using the fact that $\phi \in \bar{\mathscr{K}}_{\mathbf{u}^0}$, let $k \in Z$ be chosen such that: $(\phi - \phi_k)(\mathbf{f}(x, \mathbf{u}, \mathbf{u}') - \mathbf{f}(x, \mathbf{v}, \mathbf{v}')) \geqq -\varepsilon/3$, $(u_k - s) - v_k \leqq \phi(\mathbf{u} - \mathbf{v})$, and $|t| \leqq \min\{\varepsilon/3L, 1\}$ where $t = u_k' - v_k'$. It follows that

$$
\phi(\mathbf{f}(x, \mathbf{u}, \mathbf{u}') - \mathbf{f}(x, \mathbf{v}, \mathbf{v}')) \geqq -\frac{\varepsilon}{3} - Lt - d(s) \geqq -\varepsilon
$$

where we have used (5.4) on the last term on the right-hand side of (5.6). Since $\varepsilon > 0$ was arbitrary, it follows that $\phi(\mathbf{f}(x, \mathbf{u}, \mathbf{u}')) \geqq \phi(\mathbf{f}(x, \mathbf{v}, \mathbf{v}'))$.

Equations satisfying conditions (5.4) and (5.5) have arisen in applications of the method of lines to boundary value problems for elliptic partial differential equations in unbounded regions (see, for example, [12]).

In a final example, we give a function satisfying the hypotheses of Theorem 3.2 for the case of a cone with empty interior.

Example 5.6. Let $B = \ell^p(Z, R)$, $1 \leqq p < \infty$. B is the collection of all doubly infinite sequences $\mathbf{u} = \{u_i\}_{i \in Z}$ such that $\sum_{i \in Z} |u_i|^p < \infty$. Let K be the cone in B defined by $K = \{\mathbf{u} | u_i \geqq 0,\ i \in Z\}$. K is generated by the collection of positive linear functionals $\mathscr{K} = \{\phi_i \in B^* : \phi_i(\mathbf{u}) = u_i,\ i \in Z\}$. Let $\{a_i\}_{i \in Z}$ be a bounded sequence of real functions and let $\mathscr{B}$ be an infinite matrix function $\mathscr{B} = (b_{ij})$, $i, j \in Z$, with $b_{ij} \geqq 0$ if $i \neq j \in I$. Let $\bar{a}$ be such that $|a_i| \leqq \bar{a}$, $i \in Z$; and let $\mathscr{B}$ satisfy

$$
\sum_{j \in Z} b_{kj} \leqq \underline{b} < 0 \quad \text{and} \quad \sum_{j \in Z} |b_{kj}| \leqq \bar{b} < \infty
$$

for all $x \in I$, $k \in Z$, where $\underline{b}$ and $\bar{b}$ are constants. Under these conditions, the function $\mathbf{f}(x, \mathbf{u}, \mathbf{u}') = \mathscr{A}(x)\mathbf{u}' + \mathscr{B}(x)\mathbf{u}$ defined for $k \in Z$ by

$$
f_k(x, \mathbf{u}, \mathbf{u}') = a_k(x)u_k' + \sum_{j \in Z} b_{kj}(x)u_j
$$

satisfies (3.7). To see this, let $\phi(\mathbf{w}) = \inf\{w_k | k \in Z\} < 0$ and $\phi(\mathbf{w}') = 0$, where $\phi \in \bar{\mathscr{K}}$ and $\mathbf{w} = \mathbf{u} - \mathbf{v}$. Define $\alpha = \phi(\mathbf{w})$. Because $\phi \in \bar{\mathscr{K}}$ there exists an integer $k \in Z$ such that $|w_k'| < \varepsilon/3\bar{a}$; $|(\phi - \phi_k)(A\mathbf{w}' + B\mathbf{w})| < \varepsilon/3$; and $\bar{b}|\alpha - w_k| < \varepsilon/3$ where

$0<\varepsilon<b\alpha$ (note: both $b<0$ and $\alpha<0$). It follows that

$$\begin{aligned}\phi(A\mathbf{w}'+B\mathbf{w}) &= (\phi-\phi_k)(A\mathbf{w}'+B\mathbf{w})+a_k w_k'+\sum_{j\in Z} b_{kj}w_j \\ &\geqq \frac{-2\varepsilon}{3}+\sum_{\substack{j\in Z\\ j\neq k}} b_{kj}\alpha+b_{kk}u_k \\ &\geqq -\varepsilon+\Big(\sum_{j\in Z} b_{kj}\Big)\alpha \\ &\geqq b\alpha-\varepsilon>0.\end{aligned}$$

This last inequality verifies that (3.7) is satisfied.

As a particular example of this last result, consider the expression

$$f_k(x,\mathbf{u}) = h^{-2}\{u_{k+1}+u_{k-1}-(2+h^2\lambda^2)u_k\} \tag{5.7}$$

where h and λ are constants. This equation arises when the method of lines is applied to the steady state Klein–Gordon equation $-\Delta u+\lambda^2 u=0$ defined on an infinite strip in R^2 (cf. [13]). It is easy so show that (5.7) satisfies all the hypotheses of Example 5.6.

Acknowledgment. The author is grateful to the referees for their helpful comments on Lemma 2.3 and Example 5.3.

REFERENCES

[1] N. DUNFORD AND J. T. SCHWARTZ, *Linear Operators, Part One*, Interscience, New York, 1957.

[2] K. A. HEIMES, *Boundary value problems for ordinary nonlinear second order systems*, J. Differential Equations, 2 (1966), pp. 449–463.

[3] L. K. JACKSON, *Subfunctions and boundary value problems for second order ordinary differential equations*, Advances in Math., 2 (1968), pp. 307–363.

[4] M. A. KRASNOSELSKII, *Positive Solutions of Operator Equations*, P. Noordhoff, Groningen, the Netherlands, 1964.

[5] S. G. KREIN, ed., *Functional Analysis*, Wolters–Noordhoff, Groningen, the Netherlands, 1972.

[6] M. H. PROTTER AND H. F. WEINBURGER, *Maximum Principles in Differential Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1967.

[7] R. M. REDHEFFER AND W. WALTER, *Flow-invariant sets and differential inequalities in normed spaces*, Applicable Anal., to appear.

[8] K. SCHMITT AND P. VOLKMANN, *Boundary value problems for second order differential equations in convex subsets of a Banach space*, Trans. Amer. Math. Soc., 218 (1976), pp. 397–405.

[9] J. SCHRÖDER, *Inverse-positive linear operators—an abstract theory*, Rep. 75-4, Mathematisches Institut, Universität zu Köln, Köln, Germany, 1975.

[10] ———, *Inverse-positive ordinary differential operators of the first and second order*, Rep. 75-5, Mathematisches Institut, Universität zu Köln, Köln, Germany, 1975.

[11] ———, *Upper and lower bounds for solutions of generalized two-point boundary value problems*, Numer. Math., 23 (1975), pp. 433–457.

[12] R. THOMPSON, *Convergence and error estimates for the method of lines for certain nonlinear elliptic and elliptic-parabolic equations*, SIAM J. Numer. Anal., 13 (1976), pp. 27–43.

[13] ———, *Differential inequalities for infinite second order systems and an application to the method of lines*, J. Differential Equations, 17 (1975), pp. 421–434.

[14] P. VOLKMANN, *Gewöhnliche Differential-ungleichungen mit quasimonoton washsenden Funktionen in Banach Räumen*, Ordinary and Partial Differential Equations, Proc. Conf. Univ. Dundee, Dundee, 1975, Lecture Notes in Math., vol. 415, Springer, Berlin, 1974, pp. 439–443.

[15] ———, *Gewöhnliche Differential-ungleichungen mit quasimonoton wachsenden Funktionen in topologischen Vektorraümen*, Math. Z., 127 (1972), pp. 157–164.

[16] ———, *Über die Invarianz konvexer Mengen und Differential-ungleichungen in einem normierten Räume*, Math. Ann., 203 (1973), pp. 201–210.

[17] W. WALTER, *Differential and Integral Inequalities*, Springer, New York, 1970.

[18] J. WERNER, *Einschlieszungssätze bei nichtlinearen gewöhnlichen Randwertaufgaben und erzwungenen Schwingungen*, Numer. Math., 13 (1969), pp. 24–38.

SIAM J. MATH. ANAL.
Vol. 8, No. 4, August 1977

HEAT FLOW INEQUALITIES WITH APPLICATIONS TO HEAT FLOW OPTIMIZATION PROBLEMS*

ANDREW ACKER†

Abstract. Inequalities are derived which compare the rates of heat flow across regions with different boundaries. The inequalities form the basis for an existence and uniqueness theory for several closely related free-boundary optimization problems involving the Laplace equation.

1. Introduction. Minimizing the flow of heat or electricity across a region of specified area by a proper choice of a free boundary is a problem of natural importance in engineering. For example, assume the fluid in an infinite cylindrical pipe is to be maintained at a high constant temperature; see Fig. 1. (Here and elsewhere, the word cylindrical pertains to general cylinders, not necessarily

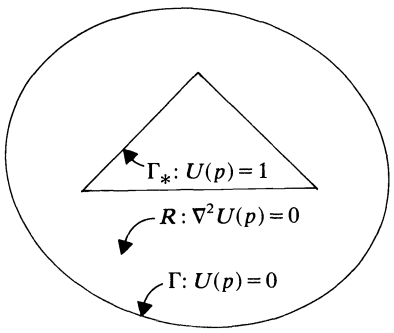


FIG. 1. *The cross-section of a cylindrical pipe. Problem: If Γ_* and the area of R are fixed, what boundary Γ minimizes the heat flow across R?*

circular.) The question arises: if the cylindrical inner boundary and the cross-sectional area of the pipe are specified, how should the cylindrical outer boundary be chosen so that the rate of heat loss (per unit length of pipe) into a cold, constant-temperature environment is minimized? A mathematically equivalent question can be formulated in terms of current leakage from an insulated wire or the capacitance of a generalized coaxial cable.

Despite its engineering interest, very little work has been done on this problem. The only results known to me are those of T. Carleman [4] and G. Szegö [9]. Carleman showed (in terms of Fig. 1) that if the area of R and the area within Γ_* are fixed, then the heat flow across R is minimized when Γ_* and Γ are concentric circles. Szegö's result is the natural three-dimensional generalization. The proofs are based on special properties of the circle and sphere; for example, Carleman uses a series expansion valid only in the annulus, and Szego uses the isoperimetric property. These methods are not applicable to the unsymmetric problem posed above.

* Received by the editors March 27, 1975, and in revised form April 15, 1976.

† Mathematisches Institut I, Universität Karlsruhe (TH), 75 Karlsruhe 1, Federal Republic of Germany.

In this paper, we prove that if the cross-section of the interior of the pipe is starlike, then there exists a unique heat-flow-minimizing outer boundary. The optimal outer boundary of the pipe is characterized as a boundary of constant heat flow density.

The proof of this result is based on a general heat flow inequality (Theorem 1) which compares the heat flow across any fixed boundary with the heat flow across any member of an appropriate monotone, continuously varying family of boundaries. Theorem 1 essentially reduces the minimum problem to the results of A. Beurling [2] and D. E. Tepper [10], [11]. What we show is that if the family of outer boundaries having constant heat flow density is elliptic in the sense of Beurling, then each of these boundaries is heat-flow minimizing at its respective area.

When applied in conjunction with the Lindelöf principle, Theorem 1 also leads to interesting heat-flow inequalities for certain classes of regions. Further results include existence theorems for a dual heat-flow maximization problem and for a problem involving area minimization on classes of conformally equivalent regions.

In the three-dimensional generalization of the pipe problem, the hot fluid occupies a finite, irregularly shaped cavity and one seeks the form of the insulation layer (or specified volume) about the cavity which minimizes the heat loss into a cold fluid bath. Although details are not given here, our inequalities can be easily generalized to three dimensions. These inequalities again reduce the minimization problem to the existence problem for surfaces of constant heat-flow density. However, the solution to the latter problem is not at present available, since Beurling's methods are restricted to two dimensions.

The problem treated in detail here is just one example of the many free boundary optimization problems that occur naturally in an engineering context. Such problems include the determination of the shape of a cantilever of specified length and weight which supports a given static load distribution with minimum deflection of the free end, or the determination of the wing profile of specified cross-sectional area which maximizes the ratio of lift to drag under given conditions, or the optimal geometry of a cooling fin. Three papers from the extensive literature are [3], [5] and [7]. To the author's knowledge, few if any of these naturally occurring engineering problems have been solved by an exact analysis under realistic assumptions. It is hoped that the methods used here may provide a first step in this direction.

2. Temperature problems. The following two temperature problems are the setting for the results in §§ 3–5.

Problem 1. (See Fig. 2.) A fixed continuous function $a(p)>0$ is defined on R^2. Let $\{S_*, S, R(\Gamma_*, \Gamma)\}$ represent any partition of R^2 with the following properties: S_* and S are disjoint closed sets with respective boundaries Γ_* and Γ. The complement of $S_* \cup S$ in R^2 is the bounded region $R(\Gamma_*, \Gamma)$. Finally, $\Gamma_* \cup \Gamma$ is of measure zero and each point on $\Gamma_* \cup \Gamma$ is the endpoint of some arbitrarily short line segment which (except for its endpoints) lies entirely in the interior of $S_* \cup S$. The weighted area $A(\Gamma_*, \Gamma)$ of $R(\Gamma_*, \Gamma)$ is defined as the integral of $a^2(p)$ over $R(\Gamma_*, \Gamma)$. By Perron's method (Ahlfors [1, pp. 237–243]), there exists a unique

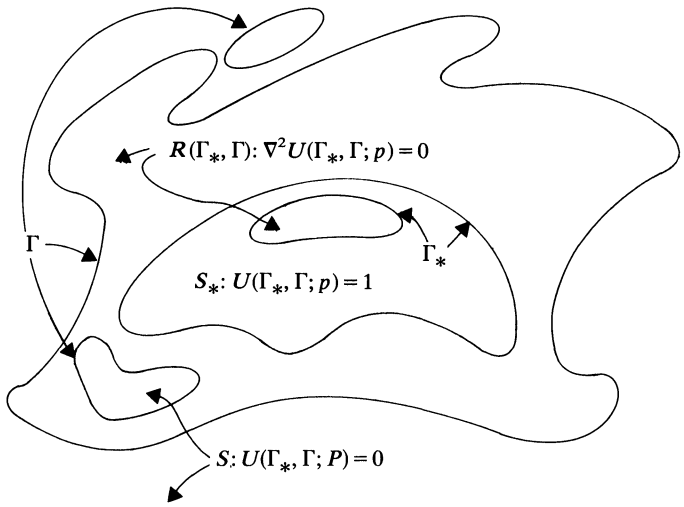


FIG. 2. *Problem* 1

continuous function $U(\Gamma_*, \Gamma; p)$ defined on R^2 such that $U(\Gamma_*, \Gamma; p) = 1$ on S_*, $U(\Gamma_*, \Gamma; p) = 0$ on S, and $U(\Gamma_*, \Gamma; p)$ is harmonic on $R(\Gamma_*, \Gamma)$. $U(\Gamma_*, \Gamma; p)$ represents the steady-state temperature in R^2 when $R(\Gamma_*, \Gamma)$ is occupied by a homogeneous, partially thermally insulating substance and S_* and S are perfectly conducting and held at temperatures 1 and 0. $H(\Gamma_*, \Gamma)$ represents the rate of steady-state heat flow across $R(\Gamma_*, \Gamma)$ from S_* into S. It is sufficient for our purposes to define $H(\Gamma_*, \Gamma)$ in the case where S_* is a finite union of bounded, simply connected sets. In this case, let γ be a finite union of positively oriented, smooth closed curves in $R(\Gamma_*, \Gamma)$ such that the collecting winding number is 1 about each point in S_* and 0 about each point in S. Then $H(\Gamma_*, \Gamma) = \int_\gamma \mathbf{D}_n U(\Gamma_*, \Gamma; p) \cdot |dp|$, where at each $p \in \gamma$, $\mathbf{D}_n U(\Gamma_*, \Gamma; p)$ is the derivative of $U(\Gamma_*, \Gamma; p)$ in the direction normal to γ at p and toward S_*.

Problem 2. (See Fig. 3.) In this case, the positive, continuous function $a(p)$ is defined on $[0, 1] \times R$ and $\{S_*, S, R(\Gamma_*, \Gamma)\}$ is any partition of $[0, 1] \times R$ in which S_* and S are closed, Γ_* and Γ are the boundaries of S_* and S relative to $[0, 1] \times R$, $R(\Gamma_*, \Gamma)$ is bounded and relatively open, $\Gamma_* \cup \Gamma$ has measure 0, and each point of $\Gamma_* \cup \Gamma$ is the endpoint of a line segment lying in the interior of $S_* \cup S$. A unique continuous temperature function $U(\Gamma_*, \Gamma; p)$ exists on $[0, 1] \times R$ such that $U(\Gamma_*, \Gamma; p) = 1$ on S_*, $U(\Gamma_*, \Gamma; p) = 0$ on S, $U(\Gamma_*, \Gamma; p)$ is harmonic in the interior of $R(\Gamma_*, \Gamma)$, and $\mathbf{D}_x U(\Gamma_*, \Gamma; p) = 0$ at all points $p = (x, y) \in R(\Gamma_*, \Gamma)$ for which $x = 0$ or $x = 1$. The definitions of $A(\Gamma_*, \Gamma)$ and $H(\Gamma_*, \Gamma)$ are analogous.

3. A heat flow inequality principle. This section is devoted to the lengthy proof of the following Theorem 1. The remainder of the paper consists mainly of applications of Theorem 1.

THEOREM 1. *In Problem* 1, *let* $\{S_*, S_\alpha, R(\Gamma_*, \Gamma_\alpha)\}$ *be a partition of* R^2 *for each* α *in an open real interval I. Define* $B(I) = \cup_{\alpha \in I} \Gamma_\alpha$. *Assume the following*:

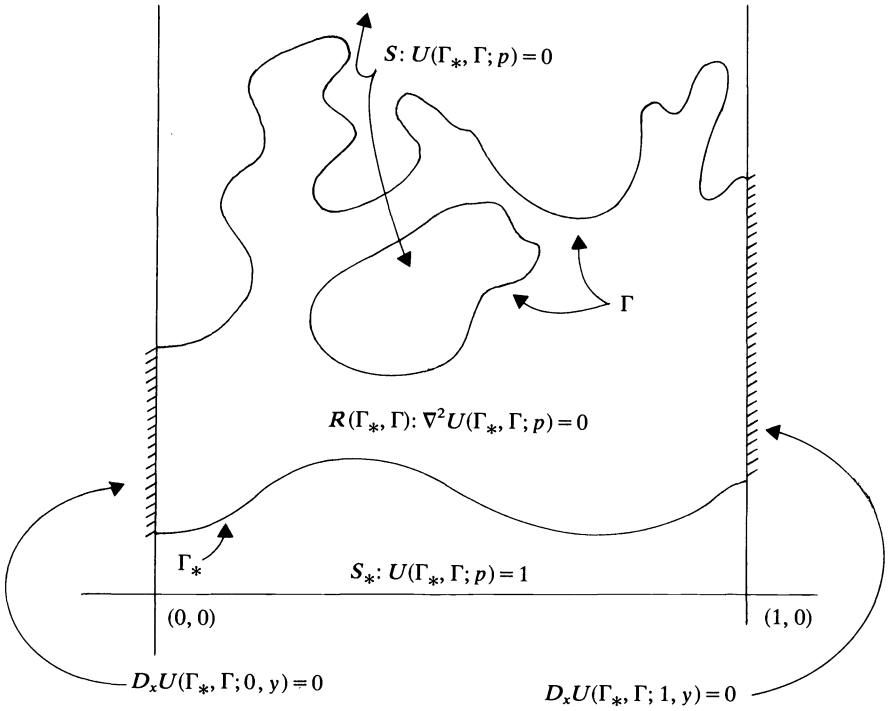


FIG. 3. *Problem* 2

(a) $R(\Gamma_*, \Gamma_\alpha) \supset R(\Gamma_*, \Gamma_\beta)$ *for all* $\alpha, \beta \in I$ *with* $\alpha \leqq \beta$. *Further, if* $\alpha \neq \beta$, *then* $\Gamma_\alpha \cap \Gamma_\beta = \varnothing$.

(b) *For any* $\alpha \in I$, $\lim_{\beta \to \alpha} \Gamma_\beta = \Gamma_\alpha$ *in the sense of set convergence.*

(c) *For each* $\alpha \in I$, Γ_α *is a finite disjoint union of smooth closed Jordan curves. Further, there is a Lipschitz continuous unit vector function* $n(p)$ *defined on* $B(I)$ *such that if* $p \in \Gamma_\alpha$, *then* $n(p)$ *is perpendicular to* Γ_α *at* p.

(d) *There exists a continuous function* $\phi(p)$ *defined on* $B(I)$ *such that for any* $p \in B(I)$, $\lim_{p' \to p} |\nabla U(\Gamma_*, \Gamma_\alpha; p')| = \phi(p)$, *where* Γ_α *is the unique boundary containing* p, *and* $p' \in R(\Gamma_*, \Gamma_\alpha)$.

Let $\{S_*, \bar{S}, R(\Gamma_*, \bar{\Gamma})\}$ *be any other partition of* R^2 *with the same set* S_* *and with* $\bar{\Gamma} \subset B(I)$. *Then for any* $c \in I$: $H(\Gamma_*, \bar{\Gamma}) \geqq H(\Gamma_*, \Gamma_c) + \int_{R_1} \phi^2(p)\, dx\, dy - \int_{R_2} \phi^2(p)\, dx\, dy$, *where* $R_1 = R(\Gamma_*, \Gamma_c) \cap \bar{S}$ *and* $R_2 = R(\Gamma_*, \bar{\Gamma}) \cap S_c$.

Proof. The theorem will first be proven under the following additional assumptions. 1. $R(\Gamma_*, \Gamma_c) \subset R(\Gamma_*, \bar{\Gamma})$. 2. Γ_* and $\bar{\Gamma}$ are each a finite disjoint union of Lipschitz continuously differentiable (L.c.d.) closed Jordan curves. 3. For any $\alpha \in I$, $\bar{\Gamma} \cap \Gamma_\alpha$ contains at most a finite number of points.

Let $\underline{c} = \sup\{\alpha \in I | R(\Gamma_*, \bar{\Gamma}) \subset R(\Gamma_*, \Gamma_\alpha)\}$. For each $\alpha \in [\underline{c}, c] \subset I$, define $R(\Gamma_*, \bar{\Gamma}_\alpha) = R(\Gamma_*, \Gamma_\alpha) \cap R(\Gamma_*, \bar{\Gamma})$ and $\bar{S}_\alpha = S_\alpha \cup \bar{S}$, and let $\bar{\Gamma}_\alpha$ be the boundary of $\bar{S}_\alpha$. Also define $\gamma_\alpha = \Gamma_\alpha \cap R(\Gamma_*, \bar{\Gamma})$ and let $\partial\gamma_\alpha = \bar{\Gamma} \cap \Gamma_\alpha$ be the boundary of γ_α relative to Γ_α. (These definitions are illustrated in Fig. 4.) For any constant $\rho > 0$,

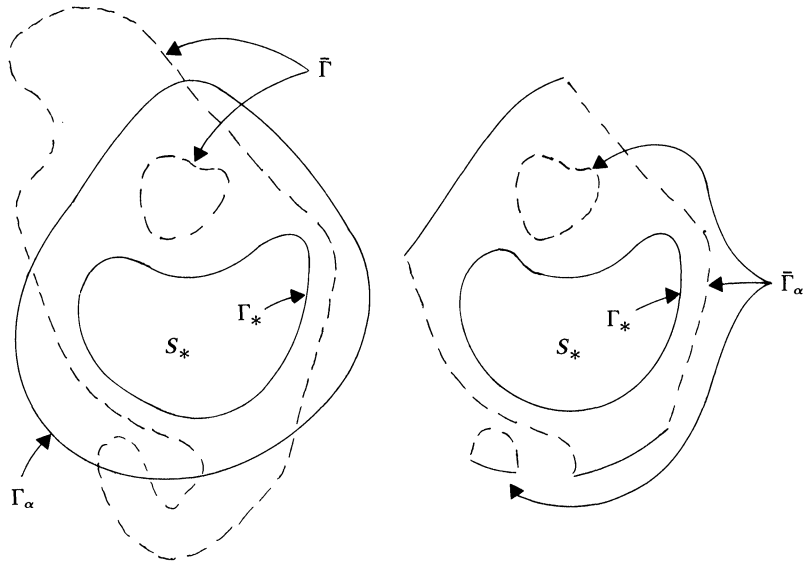


FIG. 4. *In the right hand diagram, the solid boundary curve (other than* Γ_**) is* γ_α.

let the function $\lambda_\rho(\alpha)$ be defined on $[\underline{c}, c]$ by $\lambda_\rho(\alpha) = H(\Gamma_*, \bar{\Gamma}_\alpha) - (1+\rho)\int_{R_\alpha} \phi^2(p)\, dx\, dy$, where $R_\alpha = R(\Gamma_*, \bar{\Gamma}_\alpha) \cap S_c$. The desired inequality (in the case where $R(\Gamma_*, \Gamma_c) \subset R(\Gamma_*, \bar{\Gamma})$) will follow if it is shown that $\lambda_\rho(\underline{c}) \geqq \lambda_\rho(c)$ for all $\rho > 0$. We will show (for any $\rho > 0$) that $\lambda_\rho(\underline{c}) < \lambda_\rho(c)$ leads to a contradiction.

We first note that $\lambda_\rho(\alpha)$ is continuous on $[\underline{c}, c]$. In fact if $\underline{c} \leqq \alpha \leqq \beta \leqq c$, then it can be shown by applying the maximum principle in the manner to be used in the next paragraph that $0 \leqq H(\Gamma_*, \bar{\Gamma}_\beta) - H(\Gamma_*, \bar{\Gamma}_\alpha) \leqq H(\Gamma_*, \Gamma_\beta) - H(\Gamma_*, \Gamma_\alpha)$, whereas the continuity of $H(\Gamma_*, \Gamma_\alpha)$ as a function of α follows from assumptions (b), (c) and (d). Therefore, if $\lambda_\rho(\underline{c}) < \lambda_\rho(c)$ then essentially by the intermediate value theorem there exists an $\alpha \in (\underline{c}, c)$ such that $\lambda_\rho(\alpha + \delta) \geqq \lambda_\rho(\alpha)$ for all $\delta \in (0, c - \alpha)$. We will show for every $\alpha \in (\underline{c}, c)$ that $\lambda_\rho(\alpha + \delta) < \lambda_\rho(\alpha)$ provided that $\delta > 0$ is sufficiently small.

For $\underline{c} \leqq \alpha \leqq \alpha + \delta \leqq c$ fixed, define $\Delta(\alpha, \delta) = H(\Gamma_*, \bar{\Gamma}_{\alpha+\delta}) - H(\Gamma_*, \bar{\Gamma}_\alpha) \geqq 0$. Let $V(p)$ be the harmonic function on $R(\Gamma_*, \bar{\Gamma}_{\alpha+\delta})$ whose continuous extension to the boundary satisfies $V(p) = U(\Gamma_*, \bar{\Gamma}_\alpha; p)$ on $\bar{\Gamma}_{\alpha+\delta}$ and $V(p) = 0$ on Γ_*. Then $\Delta(\alpha, \delta) = \int_{\Gamma_*} |\nabla V(p)| \cdot |dp|$, where $\nabla V(p)$ has a continuous extension to Γ_* due to assumption 2. Let $W(p)$ be the harmonic function on $R(\Gamma_*, \Gamma_{\alpha+\delta})$ whose continuous extension to the boundary (minus $\partial\gamma_{\alpha+\delta}$) satisfies $W(p) = U(\Gamma_*, \Gamma_\alpha; p)$ on $\gamma_{\alpha+\delta}$ and $W(p) = 0$ on $\Gamma_* \cup \Gamma_{\alpha+\delta} \backslash \gamma_{\alpha+\delta}$. The maximum principle shows that $0 \leqq U(\Gamma_*, \bar{\Gamma}_\alpha; p) \leqq U(\Gamma_*, \Gamma_\alpha; p)$ on $\gamma_{\alpha+\delta}$. Therefore $0 \leqq V(p) \leqq W(p)$ on $\Gamma_* \cup \bar{\Gamma}_{\alpha+\delta} \backslash \partial\gamma_{\alpha+\delta}$, and it follows essentially by the maximum principle that $0 \leqq V(p) \leqq W(p)$ on $R(\Gamma_*, \bar{\Gamma}_{\alpha+\delta})$. Therefore, $|\nabla V(p)| \leqq |\nabla W(p)|$ on Γ_*, and it follows that $0 \leqq \Delta(\alpha, \delta) \leqq \int_{\Gamma_*} |\nabla W(p)| \cdot |dp|$. Green's second theorem, applied to the

functions $W(p)$ and $U(\Gamma_*, \Gamma_{\alpha+\delta}; p)$ on $R(\Gamma_*, \Gamma_{\alpha+\delta})$, implies that:

$$0 \leqq \Delta(\alpha, \delta) \leqq \int_{\Gamma_*} |\nabla W(p)| \cdot |dp| = \int_{\Gamma_*} U(\Gamma_*, \Gamma_{\alpha+\delta}; p) \cdot |\nabla W(p)| \cdot |dp|$$

$$= \int_{\Gamma_{\alpha+\delta}} W(p) \cdot |\nabla U(\Gamma_*, \Gamma_{\alpha+\delta}; p)| \cdot |dp|$$

$$= \int_{\Gamma_{\alpha+\delta}} W(p) \cdot \phi(p) \cdot |dp|$$

$$= \int_{\gamma_{\alpha+\delta}} U(\Gamma_*, \Gamma_\alpha; p) \cdot \phi(p) \cdot |dp|.$$

Let $\alpha \in (\underline{c}, c)$ be fixed and let δ vary in $(0, c-\alpha)$. For each $0 < \delta < c - \alpha$, let $\bar{\delta} = \max\{d(p, \Gamma_\alpha) | p \in \Gamma_{\alpha+\delta}\}$ and $\underline{\delta} = \min\{d(p, \Gamma_\alpha) | p \in \Gamma_{\alpha+\delta}\}$. (Here, $d(p, \Gamma_\alpha) = \min\{|p-q| \, | \, q \in \Gamma_\alpha\}$.) By assumptions (b) and (c), $\bar{\delta} \to 0$ as $\delta \to 0$ and there is a uniform constant $C > 0$ such that $\underline{\delta} > C\bar{\delta}$. For each point $\bar{p} \in \Gamma_{\alpha+\delta}$, let p be the closest point to $\bar{p}$ on Γ_α and let $L(p)$ be the line segment connecting p to $\bar{p}$. By the theorem of the mean, $U(\Gamma_*, \Gamma_\alpha; \bar{p}) = |\nabla U(\Gamma_*, \Gamma_\alpha; p^*) \cdot n(p)| \cdot |\bar{p} - p|$, where $p^* \in L(p)$. Let $\varepsilon(\bar{\delta})$ represent any function for which $\varepsilon(\bar{\delta}) \to 0$ as $\bar{\delta} \to 0$. It follows from assumption (c) that $U(\Gamma_*, \bar{\Gamma}_\alpha; \bar{p}) = \phi(p) \cdot |\bar{p} - p| + \bar{\delta}\varepsilon(\bar{\delta})$. By using assumptions (c), (d), 2, and 3, we obtain

$$\Delta(\alpha, \delta) \leqq \int_{\gamma_{\alpha+\delta}} U(\Gamma_*, \Gamma_\alpha; \bar{p}) \cdot \phi(\bar{p}) \cdot |d\bar{p}| = \int_{\gamma_\alpha} \phi^2(p) \cdot |\bar{p} - p| \cdot |dp| + \bar{\delta}\varepsilon(\bar{\delta}).$$

Also, if $R_\delta = R_\alpha \backslash R_{\alpha+\delta} = R(\Gamma_*, \bar{\Gamma}_\alpha) \cap \bar{S}_{\alpha+\delta}$, then the same assumptions lead to the estimate: $\int_{R_\delta} \phi^2(p)\, dx\, dy = \int_{\gamma_\alpha} \phi^2(p) \cdot |\bar{p} - p| \cdot |dp| + \bar{\delta}\varepsilon(\bar{\delta})$. Therefore,

$$\lambda_\rho(\alpha + \delta) - \lambda_\rho(\alpha) = \Delta(\alpha, \delta) - (1+\rho) \int_{R_\delta} \phi^2(p)\, dx\, dy$$

$$\leqq -\rho \int_{\gamma_\alpha} \phi^2(p) \cdot |\bar{p} - p| \cdot |dp| + \bar{\delta}\varepsilon(\bar{\delta}).$$

However,

$$\int_{\gamma_\alpha} \phi^2(p) \cdot |\bar{p} - p| \cdot |dp| \geqq (\min\{\phi^2(p) | p \in \Gamma_\alpha\}) \cdot (\text{length}\,(\gamma_\alpha))\underline{\delta},$$

from which it follows that $\lambda_\rho(\alpha + \delta) - \lambda_\rho(\alpha) < 0$ for any $\rho > 0$ provided that $\delta > 0$ is sufficiently small.

This completes the proof under the additional assumptions 1–3. However, under the general assumptions of the theorem one can choose two sequences of partitions $\{S_{*k}, \bar{S}_k, R[\Gamma_{*k}, \bar{\Gamma}_k)\}$, $\{S_{*k}, S_c, R(\Gamma_{*k}, \Gamma_c)\}$ such that assumptions 2 and 3 are fullfilled for each k and such that as $k \to \infty$, $R(\Gamma_{*k}, \bar{\Gamma}_k) \to R(\Gamma_*, \bar{\Gamma})$ (in the sense of set convergence), $H(\Gamma_{*k}, \Gamma_c) \to H(\Gamma_*, \Gamma_c)$, and $H(\Gamma_{*k}, \bar{\Gamma}_k) \to H(\Gamma_*, \bar{\Gamma})$. This eliminates assumptions 2 and 3. Thus, the inequality in Theorem 1 has been shown to hold whenever $R(\Gamma_*, \Gamma_c) \subset R(\Gamma_*, \bar{\Gamma})$. However, for any $\alpha \in I$, $\alpha \leqq c$, the same methods which were used above show that $H(\Gamma_*, \Gamma_c) =$

$H(\Gamma_*, \Gamma_\alpha) + \int_{R_{\alpha,c}} \phi^2(p)\, dx\, dy$, where $R_{\alpha,c} = R(\Gamma_*, \Gamma_\alpha) \cap S_c$. The general inequality follows by combining these results.

Remark 2. Theorem 1 also applies, after minor changes of statement, to Problem 2. Namely, in assumption (c) each boundary Γ_α is the finite disjoint union of smooth closed Jordan curves in $(0, 1) \times R$ and smooth Jordan arcs whose endpoints lie in $\{0, 1\} \times R$. Γ_α is assumed to be horizontal at its endpoints in $\{0, 1\} \times R$.

Remark 3. The inequality in Theorem 1 bears a resemblance to formulas of the Hadamard–Schiffer type for the variation of domain functionals. See, for example, Schiffer [8].

4. Heat flow inequalities related to the Lindelöf principle. For convenience, we state the suitable versions of the Lindelöf principle for Problems 1 and 2.

LEMMA 4 (Lindelöf principle). *Assume in Problem* 1 *or* 2 *that* $R(\Gamma_*, \bar{\Gamma}) \subset R(\Gamma_*, \Gamma)$. *Then*:

(a) *If* $p \in \Gamma_*$, *then* $|\nabla U(\Gamma_*, \bar{\Gamma}; p)| \geqq |\nabla U(\Gamma_*, \Gamma; p)|$ *if the derivatives exist.*

(b) *If* $p \in \Gamma \cap \bar{\Gamma}$, *then* $|\nabla U(\Gamma_*, \bar{\Gamma}; p)| \leqq |\nabla U(\Gamma_*, \Gamma; p)|$ *if the derivatives exist.*

(c) *If (in Problem* 2*)* Γ_* *and* Γ *are respectively the graphs of continuous functions* $y_*(x) < y(x)$ *on* $[0, 1]$ *and if* $p^* = (x^*, y^*)$ *is a point on* $\bar{\Gamma}$ *such that* $y(x^*) - y^* \geqq y(x) - y$ *for all* $(x, y) \in \bar{\Gamma}$, *then* $|\nabla U(\Gamma_*, \bar{\Gamma}; p^*)| \geqq |\nabla U(\Gamma_*, \Gamma; x^*, y(x^*))|$ *if both derivatives exist.*

(c′) *If (in Problem* 1*)* Γ_* *and* Γ *are respectively the graphs in polar coordinates of the continuous functions* $0 < r_*(\theta) < r(\theta)$ *on* $[0, 2\pi]$ *and if* $p^* = (r^*, \theta^*)$ *is a point on* $\bar{\Gamma}$ *such that* $\lambda := (r(\theta^*)/r^*) \geqq (r(\theta)/r)$ *for all* $(r, \theta) \in \bar{\Gamma}$, *then* $|\nabla U(\Gamma_*, \bar{\Gamma}; p^*)| \geqq \lambda \cdot |\nabla U(\Gamma_*, \Gamma; r(\theta^*), \theta^*)|$ *if both derivatives exist.*

Remark 5. The Lindelöf principle is discussed in [6, pp. 16–21]. Parts (c) and (c′) in Lemma 4 generalize the final statement of the Lindelöf principle by eliminating the assumption that $\bar{\Gamma}$ is the graph of a continuous function $\bar{y}(x) > y_*(x)$ or $\bar{r}(\theta) > r_*(\theta)$.

The method of proving the heat flow inequalities in this section is based on Theorem 1. To determine the inequality relating $H(\Gamma_*, \bar{\Gamma})$ to $H(\Gamma_*, \Gamma)$, one chooses a monotone class of boundaries $\{\Gamma_\alpha | \alpha \in I\}$ such that $\Gamma = \Gamma_c$ for some $c \in I$, $\bar{\Gamma} \subset \bigcup_{\alpha \in I} \Gamma_\alpha$, and such that the individual boundaries Γ_α admit to the most favorable application of the Lindelöf principle.

THEOREM 6. *In Problem* 2, *let* $\{S_*, S, R(\Gamma_*, \Gamma)\}$ *and* $\{S_*, \bar{S}, R(\Gamma_*, \bar{\Gamma})\}$ *represent two partitions of* $[0, 1] \times R$ *in which* S_* *is the same set and in which* Γ_* *and* Γ *are respectively the graphs of continuous functions* $y_*(x) < y(x)$ *on* $[0, 1]$. *Assume a continuous function* $\phi(p)$ *is defined on* Γ *such that for* $p \in \Gamma$ (*and* $p' \in R(\Gamma_*, \Gamma)$) *we have*: $\lim_{p' \to p} |\nabla U(\Gamma_*, \Gamma; p')| = \phi(p)$. *Then*:

$$H(\Gamma_*, \bar{\Gamma}) \geqq H(\Gamma_*, \Gamma) + \int_{R_1} \phi^2(x, y(x))\, dx\, dy - \int_{R_2} \phi^2(x, y(x))\, dx\, dy,$$

where $R_1 = R(\Gamma_*, \Gamma) \cap \bar{S}$ *and* $R_2 = R(\Gamma_*, \bar{\Gamma}) \cap S$. *If* $\bar{\Gamma}$ *is the graph of a continuous function* $\bar{y}(x) > y_*(x)$ *on* $[0, 1]$, *then the inequality can be written*:

$$H(\Gamma_*, \bar{\Gamma}) \geqq H(\Gamma_*, \Gamma) + \int_0^1 \phi^2(x, y(x)) \cdot (y(x) - \bar{y}(x))\, dx.$$

Proof. Assume in addition that the function $y(x)$ is L.c.d. on $[0, 1]$. Let $R_{1,2} = \text{Closure}(R_1 \cup R_2)$. It is sufficient in Theorem 1 for the continuous function $\phi(p)$ to be defined on $R_{1,2} \subset B(I)$ and for assumption (c) and the condition "$\Gamma_\alpha \cap \Gamma_\beta = \varnothing$ whenever $\alpha \neq \beta$" to hold relative to $R_{1,2}$. A class of boundaries $\{\Gamma_\alpha | \alpha \in I\}$ which satisfies the conditions of Theorem 1 in this generalized sense will now be defined. Choose $\varepsilon > 0$ such that $R(\Gamma_*, \Gamma_* + 3\varepsilon) \subset R(\Gamma_*, \bar{\Gamma})$. (Here $\Gamma_* + 3\varepsilon = \{(x, y) \in [0, 1] \times R | (x, y - 3\varepsilon) \in \Gamma_*\}$.) Let $I = (\alpha_0, \infty)$, where $\alpha_0 = \varepsilon - \max\{y(x) - y_*(x) | 0 \leqq x \leqq 1\} \leqq 0$. For each $\alpha \in I$, define Γ_α to be the graph of the function $y_\alpha(x) = \max\{y(x) + \alpha, y_*(x) + \varepsilon\}$. Then $\Gamma_0 = \Gamma$, the boundaries $\{\Gamma_\alpha | \alpha \in I\}$ are monotone and continuous in α (where the order of indexing has been reversed), and a positive continuous extension $\phi(p)$ of the function $\phi(p)$ defined on Γ exists on $R(\Gamma_* + 2\varepsilon, \infty) = \{(x, y) \in [0, 1] \times R | y > y_*(x) + 2\varepsilon\} \supset R_{1,2}$ such that if $p \in \Gamma_\alpha$, then $\phi(p) = |\nabla U(\Gamma_*, \Gamma_\alpha; p)|$. Therefore, Theorem 1 implies that:

$$H(\Gamma_*, \bar{\Gamma}) \geqq H(\Gamma_*, \Gamma) + \int_{R_1} \phi^2(p)\, dx\, dy - \int_{R_2} \phi^2(p)\, dx\, dy.$$

Application of the Lindelöf principle to the pair of regions $R(\Gamma_*, \Gamma)$ and $R(\Gamma_*, \Gamma_\alpha)$ for each $\alpha \in I$ shows that if $y_*(x) + 2\varepsilon < y \leqq y(x)$, then $\phi(x, y) \geqq \phi(x, y(x))$, whereas if $y \geqq y(x)$, then $0 < \phi(x, y) \leqq \phi(x, y(x))$. The asserted inequality follows from this.

The assumption that $y(x)$ is L.c.d. can be eliminated as follows. For each $0 < \delta < 1$, define $R(\Gamma_*, \Gamma_\delta) = \{p \in R(\Gamma_*, \Gamma) | U(\Gamma_*, \Gamma; p) > \delta\}$. For $0 < \delta < 1$, Γ_δ is the graph of a L.c.d. function $y_\delta(x)$ on $[0, 1]$. Further, if $\phi_\delta(x, y_\delta(x)) = |\nabla U(\Gamma_*, \Gamma_\delta; x, y_\delta(x))|$, then $y_\delta(x) \to y(x)$ and $\phi_\delta(x, y_\delta(x)) \to \phi(x, y(x))$ both uniformly on $[0, 1]$ as $\delta \to 0$. For each δ we have the inequality:

$$H(\Gamma_*, \bar{\Gamma}) \geqq H(\Gamma_*, \Gamma) + \int_{R_{1,\delta}} \phi_\delta^2(x, y_\delta(x))\, dx\, dy - \int_{R_{2,\delta}} \phi_\delta^2(x, y_\delta(x))\, dx\, dy,$$

where $R_{1,\delta} = R(\Gamma_*, \Gamma_\delta) \cap \bar{S}$ and $R_{2,\delta} = R(\Gamma_*, \bar{\Gamma}) \cap S_\delta$. The general inequality is obtained by taking the limit as $\delta \to 0$.

THEOREM 7. *In Problem* 2, *let* $\{S_*, S, R(\Gamma_*, \Gamma)\}$ *and* $\{\bar{S}_*, \bar{S}, R(\bar{\Gamma}_*, \bar{\Gamma})\}$ *be two partitions of* $[0, 1] \times R$ *in which* Γ_* *and* Γ *are respectively the graphs of continuous functions* $y_*(x) < y(x)$ *on* $[0, 1]$. *Assume there exists a continuous function* $\phi(p)$ *on* $\Gamma_* \cup \Gamma$ *such that for* $p \in \Gamma_* \cup \Gamma$ (*and for* $p' \in R(\Gamma_*, \Gamma)$) *we have*: $\lim_{p' \to p} |\nabla U(\Gamma_*, \Gamma; p)| = \phi(p)$. *Then the following two cases occur*:

(a) *Assume* $R(\bar{\Gamma}_*, \bar{\Gamma}) \subset R(\Gamma_*, \Gamma)$, *and let* $R_1 = R(\Gamma_*, \Gamma) \cap \bar{S}$ *and* $R_2 = R(\Gamma_*, \Gamma) \cap \bar{S}_*$. *Then*:

$$H(\bar{\Gamma}_*, \bar{\Gamma}) \geqq H(\Gamma_*, \Gamma) + \int_{R_1} \phi^2(x, y(x))\, dx\, dy + \int_{R_2} \phi^2(x, y_*(x))\, dx\, dy.$$

(b) *Assume* $R(\bar{\Gamma}_*, \bar{\Gamma}) \supset R(\Gamma_*, \Gamma)$, *and let* $R_1 = R(\bar{\Gamma}_*, \bar{\Gamma}) \cap S$ *and* $R_2 = R(\bar{\Gamma}_*, \bar{\Gamma}) \cap S_*$. *Then*:

$$H(\bar{\Gamma}_*, \bar{\Gamma}) \geqq H(\Gamma_*, \Gamma) - \int_{R_1} \phi^2(x, y(x))\, dx\, dy - \int_{R_2} \phi^2(x, y_*(x))\, dx\, dy.$$

If $\bar{\Gamma}_*$ *and* $\bar{\Gamma}$ *are also respectively the graphs of continuous functions* $\bar{y}_*(x) < \bar{y}(x)$ *on* $[0, 1]$, *then both the above inequalities reduce to*

$$H(\bar{\Gamma}_*, \bar{\Gamma}) \geqq H(\Gamma_*, \Gamma) + \int_0^1 \phi^2(x, y(x)) \cdot (y(x) - \bar{y}(x))\, dx + \int_0^1 \phi^2(x, y_*(x)) \cdot (\bar{y}_*(x) - y_*(x))\, dx.$$

Proof (for case (a)). If $R(\bar{\Gamma}_*, \bar{\Gamma}) \subset R(\Gamma_*, \Gamma)$, then it can be shown from the maximum principle that $H(\bar{\Gamma}_*, \bar{\Gamma}) - H(\bar{\Gamma}_*, \Gamma) \geqq H(\Gamma_*, \bar{\Gamma}) - H(\Gamma_*, \Gamma)$. Therefore,

$$\begin{aligned} H(\bar{\Gamma}_*, \bar{\Gamma}) &\geqq H(\Gamma_*, \bar{\Gamma}) + H(\bar{\Gamma}_*, \Gamma) - H(\Gamma_*, \Gamma) \\ &\geqq \left(H(\Gamma_*, \Gamma) + \int_{R_1} \phi^2(x, y(x))\, dx\, dy \right) \\ &\quad + \left(H(\Gamma_*, \Gamma) + \int_{R_2} \phi^2(x, y_*(x))\, dx\, dy \right) - H(\Gamma_*, \Gamma) \\ &= H(\Gamma_*, \Gamma) + \int_{R_1} \phi^2(x, y(x))\, dx\, dy + \int_{R_2} \phi^2(x, y_*(x))\, dx\, dy, \end{aligned}$$

by double application of Theorem 6.

THEOREM 8. *In Problem* 1, *let* $\{S_*, S, R(\Gamma_*, \Gamma)\}$ *and* $\{S_*, \bar{S}, R(\Gamma_*, \bar{\Gamma})\}$ *be two partitions of* R^2 *in which* S_* *is the same set. Assume there exists a point* $p_0 \in S_*$ *such that* Γ_* *and* Γ *are respectively the graphs in polar coordinates about* p_0 *of the continuous functions* $0 < r_*(\theta) < r(\theta)$ *on* $[0, 2\pi]$. *Assume a continuous function* $\phi(p)$ *is defined on* Γ *such that for each* $p \in \Gamma$ (*and for* $p' \in R(\Gamma_*, \Gamma)$) *we have*: $\lim_{p' \to p} |\nabla U(\Gamma_*, \Gamma; p')| = \phi(p)$. *Then*:

$$H(\Gamma_*, \bar{\Gamma}) \geqq H(\Gamma_*, \Gamma) + \int_{R_1} ((r(\theta) \cdot \phi(r(\theta), \theta))/r)^2 r\, dr\, d\theta - \int_{R_2} ((r(\theta) \cdot \phi(r(\theta), \theta))/r)^2 r\, dr\, d\theta,$$

where $R_1 = R(\Gamma_*, \Gamma) \cap \bar{S}$ *and* $R_2 = R(\Gamma_*, \bar{\Gamma}) \cap S$. *If* $\bar{\Gamma}$ *is the graph in polar coordinates about* p_0 *of a continuous function* $\bar{r}(\theta) > r_*(\theta)$ *on* $[0, 2\pi]$, *then the inequality reduces to*

$$H(\Gamma_*, \bar{\Gamma}) \geqq H(\Gamma_*, \Gamma) + \int_0^{2\pi} \phi^2(r(\theta), \theta) \cdot r^2(\theta) \cdot \log (r(\theta)/\bar{r}(\theta))\, d\theta.$$

Proof. The proof is analogous to the proof of Theorem 6. Assume $r(\theta)$ is L.c.d. and choose $\rho > 1$ such that $R(\Gamma_*, \rho^3 \Gamma_*) \subset R(\Gamma_*, \bar{\Gamma})$. (Here, $\rho^3 \Gamma_*$ is the graph in polar coordinates about p_0 of the function $\rho^3 r_*(\theta)$.) Let $I = (\alpha_0, \infty)$, where $\alpha_0 = \min \{\rho r_*(\theta)/r(\theta) | \theta \in [0, 2\pi]\} < 1$. For each $\alpha \in I$, let Γ_α be the graph in polar coordinates about p_0 of the function $r_\alpha(\theta) = \max \{\alpha r(\theta), \rho r_*(\theta)\}$. Then $\Gamma_1 = \Gamma$, the boundaries $\{\Gamma_\alpha | \alpha \in I\}$ are monotone and continuous in α, and the function $\phi(p)$ such that $\phi(p) = |\nabla U(\Gamma_*, \Gamma_\alpha; p)|$ on Γ_α is a positive continuous extension to

$R(\rho^2\Gamma_*, \infty) = \{(r, \theta | r > \rho^2 r_*(\theta)\}$ of the function $\phi(p)$ defined on Γ. Therefore, Theorem 1 implies that: $H(\Gamma_*, \bar{\Gamma}) \geqq H(\Gamma_*, \Gamma) + \int_{R_1} \phi^2(p) r\, dr\, d\theta - \int_{R_2} \phi^2(p) \cdot r\, dr\, d\theta$. Application of the Lindelhöf principle to the pair of regions $R(\Gamma_*, \Gamma)$ and $R(\Gamma_*, \Gamma_\alpha)$ for each $\alpha \in I$ shows that if $\rho^2 r_*(\theta) < r \leqq r(\theta)$, then $\phi(r, \theta) \geqq (r(\theta) \cdot \phi(r(\theta), \theta)/r)$, whereas if $r \geqq r(\theta)$, then $0 < \phi(r, \theta) \leqq (r(\theta) \cdot \phi(r(\theta), \theta)/r)$.

In order to eliminate the assumption that $r(\theta)$ is L.c.d., we again define $R(\Gamma_*, \Gamma_\delta) = \{p \in R(\Gamma_*, \Gamma) | U(\Gamma_*, \Gamma; p) > \delta\}$ (where $0 < \delta < 1$). Again, Γ_δ is the graph in polar coordinates about p_0 of a L.c.d. function $r_\delta(\theta)$, so that the procedure used in the proof of Theorem 6 carries over to this case.

Remark 9. Assume in Theorems 6 and 8 that $\phi(p) > 0$ on Γ. Then the inequality cannot reduce to equality except in the case where $R_1 \cup R_2$ has zero area. The proof in the case of Theorem 6 will now be sketched. In the notation of Theorem 6, choose α_1 such that $R(\Gamma_*, \bar{\Gamma}) \subset R(\Gamma_*, \Gamma_{\alpha_1})$ and $\alpha_1 > \max\{y(x) - y_*(x) | x \in [0, 1]\}$. Define the continuous function $0 < \mu(\alpha) < 1$ on $(0, \alpha_1)$ by: $\mu(\alpha) = \max\{U(\Gamma_*, \Gamma_{\alpha_1}; p) | p \in \Gamma_* + \alpha\}$. One can show, if $y(x)$ is L.c.d. and $\phi(x, y)$ is the extension to $R_1 \cup R_2$ of $\phi(x, y(x))$, that $\phi(x, y(x) + \alpha) \leqq \mu(\alpha) \cdot \phi(x, y(x))$ for all $(x, y(x) + \alpha) \in R_2$ and $\phi(x, y(x) - \alpha) \geqq (1/\mu(\alpha)) \cdot \phi(x, y(x))$ for all $(x, y(x) - \alpha) \in R_1$. This leads to the inequality:

$$H(\Gamma_*, \bar{\Gamma}) \geqq H(\Gamma_*, \Gamma) + \int_{R_1} (1/\mu(y(x) - y))^2 \cdot \phi^2(x, y(x))\, dx\, dy - \int_{R_2} \mu^2(y - y(x)) \cdot \phi^2(x, y(x))\, dx\, dy.$$

This stronger inequality continues to hold under the general assumptions of Theorem 6, since the function $\mu(\alpha)$ can be held fixed as $\delta \to 0$ in the final stage of the proof. (Note that the function $\mu(\alpha)$ is dependent on Γ_*, Γ and $\bar{\Gamma}$.)

5. Heat flow minimizing boundaries. The basic heat flow minimization problem which we consider is the following. *For S_* (and Γ_*) and $A > 0$ fixed, determine the set S (or its boundary Γ) such that the heat flow $H(\Gamma_*, \Gamma)$ from S_* into S is minimized subject to the constraint that $A(\Gamma_*, \Gamma) \leqq A$.*

This problem is essentially reduced to a corresponding free boundary problem for the Laplace equation by means of the following theorem.

THEOREM 10. *In Problem* 1 *or* 2, *let S_* and the function $a(p) > 0$ be fixed. Assume for each c in a positive open interval I that there exists a region $R(\Gamma_*, \Gamma_c)$ such that whenever $p \in \Gamma_c$ (and $p' \in R(\Gamma_*, \Gamma_c)$): $\lim_{p' \to p} |\nabla U(\Gamma_*, \Gamma_c; p)| = c \cdot a(p)$. Assume the class of boundaries $\{\Gamma_c | c \in I\}$ satisfies assumptions* (a), (b) *and* (c) *of Theorem* 1. (*In Problem* 2, *assumption* (c) *is altered according to Remark* 2.) *Then for each $c \in I$, Γ_c has the following properties*:

(a) *Let $\{S_*, \bar{S}, R(\Gamma_*, \bar{\Gamma})\}$ be any partition with $\bar{\Gamma} \subset B(I) = \cup_{c \in I} \Gamma_c$ and $\bar{\Gamma} \neq \Gamma_c$. Then: $H(\Gamma_*, \bar{\Gamma}) > H(\Gamma_*, \Gamma_c) + c^2 \cdot (A(\Gamma_*, \Gamma_c) - A(\Gamma_*, \bar{\Gamma}))$.*

(b) *Let $A = A(\Gamma_*, \Gamma_c)$. Then for any boundary $\bar{\Gamma} \neq \Gamma_c$ for which $\bar{\Gamma} \subset B(I)$ and $A(\Gamma_*, \bar{\Gamma}) \leqq A$, the inequality: $H(\Gamma_*, \bar{\Gamma}) > H(\Gamma_*, \Gamma_c)$ holds. Therefore Γ_c is uniquely heat flow minimizing in the class of all boundaries $\Gamma \subset B(I)$ for which $A(\Gamma_*, \Gamma) \leqq A$.*

Proof. Let $c(p)$ be a function on $B(I)$ such that $c(p) = c$ on Γ_c for each $c \in I$. If $c, c' \in I$ and $c \neq c'$, then the distance between Γ_c and $\Gamma_{c'}$ is positive. It follows from this that $c(p) > 0$ is a continuous function on $B(I)$. The boundaries $\{\Gamma_c | c \in I\}$ and

the function $\phi(p) = c(p) \cdot a(p)$ satisfy the conditions of Theorem 1. Therefore, for any $c \in I$ and any region $R(\Gamma_*, \bar{\Gamma})$ we have:

$$H(\Gamma_*, \bar{\Gamma}) \geqq H(\Gamma_*, \Gamma_c) + \int_{R_1} c^2(p) \cdot a^2(p)\, dx\, dy - \int_{R_2} c^2(p) \cdot a^2(p)\, dx\, dy,$$

where $R_1 = R(\Gamma_*, \Gamma_c) \cap \bar{S}$ and $R_2 = R(\Gamma_*, \bar{\Gamma}) \cap S_c$. However, $c(p) > c$ in the interior of $R(\Gamma_*, \Gamma_c)$ and $0 < c(p) < c$ in the interior of S_c. Therefore, if $R_1 \cup R_2$ has positive area, then:

$$\begin{aligned} H(\Gamma_*, \bar{\Gamma}) - H(\Gamma_*, \Gamma_c) &> c^2 \cdot \left(\int_{R_1} a^2(p)\, dx\, dy - \int_{R_2} a^2(p)\, dx\, dy \right) \\ &= c^2 \cdot (A(\Gamma_*, \Gamma_c) - A(\Gamma_*, \bar{\Gamma})). \end{aligned}$$

The existence of free bounary solutions Γ_c such that whenever $p \in \Gamma_c$ (and $p' \in R(\Gamma_*, \Gamma_c)$): $\lim_{p' \to p} |\nabla U(\Gamma_*, \Gamma_c; p')| = c \cdot a(p)$ has been investigated by A. Beurling [2]. Beurling calls a class of free boundary solutions $\{\Gamma_c | c \in I\}$ elliptic if $R(\Gamma_*, \Gamma_c) \subset R(\Gamma_*, \Gamma_{c'})$ whenever $c \geqq c'$ and hyperbolic if the reverse inclusion always holds. For example, if S_* is the exterior of the unit circle and $a(p) = 1$ within the circle, then a hyperbolic class of solutions $\{\Gamma_c | e < c < \infty\}$ exists as concentric circles of radius less than e^{-1}. Beurling's existence results, as well as the present Theorem 10, apply only to the elliptic case. Reasonably general and easily applied conditions on S_* and $a(p)$ under which a suitable elliptic class of free boundary solutions exists are provided by the following generalization of a result of D. E. Tepper [10], [11].

LEMMA 11. *In Problem* 1, *let* S_* *be starlike with respect to an interior point* p_0. *Assume for each* $p \in R^2$ *that* $\lambda \cdot a(p_0 + \lambda p)$ *is monotone nondecreasing with increasing* λ *on* $[0, \infty)$. *Then*:

(a) *For any constant* $c > 0$ *there exists a unique doubly connected region* $R(\Gamma_*, \Gamma_c)$ *such that if* $p \in \Gamma_c$ (*and* $p' \in R(\Gamma_*, \Gamma_c)$) *then*: $\lim_{p' \to p} |\nabla U(\Gamma_*, \Gamma_c; p')| = c \cdot a(p)$.

(b) *If* $c' \geqq c > 0$, *then* $R(\Gamma_*, \Gamma_{c'}) \subset R(\Gamma_*, \Gamma_c)$. *If* $c' \neq c$, *then* $\Gamma_c \cap \Gamma_{c'} = \varnothing$.

(c) $\bigcup_{c>0} \Gamma_c =$ *the complement of* S_* *in* R^2.

(d) *For each* $c > 0$, $S_* \cup R(\Gamma_*, \Gamma_c)$ *is starlike with respect to* p_0. *Moreover*:

(d′) *Each boundary* Γ_c *is the graph in polar coordinates about* p_0 *of a continuous function* $r_c(\theta)$.

Proof. In the case where $a(p) = 1$, the above results (with the exception of (d′)) were obtained by D. E. Tepper [10], [11], who based his method on the results of A. Beurling [2]. In fact the arguments used in [10] and [11] exactly suffice to prove the present lemma (except (d′)), so that they are not repeated here.

Let $p_0 = 0$. One obtains (d′) from (d) by showing that no boundary Γ_c contains a radial line segment. In fact if $p \in \Gamma_c$ and $\lambda p \in \Gamma_c$, $\lambda = (1/\nu) > 1$, then $U(\nu\Gamma_*, \nu\Gamma_c; q) < U(\nu\Gamma_*, \Gamma_c; q) < U(\Gamma_*, \Gamma_c; q)$ on $R(\Gamma_*, \nu\Gamma_c)$, from which it follows (using $a(p) > 0$) that: $c\lambda \cdot a(\lambda p) = \lambda \cdot \lim_{p' \to \lambda p} |\nabla U(\Gamma_*, \Gamma_c; p')| = \lim_{p' \to p} |\nabla U(\nu\Gamma_*, \nu\Gamma_c; p')| < c \cdot a(p)$ in contradiction to the monotone property of the function $a(p)$.

An existence theorem for our heat flow minimization problem could now be obtained essentially by combining Theorem 10 and Lemma 11. However, one

would first have to find conditions on $a(p)$ under which the family of boundaries $\{\Gamma_c | c>0\}$ is smoothly varying (Theorem 1 (c)). The following existence theorem follows directly from Theorem 8, Remark 9, and Lemma 11.

THEOREM 12. *In Problem* 1, *let* Γ_* *be the graph in polar coordinates (about a point* p_0 *interior to* S_**) of a continuous function* $r_*(\theta)$. *Assume for each* $p \in R^2$ *that* $\lambda \cdot a(p_0+\lambda p)$ *is monotone nondecreasing with increasing* λ *on* $[0, \infty)$. *Let* $\{\Gamma_c | c>0\}$ *be the class of free boundary solutions whose existence is asserted in Lemma* 11. *Then for any constant* $A>0$ *there exists a unique* $c>0$ *such that* $A(\Gamma_*, \Gamma_c) = A$. *If* $\{S_*, \bar{S}, R(\Gamma_*, \bar{\Gamma})\}$ *is any partition of* R^2 *for which* $A(\Gamma_*, \bar{\Gamma}) \leqq A$ *and* $\bar{\Gamma} \neq \Gamma_c$, *then* $H(\Gamma_*, \bar{\Gamma}) > H(\Gamma_*, \Gamma_c)$. *Thus* Γ_c *is the unique heat flow minimizing boundary at the area* A.

Remark 13. If $a(p)=1$ in Theorem 12, then the heat flow minimizing boundaries are boundaries of constant heat flow density. If S_* is a circular disc and $a(p)=1$, then the solutions Γ_c, $c>0$, are circles concentric to Γ_*. In this situation, Theorem 12 reduces to a special case of the result of Carleman [4] and Szegö [9].

Remark 14. Lemma 11 and Theorem 12 have fully analogous statements in the context of Problem 2. For example, if Γ_* is the graph of a continuous function $y_*(x)$ on $[0, 1]$ and if $a(x, y)>0$ is monotone nondecreasing in y for each $x \in [0, 1]$, then the elliptic class of boundaries $\{\Gamma_c | c>0\}$ exists and fills the complement of S_* in $[0, 1] \times R$. Each boundary Γ_c is the graph of a continuous function $y_c(x)$ on $[0, 1]$ and is uniquely heat flow minimizing at its area.

Remark 15. Theorems 1, 6, 7, 8, and 10 can all be extended to analogous temperature problems in 3 dimensions without any significant changes in statement or in the proofs. This is because Green's theorem and the maximum principle, on which the proofs are fundamentally based, have analogous forms in 3 dimensions. On the other hand, conformal mapping plays an important role in the results in [2], on which Lemma 11 is based. Thus, an extension of Lemma 11 (and therefore of Theorem 12) to 3 dimensions would apparently require a fundamentally new proof technique.

Remark 16. Results analogous to those in Theorems 1 and 10 can be obtained under other boundary conditions. As an example, assume Γ_* is L.c.d. and let $\lambda>0$ be fixed. Let $U(\Gamma_*, \Gamma; p)$ be the unique continuous function on $S \cup R(\Gamma_*, \Gamma)$ such that $U(\Gamma_*, \Gamma; p)=1$ on S, $U(\Gamma_*, \Gamma; p)$ is harmonic on $R(\Gamma_*, \Gamma)$, and the radiation condition: $\mathbf{D}_n U(\Gamma_*, \Gamma; p) = \lambda \cdot U(\Gamma_*, \Gamma; p)$ holds on Γ_*. Let $H(\Gamma_*, \Gamma)$ be the heat flow from S into S_* due to $U(\Gamma_*, \Gamma; p)$. Then Theorems 1 and 10 both hold without alteration of statement in this new situation.

As a second example, let $f(p)$ be a positive continuous function on S_* and let $U(\Gamma_*, f, \Gamma; p)$ be the unique continuous function on R^2 such that $U(\Gamma_*, f, \Gamma; p) = f(p)$ for all $p \in S_*$, $U(\Gamma_*, f, \Gamma; p)=0$ on S, and $U(\Gamma_*, f, \Gamma; p)$ is harmonic on $R(\Gamma_*, \Gamma)$. Let $H(\Gamma_*, f, \Gamma)$ be the heat flow from S_* into S due to $U(\Gamma_*, f, \Gamma; p)$. In this case, assumption (d) in Theorem 1 generalizes to the assumption that positive continuous functions $\phi(p)$ and $\psi(p)$ are defined on $B(I)$ such that for any $p \in B(I)$:

$$\lim_{p' \to p} |\nabla U(\Gamma_*, 1, \Gamma_\alpha; p')| = \phi(p) \quad \text{and} \quad \lim_{p' \to p} |\nabla U(\Gamma_*, f, \Gamma_\alpha; p')| = \psi(p),$$

where Γ_α is the unique boundary containing p, and $p' \in R(\Gamma_*, \Gamma_\alpha)$. (The remaining

assumptions remain unchanged.) Then the heat flow inequality asserted in Theorem 1 has the generalized form:

$$H(\Gamma_*, f, \bar{\Gamma}) \geqq H(\Gamma_*, f, \Gamma_c) + \int_{R_1} \phi(p) \cdot \psi(p)\, dx\, dy - \int_{R_2} \phi(p) \cdot \psi(p)\, dx\, dy.$$

In the generalization of Theorem 10, one assumes that there exists a class of boundaries $\{\Gamma_c | c \in I\}$ such that assumptions (a), (b), and (c) of Theorem 1 hold, and such that positive continuous functions $\phi(p)$ and $\psi(p)$ (with the above stated definitions) exist on $B(I)$. It is further assumed that for each $c \in I$, $\phi(p) \cdot \psi(p) = c^2 \cdot a^2(p)$ for all $p \in \Gamma_c$. (This is the generalized condition for the free boundary Γ_c.) Then the conclusions of Theorem 10 (with $H(\Gamma_*, f, \Gamma)$ replacing $H(\Gamma_*, \Gamma)$) remain the same.

6. An area minimization problem on classes of conformally equivalent regions. The doubly connected regions $R(\Gamma_*, \Gamma)$ and $R(\bar{\Gamma}_*, \bar{\Gamma})$ (imbedded in the complex plane) are conformally equivalent if there exists an analytic function $F(z)$ which maps $R(\Gamma_*, \Gamma)$ conformally onto $R(\bar{\Gamma}_*, \bar{\Gamma})$ and whose continuous extension to the boundary maps Γ onto $\bar{\Gamma}$ and Γ_* onto $\bar{\Gamma}_*$.

The following free boundary optimization problem can be defined in the context of Problem 1. *Let a doubly connected region $R(\Gamma_*, \Gamma_0)$ be fixed, and let* **K** *be the class of all doubly connected regions $R(\Gamma_*, \Gamma)$ (with all properties stated in Problem* 1) *which have the boundary component Γ_* and which are conformally equivalent to $R(\Gamma_*, \Gamma_0)$. We seek that region in the class* **K** *for which the weighted area $A(\Gamma_*, \Gamma)$ is minimum.*

THEOREM 17. *In the problem described above, assume Γ_* is the graph in polar coordinates about $p_0 \in S_*$ of a continuous function $r_*(\theta) > 0$. Assume $\lambda \cdot a(p_0 + \lambda p)$ is monotone nondecreasing in λ on $[0, \infty)$ for each $p \in R^2$. Let $\{\Gamma_c | c > 0\}$ be the class of free boundary solutions whose existence was proven in Lemma* 11. *Then there is a unique $c > 0$ such that $R(\Gamma_*, \Gamma_c)$ is in the class* **K**. *If $R(\Gamma_*, \bar{\Gamma})$ is any region in* **K** *for which $\bar{\Gamma} \neq \Gamma_c$, then $A(\Gamma_*, \bar{\Gamma}) > A(\Gamma_*, \Gamma_c)$. Thus, $R(\Gamma_*, \Gamma_c)$ is uniquely area minimizing in the class* **K**.

Proof. $H(\Gamma_*, \Gamma_c)$ is a continuous, strictly monotone increasing function of c on $(0, \infty)$ such that $H(\Gamma_*, \Gamma_c) \to 0$ as $c \to 0+$ and $H(\Gamma_*, \Gamma_c) \to \infty$ as $c \to \infty$. Therefore, there is a unique $c > 0$ such that $H(\Gamma_*, \Gamma_c) = H(\Gamma_*, \Gamma_0)$. For all regions $R(\Gamma_*, \bar{\Gamma})$ in the class **K**, the conformal equivalence implies that $H(\Gamma_*, \bar{\Gamma}) = H(\Gamma_*, \Gamma_c)$. However, if $A(\Gamma_*, \bar{\Gamma}) \leqq A(\Gamma_*, \Gamma_c)$ and $\bar{\Gamma} \neq \Gamma_c$, then $H(\Gamma_*, \bar{\Gamma}) > H(\Gamma_*, \Gamma_c)$ due to Theorem 12. Thus if $\bar{\Gamma} \neq \Gamma_c$, then $A(\Gamma_*, \bar{\Gamma}) > A(\Gamma_*, \Gamma_c)$.

7. A dual heat flow maximization problem. The results in this section involve the following dual temperature problem for Problem 2.

Problem 3. Let $\{S_*, S, R(\Gamma_*, \Gamma)\}$ be any partition of $[0, 1] \times R$ having the properties stated in Problem 2. Assume in addition that $R(\Gamma_*, \Gamma)$ is simply connected. Let $W(\Gamma_*, \Gamma; p)$ be the harmonic conjugate function of $U(\Gamma_*, \Gamma; p)$ on $R(\Gamma_*, \Gamma)$ for which $W(\Gamma_*, \Gamma; p) = 0$ on $(\{0\} \times R) \cap R(\Gamma_*, \Gamma)$. Let $V(\Gamma_*, \Gamma; p) = C \cdot W(\Gamma_*, \Gamma; p)$, where C is a constant such that $V(\Gamma_*, \Gamma; p) = 1$ on $(\{1\} \times R) \cap R(\Gamma_*, \Gamma)$. $V(\Gamma_*, \Gamma; p)$ can be interpreted as the temperature in $R(\Gamma_*, \Gamma)$ under the conditions that S_* and S are perfectly insulating and $R(\Gamma_*, \Gamma)$

is homogeneous and partially conducting and is held at respective temperatures 0 and 1 along the left and right vertical sides. Thus, $R(\Gamma_*, \Gamma)$ can be viewed as a heat conduit joining two perfect conductors. $J(\Gamma_*, \Gamma)$ is the rate of heat flow through $R(\Gamma_*, \Gamma)$, i.e.,

$$J(\Gamma_*, \Gamma) = \int_\gamma \mathbf{D}_x V(\Gamma_*, \Gamma; 0, y)\, dy,$$

where $\gamma = (\{0\} \times R) \cap R(\Gamma_*, \Gamma)$ and $\mathbf{D}_x$ represents the right-hand derivative.

The heat flow maximization problem is the following: *For S_* (or its boundary Γ_*) and $A > 0$ fixed, determine the set S (or its boundary Γ) such that the rate of heat flow $J(\Gamma_*, \Gamma)$ across $R(\Gamma_*, \Gamma)$ is maximized subject to the constraint that $A(\Gamma_*, \Gamma) \leqq A$.*

The above heat flow maximization problem is essentially the dual of the heat flow minimization problem (in the context of Problem 2) in § 5. In fact $H(\Gamma_*, \Gamma) \cdot J(\Gamma_*, \Gamma) = 1$ and $|\nabla V(\Gamma_*, \Gamma; p)| = J(\Gamma_*, \Gamma) \cdot |\nabla U(\Gamma_*, \Gamma; p)|$ for each p in the interior of $R(\Gamma_*, \Gamma)$. Thus, when S_*, the function $a(p)$, and a constant $A > 0$ are fixed, the free boundary Γ which minimizes $H(\Gamma_*, \Gamma)$ under the constraint $A(\Gamma_*, \Gamma) \leqq A$, also maximizes $J(\Gamma_*, \Gamma)$ under the same constraint. Therefore, if Γ_* is the graph of a continuous function $y_*(x)$ on $[0, 1]$ and $a(x, y)$ is monotone nondecreasing in y on $(-\infty, \infty)$ for each $x \in [0, 1]$, then the existence and uniqueness and a characterization of the heat flow maximizing boundaries of the above problem all follows directly from Lemma 11, Theorem 12, and Remark 14. If Γ is heat flow maximizing at the area $A > 0$, then there is a constant $c > 0$ such that for all $p \in \Gamma$ (and $p' \in R(\Gamma_*, \Gamma)$): $\lim_{p' \to p} |\nabla V(\Gamma_*, \Gamma; p')| = c \cdot a(p)$.

All results related to Problem 2 in §§ 3, 4, and 5 have equivalent statements in the context of Problem 3. As an example, we state the result equivalent to Theorem 6.

THEOREM 18. *In Problem 3, let $\{S_*, S, R(\Gamma_*, \Gamma)\}$ and $\{S_*, \bar{S}, R(\Gamma_*, \bar{\Gamma})\}$ be two partitions of $[0, 1] \times R$ in which S_* is the same set and Γ_* and Γ are the respective graphs of the continuous functions $y_*(x) < y(x)$ on $[0, 1]$. Assume a positive continuous function $\phi(p)$ is defined on Γ, such that for all $p \in \Gamma$ (and $p' \in R(\Gamma_*, \Gamma)$): $\lim_{p' \to p} |\nabla V(\Gamma_*, \Gamma; p)| = \phi(p)$. Then:*

$$J(\Gamma_*, \bar{\Gamma}) \leqq J^2(\Gamma_*, \Gamma) \cdot \left(J(\Gamma_*, \Gamma) + \int_{R_1} \phi^2(x, y(x))\, dx\, dy - \int_{R_2} \phi^2(x, y(x))\, dx\, dy \right)^{-1},$$

where $R_1 = R(\Gamma_*, \Gamma) \cap \bar{S}$ and $R_2 = R(\Gamma_*, \bar{\Gamma}) \cap S$.

Note added in proof. The generalization of Theorem 1 used in the proof of Theorem 6 is itself relatively hard to prove. Therefore, Theorem 6 is more easily proved by defining a class of boundaries which allows the same application of the Lindelöf principle while fulfilling the requirements of Theorem 1 (Remark 2) in the original sense. Assuming (without loss of generality) that Γ_* and Γ are L.c.d. and horizontal at their endpoints, we can define such a class as follows. Choose $\varepsilon > 0$ such that $R(\Gamma_*, \Gamma_* + 3\varepsilon) \subset R(\Gamma_*, \bar{\Gamma}) \cap R(\Gamma_*, \Gamma)$, and define $\alpha_0 < 0$ as before.

For $\alpha \geqq 0$, define $\Gamma_\alpha = \Gamma + \alpha$. For $\alpha_0 < \alpha < 0$, let Γ_α be the boundary in $[0, 1] \times R$ of the union of all discs $B_\delta(p) \subset (S + \alpha) \cap R(\Gamma_* + (2 - (\alpha/\alpha_0))\varepsilon, \infty)$. Here $B_\delta(p) = \{q \in [0, 1] \times R \| q - p| < \delta\}$ and the constant $\delta > 0$ is chosen so small that $\Gamma_\alpha \cap R(\Gamma_* + 2\varepsilon, \Gamma) = (\Gamma + \alpha) \cap R(\Gamma_* + 2\varepsilon, \Gamma)$ for all $\alpha_0 < \alpha < 0$. An analogous class of boundaries can be constructed for the proof of Theorem 8.

REFERENCES

[1] LARS V. AHLFORS, *Complex Analysis*, McGraw-Hill, New York, 1966.

[2] A. BEURLING, *On free-boundary problems for the Laplace equation*, Seminars on Analytic Functions, vol. 1, Institute for Advanced Study, Princeton, N.J., 1957, pp. 248–263.

[3] S. BHARGAVA AND R. J. DUFFIN, *Dual extremum principles related to optimum beam design*, Arch. Rational Mech. Anal., 50 (1972), pp. 314–330.

[4] T. CARLEMAN, *Über ein Minimalproblem der Mathematischen Physik*, Math. Z., 1 (1918), pp. 208–212.

[5] R. J. DUFFIN AND D. K. MCLAIN, *Optimum shape of a cooling fin on a convex cylinder*, J. Math. Mech., 17 (1968), pp. 769–784.

[6] M. A. LAVRENTEV, *Variational Methods*, P. Noordhoff, Groningen, the Netherlands, 1963.

[7] ARTHUR H. LUSTY, JR. AND ANGELO MIELE, *Bodies of maximum lift to drag ratio in hypersonic flow*, AIAA, 4 (1966), pp. 2130–2135.

[8] M. SCHIFFER, *Variation of domain functionals*, Bull. Amer. Math. Soc., 60 (1954), pp. 303–328.

[9] G. SZEGÖ, *Über Einige Extremalaufgaben der Potentialtheorie*, Math. Z., 31 (1930), pp. 583–593.

[10] D. E. TEPPER, *Free boundary problem*, this Journal, 5 (1974), pp. 841–846.

[11] ———, *Free boundary problem, the starlike case*, this Journal, 6 (1975), pp. 503–505.

[12] A. ACKER, *A free boundary optimization problem*, this Journal, to appear. Abstract: Notices Amer. Math. Soc., 23 (1976), p. A-645.

[13] ———, *An isoperimetric inequality involving conformal mapping*, Proc. Amer. Math. Soc., to appear. Abstract: Notices Amer. Math. Soc., 24 (1977), p. A-12.

[14] ———, *An isoperimetric inequality involving heat flow under linear radiation conditions*, Proc. Amer. Math. Soc., to appear.

[15] ———, *Free boundary optimization problems involving geometric constraints*, submitted. Abstract: Notices Amer. Math. Soc., 24 (1977).

Note. An abstract of this paper appeared in Notices Amer. Math. Soc., 23 (1976), p. A-585.

SIAM J. MATH. ANAL.
Vol. 8, No. 4, August 1977

HIGH SPEED CONVOLUTION OF PERIODIC FUNCTIONS*

R. S. BUCY,† A. J. MALLINCKRODT‡ AND H. YOUSSEF¶

Abstract. A method is given to approximate the convolution of two-dimensional functions periodic in each variate, by the convolution of two appropriate periodic functions of one variable. An explicit bound for the error is given, and numerical result presented.

1. Introduction. In various problems, it becomes important to evaluate numerically convolutions of periodic functions of more than one variable. The problem of convolving functions of one variable has been studied extensively and various mega-Hertz bandwidth devices have been proposed to achieve high speed correlation—see [1] and [2]. Our interest in the problem arose because we were faced with the following problem, given real valued functions of real arguments; $F(x, y)$ and $G(x, y)$ periodic of period 2π in each argument evaluate $F * G(x, y)$ where

$$F * G(x, y) = \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} F(x-u, y-v) G(u, v)\, du\, dv. \tag{1.1}$$

This problem arose in the course of investigations of the problem of optimal phase demodulation—see [3]. In order to achieve the realization of the optimal demodulator, $F * G$ must be computed each time new observations arrive to find the next phase estimate. In obtaining the performance of the optimal demodulator, reported in [3], by Monte Carlo simulation, the convolution operation had to be performed 26,000 times. The digital realization of the optimal demodulation was in effect speed limited by a convolution of the type (1.1).

In this paper we will be concerned with a periodic mapping ϕ of $[0, T]$ into $S^2\{(x, y) | 0 \leqq x < 2\pi, 0 \leqq y < 2\pi\}$. We will replace a function $F(x, y)$ defined on S^2 by the function $\tilde{F}(t)$

$$\tilde{F}(t) = \phi(F) = F(x(t), y(t)). \tag{1.2}$$

As t ranges from 0 to T, the pairs $(x(t), y(t))$ range over the diagonal lines shown in Fig. 1.

We demonstrate that

$$(F * G)(x(t), y(t)) = (\tilde{F} * \tilde{G})(t) - E(t) \tag{1.3}$$

where $E(t)$, the error term, is shown to be small under certain hypotheses.

$$(\tilde{F} * \tilde{G})(t) \triangleq (\phi(F) * \phi(G))(t) \triangleq \frac{1}{T} \int_0^T \tilde{F}(t-s) \tilde{G}(s)\, ds.$$

* Received by the editors May 30, 1975, and in revised form July 26, 1976. This research was supported in part by the U.S. Air Force Office of Scientific Research, Air Force Systems Command, under Grant AFOSR-2141, and in part by the National Science Foundation under Grant O.I.P.-7420601.

† Department of Mathematics, University of Southern California, Los Angeles, California 90007.

‡ Communication Research Laboratory, Santa Ana, California 92705.

¶ Lockheed–California Company, Burbank, California 91503.

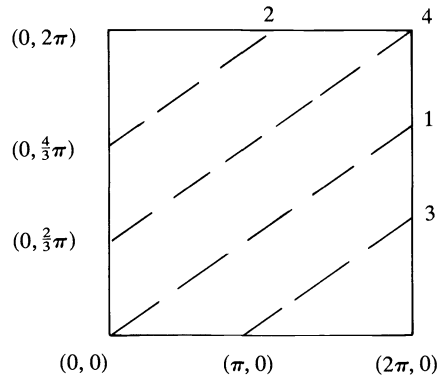


FIG. 1.

The computation of $\tilde{F} * \tilde{G}$ is much quicker using a high speed correlator than that of $(F * G)(x, y)$, and $(F * G)(x, y)$ can thus be obtained quickly for (x, y) any grid of points contained in the diagonal lines.

This paper will concern itself with such a mapping. Although the ϕ which we consider applies to functions periodic in two variates, the multivariate case is an easy generalization and is left to the reader.

2. Properties of the sweep mapping. Consider the map ϕ to be

$$\phi: t \to (\alpha t \bmod 2\pi, \beta t \bmod 2\pi).$$

It is well known that ϕ is periodic iff α/β is rational, say r_1/s_1 where r_1 and s_1 are relatively prime—see [5]. Let T^* be the period; then if $x(t) = \alpha t \bmod 2\pi$ and $y(t) = \beta t \bmod 2\pi$,

$$x(T^* + t) = x(t), \qquad y(T^* + t) = y(t),$$

or

$$\alpha T^* = 2\pi k, \qquad \beta T^* = 2\pi l \quad \text{for some integers } k \text{ and } l. \tag{2.1}$$

Now if $\lambda = \alpha s_1 = \beta r_1$, then

$$T^* = \frac{2\pi k s_1}{\lambda}, \qquad T^* = \frac{2\pi l r_1}{\lambda}$$

or $k = v r_1$ and $l = v s_1$ for v an integer. But $T = 2\pi r_1 s_1/\lambda$ is a period, and hence a minimum period, as it corresponds to $v = 1$. Hence, the following is true.

LEMMA 1. *Let $\alpha/\beta = r_1/s_1$ with r_1 and s_1 relatively prime; then the period of ϕ is $T = 2\pi r_1 s_1/\lambda$ where $\lambda = \alpha s_1 = \beta r_1$. Further,*

$$\alpha T = 2\pi r_1, \qquad \beta T = 2\pi s_1, \tag{2.2}$$

or $x(t) = 0$ and $y(t) = 0$ have r_1 and s_1 roots for $0 < t \leqq T$. The roots are

$$\begin{aligned} t_i &= 2\pi \frac{i s_1}{\lambda}, \qquad i = 1, \cdots, r_1, \\ t_j &= 2\pi j \frac{r^1}{\lambda}, \qquad j = 1, \cdots, s_1, \end{aligned} \tag{2.3}$$

respectively, and[1]

$$y(t_i) = \left(\left(\frac{is_1}{r_1}\right)\right) * 2\pi, \qquad x(t_j) = \left(\left(\frac{jr_1}{s_1}\right)\right) * 2\pi. \tag{2.4}$$

Proof. The previous argument established the period T as corresponding to $v = 1$, and hence (1.3) implies (1.4). If $x(t) = 0$, then $\alpha t = 2\pi i$ for some $0 < i \leqq r_1$ and (2.3) and (2.4) follow.

As t moves through $[0, T]$ the image point under ϕ moves through the square $0 \leqq x < 2\pi$ on straight lines of slope s_1/r_1 with velocity

$$\lambda\sqrt{(1/r_1)^2 + (1/s_1)^2}$$

reflecting from $x = 2\pi$ to $x = 0$ and from $y = 2\pi$ to $y = 0$. The total number of lines is $(r_1 + s_1 - 1)$ and distance traversed is

$$2\pi\sqrt{r_1^2 + s_1^2}.$$

The perpendicular distance between neighboring lines is constant and is equal to

$$\frac{2\pi}{\sqrt{r_1^2 + s_1^2}}.$$

The particular ϕ we have considered was chosen because it satisfies (1.3), as will become evident when we analyze the induced function map in the next section.

Consider the sweep; it is made up of line segments which we number in their order of occurrence; for example, note the numbers on the right most endpoints of Fig. 1 (which features the $s_1 = 2$ and $r_1 = 3$ sweep). Let $k(i)$ be the line number of the i hit on $y = 2\pi$, it occurs at $x_i = (ir/s)2\pi$, while $l(j)$ is line number of j hit on $x = 2\pi$, it occurs at $y_j = (js/r)\ 2\pi$. Suppose for convenience that $r > s$, and let $r = \bar{r}s + r_0$ with $0 \leqq r_0 < s$; then

$$k(i+1) = k(i) + p(i) + \bar{r},$$

where

$$p(i) = \begin{cases} 1, & x_{i+1} - x_i > 0, \\ 2 & \text{otherwise}, \end{cases}$$

$$l(i+1) = l(i) + g(i),$$

$$g(i) = \begin{cases} 1, & y_{i+1} - y_i > 0, \\ 2 & \text{otherwise}. \end{cases}$$

The above relations can be easily verified—see for example the line numbers of Fig. 1.

[1] $((z))$ denotes the fractional part of the real number z.

3. The convolution property. Let F and G have the following Fourier series:

$$F(x, y) = \sum_{n,m} f_{n,m} e^{inx} e^{imy},$$
$$G(x, y) = \sum_{n,m} g_{n,m} e^{inx} e^{imy}; \tag{3.1}$$

then

$$\tilde{F}(t) = \sum_{n,m} f_{n,m} e^{i(n\alpha+m\beta)t},$$
$$\tilde{G}(t) = \sum_{n,m} g_{n,m} e^{i(n\alpha+m\beta)t}. \tag{3.2}$$

Now

$$\widetilde{F * G}(x, y) = \sum_{n,m} g_{n,m} f_{n,m} e^{inx} e^{imy},$$
$$F * G(t) = \sum_{n,m} g_{n,m} f_{n,m} e^{i(n\alpha+m\beta)t}, \tag{3.3}$$

so that

$$\tilde{F} * \tilde{G}(t) = \sum_{n,m} \sum_{l,k} f_{n,m} g_{l,k} e^{i(n\alpha+\beta m)t} \frac{1}{T} \int_0^T e^{i((l-n)\alpha+\beta(k-m))s}\, ds. \tag{3.4}$$

But

$$\frac{1}{T} \int_0^T e^{i((l-n)\alpha+\beta(k-m))s}\, ds = \begin{cases} 0 & \text{if } (l-n)\alpha+\beta(k-m) \neq 0, \\ 1 & \text{if } (l-n)\alpha+\beta(k-m) = 0, \end{cases}$$

so that $(l-n)/s_1 = (m-k)/r_1$ and since r_1 and s_1 are relatively prime, $n-l = ps_1$ and $m-k = -pr_1$ for some integer p. It follows that

$$\tilde{F} * \tilde{G}(t) = \sum_{p} \sum_{n,m} f_{n,m} g_{n-ps_1, m+pr_1} e^{i(n\alpha+m\beta)t}. \tag{3.5}$$

Consequently, the deviation of ϕ from a true homorphism is given by $E(t)$ in

$$\tilde{F} * \tilde{G}(t) = \widetilde{F * G}(t) + E(t) \quad \text{where } E(t) = \sum_{p \neq 0} \sum_{n,m} f_{n,m} g_{n-ps_1, m+pr_1} e^{i(n\alpha+m\beta)t}. \tag{3.6}$$

We remark that the form of the error reminds one of the relation between the Fourier coefficients of a sampled function and the Fourier coefficients of the original function—see [6, p. 29].

4. Evaluation of the error $E(t)$. Despite the explicit form of the error given in (3.6), it seems difficult even when one of the functions, say F, is known, by making assumptions on the function G to obtain a sharp bound on $E(t)$. However, it is clear that as r_1 and s_1 tend to infinity, $E(t)$ tends to zero if F and G are, say, integrable and square integrable along with their derivatives. Motivated by the

physical problem—see [3]—we introduce the theta function

$$\theta(x,q)=\sum_{V=-\infty}^{\infty}\exp\left[-\frac{1}{2q}(x-2\pi v)^2\right] \tag{4.1}$$

and let

$$F(x,y)=\theta(x-\pi,.01)\theta(y-\pi,.1), \tag{4.2}$$

and let $r_1=103$, $s_1=21$. A program was written which, when G is given, evaluates $E(t)$, by finding $\phi(F)*\phi(G)(t)$.

Two cases were considered: $G=1$ and $G(x,y)=\theta(x-\pi,.5)\theta(y-\pi,.5)$. Note that in each of these cases $F*G$ is known analytically. The results we found were that

$$\|E\|_\infty<10^{-4} \quad \text{when } G=1 \quad \text{and} \quad \|E\|_\infty<10^{-5}$$

in the second case. Now

$$\|E(t)\|_\infty=\sup_{0\leqq t\leqq T}|E(t)|.$$

In order to derive a bound on the error, we introduce the sequence a_n, the Fourier coefficients of the function $\frac{1}{2}(x-\pi)^2$ on the interval $(0,2\pi)$; it is easily calculated that

$$a_n=\begin{cases}\pi^2/6, & n=0,\\ 1/n^2, & \text{otherwise.}\end{cases}$$

LEMMA 2. *Let*

$$b_n=\sum_{k=-\infty}^{\infty}a_k a_{n-k}.$$

Then

$$b_n=\begin{cases}\pi^4/20, & n=0,\\ (\pi^2n^2-6)/n^4), & n\neq 0,\end{cases}$$

and further,

$$\sum_{n=1}^{\infty}\frac{1}{n^4}=\frac{\pi^4}{90}. \tag{4.3}$$

Proof. Since the a_n are the Fourier coefficients of $\frac{1}{2}(x-\pi)^2$, by the faulting theorem, the b_n are the Fourier coefficients of $\frac{1}{4}(x-\pi)^4$, and the assertion follows by a simple computation. The evaluation of (4.3) is a consequence of the value of b_0 and the relation between b_0 and the a_n's; of course the last conclusion is well known—see [6, p. 57].

THEOREM 3. *Suppose* $|f_{n,m}|\leqq(K_1/\pi^2)a_n a_m$ *and* $|g_{n,m}|\leqq(K_2/\pi^2)a_n a_m$. *Then*

$$|E(t)|\leqq\frac{K_1K_2\pi^4}{45\cdot s_1^2r_1^2}\leqq\frac{K_1K_2\pi^6}{11T^2\lambda^2}$$

with T the period of the map.

Proof. Since

$$E(t) = \sum_{\lambda \neq 0} \sum_m \sum_n f_{n,m} g_{n+\lambda s_1, m-\lambda r_1} e^{i(\alpha n + \beta m)t}$$

it follows, using the triangle inequality, that

$$|E(t)| \leqq \sum_{\lambda \neq 0} \sum_m \sum_n a_n a_m a_{n+\lambda s_1} a_{m-\lambda r_1} \frac{K_1 K_2}{\pi^4}$$

or

$$|E(t)| \leqq \sum_{\lambda \neq 0} \left\{ \sum_n a_n a_{-\lambda s_1 - n} \right\} \left\{ \sum_m a_m a_{\lambda r_1 - m} \right\} \frac{K_1 K_2}{\pi^4}.$$

Now, using the definition of b_n and Lemma 2, it follows that

$$|E(t)| \leqq \sum_{\lambda \neq 0} \frac{b_{-\lambda s_1}}{\pi^2} \frac{b_{\lambda r_1}}{\pi^2} K_1 K_2 \leqq \frac{K_1 K_2}{r_1^2 s_1^2} 2 \sum_{\lambda=1}^{\infty} \frac{1}{\lambda^4},$$

and the assertion follows.

5. Remarks. This theorem resulted from conversations with Dr. Richard Edwards of Aerospace Engineering Department, University of Southern California. It is clear that this theorem is one of a class where growth conditions on the Fourier coefficients of F and G induce a bound on the error, although in general the error would take a more complex form. For a periodic function of one variable, there are well-known mild conditions in order that its Fourier coefficients are $O(1/n^2)$, for example, if the function possesses a derivative of bounded variation—see [7]. The assumption of the bound as a product of functions of m and n is more stringent. It is easy to see, if $|f_{n,m}| \leqq K_1/\pi^2\, a_m a_n$ and $|g_{n,m}|$ goes to zero faster, the error bound does not improve, because the asymptotic behavior of the convolution of sequences $f * g(n, m)$ is equivalent to that of its constituent $f_{n,m}$ as $g_{0,0}$ is positive.

If one assumes that $|g_{m,n}| = |g_{-m,-n}|$, then

$$|E(t)| \leqq \sum_{\lambda=1}^{\infty} \{k_{-\lambda s_1, \lambda r_1} + k_{\lambda s_1, -\lambda r_1}\} \quad \text{with } k_{l,m} = \sum_{\alpha\beta} |f_{\alpha,\beta}||g_{l-\alpha, m-\beta}|,$$

or $k_{l,m}$ are the Fourier coefficients of $F_1 G_1$ with $F_1\{G_1\}$ the function with Fourier coefficients $|f_{m,n}|$, $\{|g_{m,n}|\}$ respectively. It is easy to see that

$$k_{l,m} \geqq |f_{0,0}||g_{l,m}|, \qquad k_{l,m} \geqq |g_{0,0}||f_{l,m}|,$$

so that $k_{l,m}$ cannot approach zero faster than the slower $f_{l,m}$ and $g_{l,m}$ as long as $f_{0,0}$ and $g_{0,0}$ do not vanish.

6. Conclusions and acknowledgments. We have investigated the approximate evaluation of multi-dimensional convolutions by a one-dimensional convolution. The numerical results were derived by the use of a program written by Luis Basañez and Pedro Brunet of the Polytechnic University of Barcelona, U.P.B. The results of this paper were presented at a seminar on parallel processing the nonlinear filtering, and we thank the participants for their comments and

criticism; they were R. Huber, J. Pages, L. Basañez, P. Brunet of U.P.B., C. Hecht of Aerospace Corporation, and D. S. Miller of T.R.W. Systems.

In a future paper the application of the sweep mapping, described here, to nonlinear filtering will be investigated in detail.

REFERENCES

[1] E. DIEULESAINT AND D. ROYER, *Ondes Elastique dans Les Solides*, Masson, Paris, 1974.

[2] W. D. SQUIRE, H. J. WHITEHOUSE AND J. M. ALSUP, *Linear signal processing and ultrasonic transversal filters, IEEE Trans. Microwave Theory Tech.*, 17 (1969), pp. 1020–1040.

[3] R. S. BUCY, C. HECHT AND K. D. SENNE, *An engineer's guide to building non-linear filters*, Frank J. Seiler Res. Lab., U.S.A.F. Acad. Rep. SRL-TR-72-0004, 1972.

[4] V. I. ARNOLD AND A. AVEZ, *Théorie Ergodic des Systemes dynamiques*, Gautier-Villars, Paris, 1967.

[5] R. W. HAMMING, *Introduction to Applied Numerical Analysis*, McGraw-Hill, New York, 1972.

[6] L. B. W. JOLLEY, *Summation of Series*, Dover, New York, 1961.

[7] E. W. HOBSON, *Theory of Functions of a Real Variable*, vol. 2, Dover, New York, 1957; reprint of 3rd ed., 1927.

SIAM J. MATH. ANAL.
Vol. 8, No. 4, August 1977

UNIFORM L^1 BEHAVIOR FOR AN INTEGRODIFFERENTIAL EQUATION WITH PARAMETER*

KENNETH B. HANNSGEN†

Abstract. For a family of real integrodifferential equations with nonnegative, nonincreasing, convex, strongly positive convolution kernel $\lambda[c+a(t)]$, depending on the parameter λ, $0<\Lambda\leqq\lambda<\infty$, we show that the solutions $u(t,\lambda)$, normalized by the initial condition $u(0,\lambda)=1$, satisfy $\sup_\lambda |u(t,\lambda)|=u^0(t)$ where $u^0\in 1(0,\infty)$ and $u^0(\infty)=0$. A result of the same type holds for u_t. The proof uses a Fourier integral representation for the solution. Applications to equations in Hilbert space and to scalar equations with a complex parameter are given.

1. Introduction.

We consider the problem

$$u'(t,\lambda)+\lambda\int_0^t [c+a(t-s)]u(s,\lambda)\,ds=0, \qquad u(0,\lambda)=1, \tag{1.1}$$

(primes denote differentiation with respect to the first variable, t in this case), where c is a fixed nonnegative constant, λ is a real or complex parameter, and

$$\text{(H)}\quad \begin{aligned} &a\in C[0,\infty);\ a(t)\ \textit{is negative, nonincreasing, and convex};\\ &a(\infty)=0,\ \textit{but}\ a(t)\not\equiv 0. \end{aligned}$$

Results of S. I. Grossman and R. K. Miller [2] and of D. F. Shea and S. Wainger [10] show that

$$\int_0^\infty |u^{(j)}(t,\lambda)|\,dt<\infty, \qquad j=0,1, \quad \lambda>0, \tag{1.2}$$

if in addition to (H) we assume

$$i\tau+\lambda[A(\tau)+c/i\tau]\neq 0 \qquad (\tau>0), \tag{1.3}$$

where $A(\tau)$ is the Fourier transform of $a(t)$:

$$A(\tau)=\int_0^\infty e^{-i\tau t}a(t)\,dt. \tag{1.4}$$

This will certainly be the case if there is a constant η such that

$$\operatorname{Re} A(\tau)\geqq\frac{\eta}{1+\tau^2}>0 \qquad (\tau>0) \tag{1.5}$$

(a is then strongly positive). Fix $\Lambda>0$ and define

$$u^0(t)=\sup_{\Lambda\leqq\lambda<\infty}|u(t,\lambda)|,$$

$$u^1(t)=\sup_{\Lambda\leqq\lambda<\infty}\lambda^{-1/2}[1+\log(\lambda/\Lambda)]^{-1}|u'(t,\lambda)|.$$

* Received by the editors September 4, 1975.

† Department of Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061. This work was supported in part by the National Science Foundation under Grant MPS 74-06403 A01.

THEOREM 1. *Let* (H) *hold, and suppose* $\int_0^\infty a(t)\,dt<\infty$ *and* da' *is not a purely singular measure. Then*

$$\int_0^\infty u^j(t)\,dt<\infty \qquad (j=0,1), \tag{1.6}$$

$$u^j(t)\to 0 \qquad (t\to\infty,\quad j=0,1). \tag{1.7}$$

THEOREM 2. *Let* (H) *hold, and suppose* $-a'(t)$ *is convex. Then* (1.6) *and* (1.7) *hold.*

Clearly da' is not purely singular in Theorem 2. Thus [9, Cor. 2.1 and 2.2] in both theorems (1.5) holds.

THEOREM 3. *Let* (H) *hold, and suppose* $c=0$ *and* $a'(0+)>-\infty$. *Assume that* $\operatorname{Re} A(\tau)>0$ $(\tau>0)$ *but that no positive* η *exists for which* (1.5) *holds. Then*

$$\limsup_{\lambda\to\infty}\int_0^\infty |u(t,\lambda)|\,dt=\infty.$$

Before discussing the hypotheses of these theorems, we indicate two applications of (1.6) and (1.7).

First, assume (H) holds and let L be a self-adjoint linear operator densely defined on a Hilbert space $\mathbf{H}$ with spectral decomposition

$$L\mathbf{x}=\int_\Lambda^\infty \lambda\,dE_\lambda\mathbf{x} \qquad (\mathbf{x}\in\mathscr{D}(L)).$$

In previous work [5], [8], we have shown that the formula

$$R(t)=\int_\Lambda^\infty u(t,\lambda)\,dE_\lambda$$

defines a bounded operator on $\mathbf{H}$. $R(t)$ is the resolvent kernel for the equation

$$\mathbf{y}'(t)+\int_0^t a(t-s)L\mathbf{y}(s)\,ds=\mathbf{f}(t);$$

that is, under reasonably mild assumptions,

$$\mathbf{y}(t)=R(t)\mathbf{y}(0)+\int_0^t R(t-s)\mathbf{f}(s)\,ds.$$

If (1.6) and (1.7) hold,

$$\int_0^\infty \|R(t)\|\,dt<\infty \quad\text{and}\quad \|R(t)\|\to 0 \qquad (t\to\infty).$$

Second, for complex λ consider the series

$$U(t,\lambda)=\sum_{n=0}^\infty U_n(t)(\lambda-\mu)^n, \qquad \mu\geqq\Lambda, \tag{1.8}$$

where $U_0(t) = u(t, \mu)$,

$$U_{n+1}(t) = \frac{1}{\mu}\int_0^t u'(s, \mu)U_n(t-s)\,ds \qquad (n \geqq 0). \tag{1.9}$$

Because of (1.2), $U_n(t) \to 0$ as $t \to \infty$. Assuming (1.6) and using the convolution inequality for L^1 norms, we see that for a suitable constant $B = B(\Lambda)$,

$$\int_0^\infty |U_n(t)|\,dt \leqq \{B\mu^{-1/2}[1+\log(\mu/\Lambda)]\}^n.$$

It is known (see (1.10) below) that $|u(t, \mu)| \leqq 1$; therefore we also have

$$\sup_{t \geqq 0} |U_n(t)| \leqq \{B\mu^{-1/2}[1+\log(\mu/\Lambda)]\}^n.$$

Thus if $\lambda \in D_\mu = \{\lambda \mid B|\lambda - \mu| < \mu^{1/2}/[1+\log(\mu/\Lambda)]\}$, (1.8) is a convergent series in the space of bounded continuous integrable functions of t with the norm $\|f\| = \sup_{0 \leqq t < \infty} |f(t)| + \int_0^\infty |f(t)|\,dt$, and $\lim_{t\to\infty} U(t, \lambda) = 0$.

COROLLARY 1.1. *If $\lambda \in D_\mu$ for some $\mu \geqq \Lambda$, then $U(t, \lambda)$ is the solution $u(t, \lambda)$ of* (1.1).

This corollary, which will be proved in § 7 by comparison of Laplace transforms, shows that $u(t, \lambda) \to 0$ $(t \to \infty)$ and $\int_0^\infty |u(t, \lambda)|\,dt < \infty$ for $\lambda \in \cup D_\mu$, $\mu \geqq \Lambda$.

In [8], we proved (1.6) and (1.7) for $j = 0$ when $a(t)$ is continuous on $[0, \infty)$ and completely monotonic on $(0, \infty)$; Theorem 2 contains this result.

Nohel and Shea [9, § 4] show that functions $a(t)$ exist satisfying the hypotheses of Theorem 3.

If $c = 0$, $a(t) = e^{-t}$, (1.1) reduces to the initial value problem

$$u'' + u' + \lambda u = 0, \qquad u(0, \lambda) = 1, \qquad u'(0, \lambda) = 0$$

with solution

$$u(t, \lambda) = e^{-t/2}\left[\cos \mu t + \frac{1}{2\mu}\sin \mu t\right] \qquad (\lambda > \tfrac{1}{4}),$$

where $\mu = (4\lambda - 1)^{1/2}/2$. Using Fourier transforms we see that

$$\int_0^\infty |u(t, \lambda)|\,dt \geqq \sup_{-\infty < \tau < \infty} |\hat{u}(\tau, \lambda)| \geqq |\hat{u}(\sqrt{\lambda - 1}, \lambda)| = 1.$$

Similarly,

$$\int_0^\infty |u'(t, \lambda)|\,dt \geqq \lambda |A(\sqrt{\lambda - 1})\hat{u}(\sqrt{\lambda - 1}, \lambda)| = |1 - i\sqrt{\lambda - 1}|.$$

In this sense, (1.6) is sharp for $j = 0$; for $j = 1$, we have not determined whether the term $\log(\lambda/\Lambda)$ must be included in u^1. It will be evident from the proof that this log term is not needed for (1.7).

In [4] we showed that (H) implies

$$|u(t, \lambda)| \leqq 1 \qquad (0 \leqq t < \infty, \quad 0 < \lambda < \infty) \tag{1.10}$$

($|u| \leqq \sqrt{2}$ is given in [4], due to an arithmetic error in the proof; $c = 0$ is assumed in [4], but $a(\infty) = 0$ is not used in the proof of (1.10).) In [6] we showed that (H) implies

$$\sup_{\Lambda \leqq \lambda < \infty,\, t \geqq 0} \left| \int_0^t u(s, \lambda)\, ds \right| < \infty$$

but if we replace Λ by 0 this sup is infinite.

2. Integral representations. We assume without further mention that $c + a(t)$ has been rescaled if necessary so that $\Lambda = 1$. We define $D(\tau) \equiv D(\tau, \infty) = A(\tau) + c/i\tau$ and $D(\tau, \lambda) = D(\tau) + i\tau/\lambda$ ($\lambda \geqq 1$, $\tau > 0$). We define $a'(t)$ where necessary so that $a'(t+) = a'(t)$ ($0 \leqq t < \infty$).

LEMMA 2.1. *If* (H) *and* (1.3) *hold*

(i) $A(\tau)$ *is analytic in* $\Omega = \{\operatorname{Im} \tau < 0\}$ *and continuous in* $\bar{\Omega} \backslash \{0\}$; *if* $a \in L^1(0, \infty)$, A *is continuous in* $\bar{\Omega}$.

(ii) $A(\tau) = O(\tau^{-1})$ ($|\operatorname{Re} \tau| \to \infty$), *uniformly in* $\{0 \leqq -\operatorname{Im} \tau < \infty\}$.

(iii) $\tau A(\tau) \to 0$ ($\tau \to 0$, $\tau \in \bar{\Omega}$).

(iv) *If* $\int_0^\infty a(t)\, dt = \infty$, *then* $[A(\tau)]^{-1} \to 0$ ($\tau \to 0$, $\tau \in \bar{\Omega}$).

(v) *The representations*

$$\pi u(t, \lambda) = \frac{1}{\lambda} \int_0^\infty \operatorname{Re} \left\{ \frac{e^{i\tau t}}{D(\tau, \lambda)} \right\} d\tau, \tag{2.1}$$

$$\pi u'(t, \lambda) = -\int^\infty \operatorname{Re} \left\{ \frac{e^{i\tau t} D(\tau)}{D(\tau, \lambda)} \right\} d\tau \tag{2.2}$$

hold for $t > 0$. (The integral in (2.1) is improper at $\tau = \infty$.)

Proof. The proofs of (i) through (iv) are fairly straightforward and will be found in [3, § 2]. Equations (2.1) and (2.2) have similar proofs; we outline the latter, since (2.1) is proved in [3, § 3].

Using standard a priori estimates for linear Volterra equations [1, § 7.6], together with (1.1) and the complex inversion formula for Laplace transforms, one sees that for $t > 0$

$$2\pi i u'(t, \lambda) = -\int_C e^{zt} [D(-iz)/D(-iz, \lambda)]\, dz, \tag{2.3}$$

where C is the straight line $\{\sigma + i\tau, -\infty < \tau < \infty\}$ for sufficiently large positive σ. A contour shift, justified by (i) through (iv) allows us to change (2.3) to

$$2\pi u'(t, \lambda) = -\int_{-\infty}^\infty e^{i\tau t} [D(\tau)/D(\tau, \lambda)]\, d\tau,$$

and (2.2) follows by a change of variables.

Fix $t_1>0$ such that $a(t_1)>0$, and set $\rho=6/t_1$. For $t>0$ define

$$u_1(t)=\int_0^\rho \operatorname{Re}\left\{\frac{e^{i\tau t}}{D(\tau)}\right\}dt-\operatorname{Re}\left\{\frac{e^{i\rho t}}{itD(\rho)}\right\},$$

$$u_2(t,\lambda)=\frac{-1}{\lambda^2}\int_0^\rho \operatorname{Re}\left\{\frac{e^{i\tau t}i\tau}{D(\tau)D(\tau,\lambda)}\right\}d\tau+\operatorname{Re}\left\{\frac{\rho e^{i\rho t}}{\lambda^2 tD(\rho)D(\rho,\lambda)}\right\},$$

$$u_3(t,\lambda)=\frac{1}{\lambda}\int_\rho^\infty \operatorname{Re}\left\{\frac{e^{i\tau t}}{D(\tau,\lambda)}\right\}d\tau+\frac{1}{\lambda}\operatorname{Re}\left\{\frac{e^{i\rho t}}{itD(H,\lambda)}\right\},$$

$$v_1(t,\lambda)=-\int_0^\rho \operatorname{Re}\left\{\frac{e^{i\tau t}D(\tau)}{D(\tau,\lambda)}\right\}d\tau+\operatorname{Re}\left\{\frac{e^{i\rho t}D(\rho)}{itD(\rho,\lambda)}\right\},$$

$$v_2(t,\lambda)=-\int_\rho^\infty \operatorname{Re}\left\{\frac{e^{i\tau t}D(\tau)}{D(\tau,\lambda)}\right\}d\tau-\operatorname{Re}\left\{\frac{e^{i\rho t}D(\rho)}{itD(\rho,\lambda)}\right\}.$$

COROLLARY 2.1. *If* (H) *and* (1.3) *hold, then*

$$\pi u(t,\lambda)=\lambda^{-1}u^1(t)+u_2(t,\lambda)+u_3(t,\lambda),$$

$$\pi u'(t,\lambda)=v_1(t,\lambda)+v_2(t,\lambda).$$

The proof is immediate.

We separate $A(\tau)$ into real and imaginary parts $A(\tau)=\varphi(\tau)-i\tau\theta(\tau)$ $(\tau>0)$.

LEMMA 2.2. *Suppose* (H) *holds and* $a'(0)>-\infty$. *Then* $A(\tau)$ *is differentiable* $(\tau>0)$ *and*

$$\frac{1}{2\sqrt{2}}\int_0^{1/\tau} a(t)\,dt\leqq|A(\tau)|\leqq 4\int_0^{1/\tau} a(t)\,dt \qquad (\tau>0), \tag{2.4}$$

$$|A'(\tau)|\leqq 40\int_0^{1/\tau} ta(t)\,dt\leqq 40a(0)/\tau^2 \qquad (\tau>0), \tag{2.5}$$

$$\theta(\tau)\geqq a(t_1)/2\tau^2 \qquad (\tau\geqq\rho), \tag{2.6}$$

$$\theta'(\tau)<0 \quad (\tau>0) \quad \text{and} \quad \theta'(\tau)\leqq\frac{-a(t_1)}{2\tau^3} \qquad (\tau\geqq\rho). \tag{2.7}$$

For each $\lambda>0$ *there is at most one number* $\omega=\omega(\lambda)$ *on the interval* $\{\rho\leqq\omega<\infty\}$ *such that*

$$\theta(\omega)+\frac{c}{\omega^2}=\frac{1}{\lambda}. \tag{2.8}$$

We define $\omega(\lambda)=\rho$ *if no such* ω *exists. Then* $\omega(\lambda)$ *is a continuous, nondecreasing function on* $\{\lambda>0\}$ *and*

$$a(t_1)\lambda/2\leqq\omega^2(\lambda)\leqq\max\{\rho^2,[4a(0)+c]\lambda\}. \tag{2.9}$$

Proof. Inequalities (2.4) and (2.5) are the conclusions of [10, Lemma 1] (the assumption $b\notin L^1(0,\infty)$ of that lemma is not used here). For the other conclusions,

we integrate twice by parts in (1.4) (following [10]) to obtain

$$A(\tau)=\tau^{-2}\int_0^\infty (1-i\tau t-e^{-i\tau t})\,da'(t). \tag{2.10}$$

Here we have used several consequences of (H); in particular, $ta'(t)\to 0$ $(t\to\infty)$ and

$$\int_T^\infty t\,da'(t)=a(T)-Ta'(T) \qquad (T\geqq 0) \tag{2.11}$$

so that the integral in (2.10) converges absolutely. Thus

$$\tau^2\theta(\tau)=\int_0^\infty \left(t-\frac{\sin\tau t}{\tau}\right) da'(t)>0,$$

so if $\tau\geqq\rho=6/t_1$, (2.11) shows that $2\tau^2\theta(\tau)\geqq\int_{t_1}^\infty t\,da'(t)\geqq a(t_1)$ and (2.6) holds. Similarly $\tau^3\theta'(\tau)=\int_0^\infty t\psi(\tau t)\,da'(t)$, where $\psi(x)=-2-\cos x+3\sin x/x$. Now $\psi(x)<0$ $(x>0)$ (this is obvious for $x\geqq\pi$; note that $x\psi(x)$ vanishes, together with its first two derivatives, at $x=0$, while $[x\psi(x)]'''<0$, $0<x<\pi$). Thus $\theta'(\tau)<0$ $(\tau>0)$ and if $\tau\geqq\rho$, $2\tau^3\theta'(\tau)\leqq-\int_{t_1}^\infty t\,da'(t)\leqq-a(t_1)$, so (2.7) holds. The uniqueness and monotonicity of $\omega(\lambda)$ follow easily from (2.7) and (2.8). The implicit function theorem shows that (2.8) defines a differentiable function $\tilde{\omega}(\lambda)$ locally, so $\rho(\lambda)=\max\{\rho,\tilde{\omega}(\lambda)\}$ is continuous. Finally, (2.9) is an immediate consequence of (2.4), (2.6), and (2.8).

COROLLARY 2.2. *If* (H) *and* (1.3) *hold and* $a'(0)>-\infty$, *then* $\int_0^\rho |D'(\tau,\lambda)/D^2(\tau,\lambda)|\,d\tau<\infty$ $(1\leqq\lambda\leqq\infty)$ *and*

$$u_1(t)=\operatorname{Re}\frac{1}{it}\int_0^\infty e^{i\tau t}\frac{D'(\tau)}{D^2(\tau)}\,d\tau, \tag{2.12}$$

$$t\lambda^2u_2(t,\lambda)=\operatorname{Re}\int_0^\rho e^{i\tau t}\left[\frac{1}{D(\tau)D(\tau,\lambda)}-\frac{\tau D'(\tau)}{D^2(\tau)D(\tau,\lambda)}-\frac{\tau D'(\tau,\lambda)}{D(\tau)D^2(\tau,\lambda)}\right]d\tau, \tag{2.13}$$

$$u_3(t,\lambda)=\operatorname{Re}\frac{1}{it\lambda}\int_\rho^\infty e^{i\tau t}\frac{D'(\tau,\lambda)}{D^2(\tau,\lambda)}d\tau, \tag{2.14}$$

$$v_1(t,\lambda)=\operatorname{Re}\frac{1}{t\lambda}\int_0^\rho e^{i\tau t}\left[\frac{\tau D'(\tau,\lambda)}{D^2(\tau,\lambda)}-\frac{1}{D(\tau,\lambda)}\right]d\tau, \tag{2.15}$$

$$v_2(t,\lambda)=\operatorname{Re}\frac{1}{it}\int_\rho^\infty e^{i\tau t}\left[\frac{D'(\tau)}{D(\tau,\lambda)}-\frac{D(\tau)D'(\tau,\lambda)}{D^2(\tau,\lambda)}\right]d\tau. \tag{2.16}$$

Proof. Integrate by parts. Vanishing of the boundary terms at $\tau=0,\infty$ and absolute convergence of the integrals are assured by Lemmas 2.1 and 2.2. (Note: if the factors λ and i of (2.15) and (2.16) are brought under the integral signs, the two integrands are the same.)

3. Proof of Theorem 3. Let $\hat{u}(\tau,\lambda)=\int_0^\infty e^{-i\tau t}u(t,\lambda)\,dt$. Then $\sup_{-\infty<\tau<\infty}|\hat{u}(\tau,\lambda)|\leqq\int_0^\infty|u(t,\lambda)|\,dt<\infty$. From (1.1), $\hat{u}(\tau,\lambda)=[\lambda D(\tau,\lambda)]^{-1}$ in $\{\operatorname{Im}\tau<0\}$ and by continuity for real τ as well. Let n be a positive integer and

choose $\tau_n > \rho + n$ such that $n\tau_n^2\varphi(\tau_n) \leqq 1$. Since $\omega(\lambda)$ is continuous with $\omega(0+) = \rho$ and $\omega(\infty) = \infty$, there exists $\lambda = \lambda_n$ such that $\omega(\lambda) = \tau_n$; that is $\theta(\tau_n) = \lambda_n^{-1}$. By (2.9), $\lambda_n \geqq \tau_n^2/(4a(0)) \to \infty$ $(n \to \infty)$. Moreover $\int_0^\infty |u(t, \lambda_n)|\, dt \geqq |\hat{u}(\tau_n, \lambda_n)| = 1/\lambda_n\varphi(\tau_n) \geqq n\tau_n^2/\lambda_n = n\omega^2(\lambda_n)/\lambda_n \geqq na(t_1)/2 \to \infty$ $(n \to \infty)$. This proves Theorem 3.

4. Uniform estimates and approximating sequences. The following considerations are needed to deal with cases where $a'(0)$ or $a''(0)$ is infinite. The letter M will denote a finite positive a priori constant which can be chosen uniformly in λ, $1 \leqq \lambda < \infty$, and uniformly over any class $\mathscr{A}$ of kernels $a(t)$ for which (4.1) through (4.3) hold.

(4.1) *The hypotheses of Theorem* 1 (*Theorem* 2) *hold, and* $a'(0) > -\infty [a'(0) > -\infty$ *and* $a''(0+) < \infty]$ *for all* a *in* $\mathscr{A}$.

(4.2) *There are fixed positive numbers* t_1, α, β *such that* $a(t_1) = \alpha$, $a(0) = \beta$ (*all* a *in* $\mathscr{A}$), *and the number* η *of* (1.5) *may be chosen the same for all* a *in* $\mathscr{A}$.

(4.3) *In the case of Theorem* 1, *the set of numbers* $\int_0^\infty a(t)\, dt$, $a \in \mathscr{A}$, *is bounded.*

Such a class $\mathscr{A}$ is constructed as follows. For each positive integer n, let $b_n(t)$ be any function with the properties (i) $b_n \in C^3[0, 1/n]$, (ii) $(-1)^k b_n^{(k)}(t) \geqq 0$ $(0 \leqq t \leqq 1/n, k = 0, 1, 2, 3)$, (iii) $b_n(0) = a(0)$, (iv) $b_n(1/n) = a(1/n)$, (v) $b_n'(1/n) = a'(1/n)$, (vi) $b_n''(1/n) = a''(1/n+)$ if the latter exists. Let $a_n(t) = b_n(t)$ $(0 \leqq t < 1/n)$, $a_n(t) = a(t)$ $(t \geqq 1/n)$, and set $\mathscr{A}_N = \{a_n | n \geqq N\}$. The Fourier transform of a_n will be denoted $A_n = \varphi_n - i\tau\theta_n$.

LEMMA 4.1. *Let* $a(t)$ *satisfy the hypotheses of Theorem* 1 [*Theorem* 2], *and define* $\mathscr{A}_N$ *as above. For sufficiently large* N, *the set* $\mathscr{A} = \mathscr{A}_N$ *has properties* (4.1), (4.2) *and* (4.3).

Proof. The only nontrivial assertion is that η may be chosen uniformly over A_N. Using the Riemann–Lebesgue lemma, choose positive numbers τ_0, T, γ so that

$$\int_T^\infty (1 - \cos \tau t)\, d\nu(t) \geqq \gamma \qquad (\tau \geqq \tau_0),$$

where $d\nu$ is the absolutely continuous part of da'. Choose $N > T^{-1}$ with $N^{-1}[a(0) - a(1/N)] < \min_{0 \leqq \tau \leqq \tau_0} \varphi(\tau) \equiv 2\delta$. Then for $n \geqq N$, $\varphi_n(\tau) \geqq \delta$ $(0 \leqq \tau \leqq \tau_0)$, and from (2.10) we see that $\varphi_n(\tau) \geqq \gamma\tau^{-2}$ $(\tau \geqq \tau_0)$. Thus (1.5) holds with $\eta = \min\{\delta, \gamma\}$ for all a in $\mathscr{A}_N$.

5. Proof of Theorems 1 and 2. The following two lemmas, which will be proved in § 6, give the main estimates of the proof. The constant M has the special meaning discussed in § 4.

LEMMA 5.1. *Under the hypotheses of Theorem* 1, *if* $a'(0) > -\infty$, *the estimates*

(5.1) $$|u_2(t, \lambda)| \leqq M(1 + \log t)/t^2,$$

(5.2) $$|u_3(t, \lambda)| \leqq M/t^2,$$

(5.3) $$|v_1(t, \lambda)| \leqq M(1 + \log t)/t^2,$$

(5.4) $$|v_2(t, \lambda)| \leqq M\sqrt{\lambda}/t^2$$

are valid for $t \geqq 1$. *The same estimates are valid for* $t \geqq 1$ (*even with the term* $\log t$ *omitted in* (5.1) *and* (5.3)) *if the hypotheses of Theorem* 2 *hold and* $|a'(0)| + a''(0+) < \infty$.

LEMMA 5.2. *If* (H) *and* (1.5) *hold and* $a'(0) > -\infty$, *then*

$$|u'(t, \lambda)| \leqq M\lambda^{1/2}(1 + \log \lambda) \qquad (0 \leqq t \leqq 1). \tag{5.5}$$

Then (5.3), (5.4) and (5.5), together with Corollary 2.1, show that (1.6) and (1.7) hold for $j = 1$ provided $a'(0) > -\infty$ and, in Theorem 2, $a''(0+) < \infty$. With the same proviso, (5.1), (5.2), (1.2), and Corollary 2.1 imply that $u_1(t) \to 0$ $(t \to \infty)$ and $\int_1^\infty |u_1(t)|\, dt < \infty$. Hence (1.7) is valid for $j = 0$ and $\int_1^\infty u^0(t)\, dt < \infty$. Since (1.10) holds, (1.6) is established for $j = 0$, and the proof is complete.

In the general case, where a' or a'' may be infinite at $t = 0$, consider the set $\mathscr{A} = \mathscr{A}_N = \{a_n\}$ of § 4. To each a_n there corresponds a solution $u_n(t, \lambda)$ and functions $u_j^n(t, \lambda)$, $v_j^n(t, \lambda)$ as in Corollary 2.1. A final lemma, also proved in § 6, will allow us to finish the proof.

LEMMA 5.3. *Fix* $\lambda \geqq 1$, $t > 0$. *Let* $a(t)$ *satisfy the hypotheses of Theorem* 1 *or Theorem* 2, *with* $\{a_n\}$ *as in* § 4. *Then* $u_n'(t, \lambda) \to u'(t, \lambda)$, $u_j^n(t, \lambda) \to u_j(t, \lambda)$ $(j = 2, 3)$, $v_j^n(t, \lambda) \to v_j(t, \lambda)$ $(j = 1, 2)$ *as* $n \to \infty$.

By Lemma 4.1, the estimates of Lemmas 5.1 and 5.2 hold for u_2^n, u_3^n, v_1^n, v_2^n, and u_n' with a constant M independent of n. Letting $n \to \infty$, we see from Lemma 5.3 that (5.1) through (5.5) hold even in the general case of Theorems 1 and 2, and we can complete the proof as above.

6. Proofs of Lemmas 5.1, 5.2 and 5.3.

LEMMA 6.1. (i) *If the hypotheses of Theorem* 1 *hold and* $a'(0) > -\infty$, *then* $A(\tau)$ *is twice continuously differentiable and*

$$\begin{aligned} |A''(\tau)| &\leqq 600\left(\int_0^{1/\tau} t^2 a(t)\, dt + \tau^{-2}\int_{1/\tau}^\infty a(t)\, dt\right) \\ &\leqq 600A(0)/\tau^2 \qquad (\tau > 0). \end{aligned} \tag{6.1}$$

(ii) *If the hypotheses of Theorem* 2 *hold and* $|a'(0)| + a''(0+) < \infty$, *then* $A(\tau)$ *is twice continuously differentiable and*

$$|A''(\tau)| \leqq 6000 \int_0^{1/\tau} t^2 a(t)\, dt \qquad (\tau > 0). \tag{6.2}$$

Proof. (i) Integration by parts shows that $\int_0^\infty t^2\, da'(t)\, dt < 2A(0)$. Then we may differentiate twice in (2.10) to obtain

$$A''(\tau) = \tau^{-4}\int_0^\infty J(-\tau t)\, da'(t), \tag{6.3}$$

$J(x) = 6[1 + ix - \exp(ix)] - 4ix[1 - \exp(ix)] - x^2 \exp(ix)$. We have $|J(x)| \leqq 19(1 + x^2)$ $(x \geqq 0)$ and $|J(x)| \leqq x^4$ $(0 \leqq x \leqq 1)$. Thus

$$|A''(\tau)| \leqq \int_0^{1/\tau} t^4\, da'(t) + 20\tau^{-4}\int_{1/\tau}^\infty (1 + \tau^2 t^2)\, da'(t). \tag{6.4}$$

Note that

$$\int_0^{1/\tau} t^4\, da'(t) - \tau^{-4}a'(1/\tau) + 4\tau^{-3}a(1/\tau) = 12\int_0^{1/\tau} t^2 a(t)\, dt, \tag{6.5}$$

$$\int_{1/\tau}^{\infty} da'(t) = -a'(1/\tau). \tag{6.6}$$

For $R > 1/\tau$, integration by parts and (H) show that

$$\tau^{-2}\int_{1/\tau}^{R} t^2\, da'(t) \leqq -\tau^{-4}a'(1/\tau) + 2\tau^{-3}a(1/\tau) + 2\tau^{-2}\int_{1/\tau}^{R} a(t)\, dt,$$

so, using (6.5), we see that

$$\tau^{-2}\int_{1/\tau}^{\infty} t^2\, da'(t) \leqq 18\int_0^{1/\tau} t^2 a(t)\, dt + 2\tau^{-2}\int_{1/\tau}^{\infty} a(t)\, dt. \tag{6.7}$$

Combining (6.3) through (6.7), we get (6.1). For (ii), integration by parts shows that $\int_0^{\infty} t^2\, da''(t) > -2a(0)$. Note that $t^2 a''(t) \to 0$ as $t \to \infty$. We integrate by parts in (2.10) to find that

$$A(\tau) = i\tau^{-3}\int_0^{\infty} [e^{-i\tau t} - 1 + i\tau t + \tau^2 t^2/2]\, da''(t),$$

so that

$$A''(\tau) = i\tau^{-5}\int_0^{\infty} K(-\tau t)\, da''(t),$$

where $K(x) = 12[\exp(ix) - 1 - ix - (ix)^2/2] - 6[ix\exp(ix) - ix - (ix)^2] + [(ix)^2 \cdot \exp(ix) - (ix)^2]$. Then $|K(x)| \leqq 36(1+x^2)$ $(x \geqq 0)$ and $|K(x)| \leqq u^5$ $(0 \leqq u \leqq 1)$. Since $\int_0^{1/\tau} t\, da''(t) - a''(1/\tau)/\tau^5 + 5a'(1/\tau)/\tau^4 - 20a(1/\tau)/\tau^3 = -60\int_0^{1/\tau} t^2 a(t)\, dt$, $\int_{1/\tau}^{\infty} da''(t) = -a''(1/\tau)$, and $\tau^2\int_{1/\tau}^{\infty} t^2\, da''(t) = -a''(1/\tau) + 2\tau a'(1/\tau) - 2\tau^2 a(1/\tau)$, we see as above that (6.2) holds. In this proof we have followed the method of [10, Lemma 1].

Proof of Lemma 5.1. Integration of (2.7) from $\omega(\lambda)$ to τ, together with (2.9), shows that

$$\begin{aligned} |\operatorname{Im} D(\tau,\lambda)| &\geqq \frac{a(t_1)|\omega(\lambda)-\tau|[\omega(\lambda)+\tau]}{4\tau\omega^2(\lambda)} \\ &\geqq |\omega(\lambda)-\tau|[\omega(\lambda)+\tau]/\tau\lambda M. \end{aligned} \tag{6.8}$$

Integrating by parts in (2.14) (using Lemma 2.2) we see that

$$\begin{aligned} \lambda t^2 u_3(t,\lambda) = \operatorname{Re}\Big\{ & e^{i\rho t} D'(\rho,\lambda)/D^2(\rho,\lambda) \\ & + \int_{\rho}^{\infty} e^{i\tau t}\left[\frac{D''(\tau)}{D^2(\tau,\lambda)} - \frac{2[D'(\tau,\lambda)]^2}{D^3(\tau,\lambda)}\right] d\tau\Big\}. \end{aligned} \tag{6.9}$$

Let $\varepsilon = \min\{\rho/2, \sqrt{a}(t_1)/4\}$. We use (2.5) and (6.1) or (6.2) to estimate the numerators in (6.9); inequality (6.8) gives an estimate for the denominators,

except on $\{|\tau-\omega(\lambda)|<\varepsilon\}$ and at $\tau=\rho$ where we use (1.5); inequality (2.9) permits us to compare $\omega(\lambda)$ to λ. There results the estimate

$$\begin{aligned}|t^2u_3(t,\lambda)| &\leqq \frac{M}{\lambda}+M\varepsilon+M\int_\rho^{\omega(\lambda)-\varepsilon}\frac{d\tau}{[\omega(\lambda)-\tau]^2}\\ &\quad+M\lambda\int_{\rho/2}^{\omega(\lambda)-\varepsilon}\frac{d\tau}{\tau[\omega(\lambda)-\tau]^3[\omega(\lambda)+\tau]}\\ &\quad+M\int_{\omega(\lambda)+\varepsilon}^{\infty}\left[\frac{\lambda}{[\omega(\lambda)-\tau]^2[\omega(\lambda)+\tau]^2}+\frac{\tau^3}{[\tau-\omega(\lambda)]^3[\omega(\lambda)+\tau]^3}\right]d\tau\\ &\leqq M+M\sqrt{\lambda}\int_{\rho/2}^{\omega(\lambda)/2}\frac{d\tau}{\tau\omega^3(\lambda)}+M\int_{\omega(\lambda)/2}^{\omega(\lambda)-\varepsilon}\frac{d\tau}{[\omega(\lambda)-\tau]^3}\\ &\quad+M\int_{\omega(\lambda)+\varepsilon}^{\infty}\left[\frac{1}{[\tau-\omega(\lambda)]^2}+\frac{1}{[\tau-\omega(\lambda)]^3}\right]d\tau\leqq M.\end{aligned}$$

Thus (5.2) holds.

Integration by parts in (2.16) yields

$$\begin{aligned}t^2v_2(t,\lambda)=\operatorname{Re}\Bigg\{&e^{i\rho t}\left[\frac{D'(\rho)}{D(\rho,\lambda)}-\frac{D(\rho)D'(\rho,\lambda)}{D^2(\rho,\lambda)}\right]\\ &+\frac{i}{\lambda}\int_\rho^{2\omega(\lambda)}e^{i\tau t}\left[\frac{2D'(\tau,\lambda)}{D^2(\tau,\lambda)}+\frac{\tau D''(\tau)}{D^2(\tau,\lambda)}-\frac{2\tau[D'(\tau,\lambda)]^2}{D^3(\tau,\lambda)}\right]d\tau\\ &+\int_{2\omega(\lambda)}^{\infty}\frac{e^{i\tau t}}{D(\tau,\lambda)}\Bigg[D''(\tau)-\frac{2D'(\tau)D'(\tau,\lambda)-D(\tau)D''(\tau)}{D(\tau,\lambda)}\\ &\qquad+\frac{2D(\tau)[D'(\tau,\lambda)]^2}{D^2(\tau,\lambda)}\Bigg]d\tau\Bigg\}.\end{aligned}$$

Then similar estimates to those above show that

$$|t^2v_2(t,\lambda)|\leqq M\sqrt{\lambda}+\int_{2\omega(\lambda)}^{\infty}M\left[\frac{\lambda}{\tau^3}+\frac{1}{\tau^2}\right]d\tau\leqq M\sqrt{\lambda}.$$

For the other estimates, first let

$$f(\tau)=\frac{\int_0^{1/\tau}ta(t)\,dt}{\left[\int_0^{1/\tau}a(t)\,dt\right]^2}.$$

Note that

$$\tau f(\tau)\leqq\left[\int_0^{1/\tau}a(t)\,dt\right]^{-1}. \tag{6.10}$$

An argument of Shea and Wainger [10] (see [7, p. 696] for the version needed here) shows that for $0<\delta\leqq\rho$,

$$\int_\delta^\rho f(\tau)\,d\tau\leqq 2\left[\int_0^{1/\rho}a(t)\,dt\right]^{-1}\leqq\frac{12}{t_1a(t_1)}.$$

Thus

$$\int_0^\rho f(\tau)\,d\tau \leqq M. \tag{6.11}$$

Now if $c>0$, (2.4), (2.5) and $a(\infty)=0$ imply that

$$\frac{\tau D'(\tau,\lambda)}{D^2(\tau,\lambda)}\to 0 \qquad (\tau\to 0+, \quad 1\leq\lambda\leqq\infty). \tag{6.12}$$

If $c=0$ and $A(0)=\infty$, (6.12) is still true, because of (2.4), (2.5) and (6.10). If $c=0$ and $A(0)<\infty$,

$$\tau A'(\tau)=-\int_0^\infty\left[\frac{e^{-i\tau t}-1+i\tau t}{i\tau t}\right]ta'(t)\,dt\to 0$$

as $\tau\to 0+$ by Lebesgue's dominated convergence theorem, so (6.12) holds in all cases. Moreover, $D(0+,\lambda)$ is real or infinite ($1\leq\lambda\leqq\infty$). We may then integrate by parts in (2.13) to obtain

$$\begin{aligned} it^2\lambda^2 u_2(t,\lambda)=\operatorname{Im}\Bigg\{&\frac{e^{i\rho t}}{D(\rho)D(\rho,\lambda)}\left[1-\frac{\rho D'(\rho)}{D(\rho)}-\frac{\rho D'(\rho,\lambda)}{D(\rho,\lambda)}\right]\\ &-\int_0^\rho\frac{e^{i\tau t}}{D(\tau)D(\tau,\lambda)}\left[\frac{-2D'(\tau)}{D(\tau)}-\frac{2D'(\tau,\lambda)}{D(\tau,\lambda)}-\frac{\tau D''(\tau)}{D(\tau)}\right.\\ &\qquad-\frac{\tau D''(\tau)}{D(\tau,\lambda)}+\frac{2\tau[D'(\tau)]^2}{D^2(\tau)}\\ &\qquad\left.+\frac{2\tau[D'(\tau,\lambda)]^2}{D^2(\tau)}+\frac{2\tau D'(\tau)D'(\tau,\lambda)}{D(\tau)D(\tau,\lambda)}\right]d\tau\Bigg\}. \end{aligned} \tag{6.13}$$

If $c>0$, then $|\operatorname{Im} D(\tau,\lambda)|\geqq c/2\tau$ if $\tau^2<c/2$, while $\operatorname{Re} D(\tau,\lambda)\geqq\eta/(1+\rho^2)$ ($\sqrt{c/2}\leqq\tau\leqq\rho$). Thus by (2.5) and Lemma 6.1,

$$t^2\lambda^2|u_2(t,\lambda)|\leqq M+M\int_0^{\min\{\rho,\sqrt{c/2}\}}[\tau+\tau^5]\,d\tau,$$

so

$$t^2\lambda^2|u_2(t,\lambda)|\leqq M. \tag{6.14}$$

If $c=0$ let $\sigma=\min\{\rho,\int_0^{1/\rho}a(t)\,dt/8\}$; then $\sigma^{-1}\leqq M$, $|D(\tau,\lambda)|\geqq\eta/(1+\rho^2)$ ($\sigma<\tau\leqq\rho$). In Theorem 2 we use (2.5) and (6.2) to get

$$|\lambda^2t^2u_2(t,\lambda)|\leqq M+M\int_0^\sigma[1+f(\tau)]\,d\tau.$$

Using (6.11), we again get (6.14). Thus (5.1) holds (without the log term) under the hypotheses of Theorem 2 and in Theorem 1 when $c>0$.

For Theorem 1 with $c = 0$, we return to (2.13) for the estimate

$$|t\lambda^2 u_2(t,\lambda)| \leqq M\int_0^{\min\{\sigma,1/t\}}\left[1+\tau\int_0^{1/\tau} sa(s)\,ds\right]d\tau$$
$$+\left|\int_{\min\{\sigma,1/t\}}^{\rho}\frac{e^{i\tau t}}{D(\tau)D(\tau,\lambda)}\left[1-\frac{\tau D'(\tau)}{D(\tau)}-\frac{\tau D'(\tau,\lambda)}{D(\tau,\lambda)}\right]d\tau\right|$$
$$= I_1 + |I_2|.$$

Here I_1 is bounded by M/t, since $A(0) \leqq M$, while integration by parts shows that

$$I_2(t,\lambda) = \left[\frac{e^{i\tau t}}{itD(\tau)D(\tau,\lambda)}\left(1-\frac{\tau D'(\tau)}{D(\tau)}-\frac{\tau D'(\tau,\lambda)}{D(\tau,\lambda)}\right)\right]_{\min\{\sigma,1/t\}}^{\rho}$$
$$-\frac{1}{it}\left[\int_{\min\{\sigma,1/t\}}^{\sigma}+\int_{\sigma}^{\rho}\right]Q\,d\tau,$$

where $Q = Q(\tau,\lambda)$ is the same as the integrand in (6.13). Estimating as before, but with (6.1) instead of (6.2), we find that

$$|tI_2(t,\lambda)| \leqq M + M\int_{\min\{\sigma,1/t\}}^{\sigma}\frac{d\tau}{\tau} \leqq M(1+\log t).$$

Then (5.1) holds in all cases.

The proof of (5.3) is similar to that for (5.1), and we omit it.

Proof Lemma 5.2. Start with (2.2). Obviously $|D(\tau)/D(\tau,\lambda)| = |1 - i\tau/\lambda D(\tau,\lambda)| \leqq 1+\rho(1+\rho^2)/\eta\lambda \leqq M$ $(0 \leqq \tau \leqq \rho)$, so

$$\int_0^{\rho}\left|1-\frac{i\tau}{\lambda D(\tau,\lambda)}\right|d\tau \leqq M. \tag{6.15}$$

Next, using (6.8) (a consequence of (2.7) and (2.9), which hold here, by Lemma 2.2) we see that

$$\int_{\rho/2}^{\omega(\lambda)/2}\left|1-\frac{i\tau}{\lambda D(\tau,\lambda)}\right| \leqq \frac{1}{2}\omega(\lambda)+\frac{M}{\lambda}\int_{\rho/2}^{\omega(\lambda)/2}\tau^2\,d\tau \leqq M\sqrt{\lambda}, \tag{6.16}$$

$$\left[\int_{\omega(\lambda)/2}^{\omega(\lambda)-\varepsilon}+\int_{\omega(\lambda)+\varepsilon}^{2\omega(\lambda)}\right]\left|1-\frac{i\tau}{\lambda D(\tau,\lambda)}\right|d\tau \leqq M\sqrt{\lambda}\left(1+\int_{\varepsilon}^{\omega(\lambda)}\sigma^{-1}\,d\sigma\right)$$
$$\leqq M\sqrt{\lambda}\,(1+\log\lambda), \tag{6.17}$$

$$\int_{\omega(\lambda)-\varepsilon}^{\omega(\lambda)+\varepsilon}\left|1-\frac{i\tau}{\lambda D(\tau,\lambda)}\right|d\tau \leqq 2\varepsilon\left[1+\frac{M(1+\lambda)}{\eta\sqrt{\lambda}}\right] \leqq M\sqrt{\lambda}, \tag{6.18}$$

$$\int_{2\omega(\lambda)}^{\infty}|D(\tau)/D(\tau,\lambda)|\,d\tau \leqq M\lambda\int_{2\omega(\lambda)}^{\infty}\tau^{-2}\,d\tau \leqq M\sqrt{\lambda}. \tag{6.19}$$

Comparing (6.15) through (6.19) with (2.2), we see that (5.5) holds.

Proof of Lemma 5.3. Integration by parts shows that

$$|A_n(\tau)-A(\tau)| \leqq -\tau^{-1}\int_0^{1/n}[a_n'(t)+a'(t)]\,dt \tag{6.20}$$

$$= 2[a(0)-a(1/n)]/\tau.$$

With $D_n(\tau,\lambda)=i\tau\lambda^{-1}+A_n(\tau)+c/(i\tau)$, we have

$$\left|\frac{1}{D_n(\tau,\lambda)}-\frac{1}{D(\tau,\lambda)}\right| \leq \frac{2[a(0)-a(1/n)]}{\tau|D(\tau,\lambda)D_n(\tau,\lambda)|}. \tag{6.21}$$

To each n there corresponds a number $\omega_n=\omega_n(\lambda)$. By (2.9) and (6.8) there are positive numbers δ, $\Delta<\infty$., independent of n, such that $\omega_n<\Delta$ $(n>N)$ and $|D_n(\tau,\lambda)|\geqq\delta\tau$ $(n\geqq N, \tau\geqq 2\Delta)$. Since $A_n\to A$, also $|D(\tau,\lambda)|\geqq\delta\tau$ $(\tau\geqq 2\Delta)$. Thus by (6.21), $\int_{2\Delta}^{\infty}|D_n^{-1}(\tau,\lambda)-D^{-1}(\tau,\lambda)|\,d\tau\to 0$ $(n\to\infty)$. But $D_n\to D$ uniformly and $D^{-1}(\tau,\lambda)$ is continuous on $\{0\leqq\tau\leqq 2\Delta\}$, so

$$\int_0^{\infty}\left|\frac{1}{D_n(\tau,\lambda)}-\frac{1}{D(\tau,\lambda)}\right|d\tau\to 0 \qquad (n\to\infty). \tag{6.22}$$

Clearly (6.22) also holds with λ replaced by ∞. The convergence of u_2^n, u_3^n, v_1^n is now obvious. Moreover,

$$\begin{aligned}|v_2^n(t,\lambda)-v_2(\tau,\lambda)| \leqq & \left|\frac{D(\rho)}{D(\rho,\lambda)}-\frac{D_n(\rho)}{D_n(\rho,\lambda)}\right| \\ & +\int_\rho^{\infty}\left|D(\tau)\left[\frac{1}{D(\tau,\lambda)}-\frac{1}{D_n(\tau,\lambda)}\right]\right|d\tau \\ & +\int_\rho^{\infty}\left|\frac{A(\tau)-A_n(\tau)}{D(\tau,\lambda)}\right|d\tau \\ & +\int_\rho^{\infty}|A(\tau)-A_n(\tau)|\left|\frac{1}{D_n(\tau,\lambda)}-\frac{1}{D(\tau,\lambda)}\right|d\tau.\end{aligned}$$

Using Lemma 2.1 (ii), (6.20) and (6.22), we see that $v_2^n(t,\lambda)\to v_2(t,\lambda)$ as $n\to\infty$. Finally, the convergence of $u_n'(t,\lambda)$ to $u'(t,\lambda)$ is a simple consequence of (1.1) and standard continuous dependence theorems for Volterra integral equations. This completes our proof.

7. Proof of Corollary 1.1. Taking Laplace transforms in (1.8) and using (1.1) with $\lambda=\mu$, one sees that

$$U^*(z,\lambda)=\int_0^{\infty}e^{-zt}U(t,\lambda)\,dt=\frac{1}{z+\mu a^*(z)}\sum_{n=0}^{\infty}\left[\frac{a^*(z)(\mu-\lambda)}{z+\mu a^*(z)}\right]^n$$

$$=\frac{1}{z+\lambda a^*(z)} \qquad (\mathrm{Re}\, z>0).$$

Using the standard existence-uniqueness proof [1, Chap. 7] for the integrated form $u(t,\lambda)+\lambda\int_0^t[(t-s)c+\int_0^{t-s}a(x)\,dx]u(s,\lambda)\,ds=t$ of (1.1), one sees that a

unique solution $u(t, \lambda)$ exists on $\{0 \leq t < \infty\}$ and $|u(t, \lambda)| \leqq \alpha e^{\beta t}$ $(t \geqq 0)$ for suitable finite α, β. Then by (1.1), $|u'(t, \lambda)| \leqq \tilde{\alpha} e^{\beta t}$ $(t \geqq 0)$. Taking Laplace transforms we get $u^*(z, \lambda) = U^*(z, \lambda)$ (Re $z > \beta$), so $u = U$.

REFERENCES

[1] R. BELLMAN AND K. COOKE, *Differential-Difference Eauations*, Academic Press, New York, 1963.

[2] S. I. GROSSMAN AND R. K. MILLER, *Nonlinear Volterra integrodifferential systems with L^1-kernels*, J. Differential Equations, 13 (1973), pp. 551–556.

[3] K. B. HANNSGEN, *Indirect Abelian theorems and a linear Volterra equation*, Trans. Amer. Math. Soc., 142 (1969), pp. 539–555.

[4] ———, *A Volterra equation with parameter*, this Journal, 4 (1973), pp. 22–30.

[5] ———, *A Volterra equation in Hilbert space*, this Journal, 5 (1974), pp. 412–416.

[6] ———, *Note on the family of Volterra equations*, Proc. Amer. Math. Soc., 46 (1974), pp. 239–243.

[7] ———, *Uniform boundedness in a class of Volterra equations*, this Journal, 6 (1975), pp. 689–697.

[8] ———, *The resolvent kernel of an integrodifferential equation in Hilbert space*, this Journal, 6 (1975), pp. 689–697.

[9] J. A. NOHEL AND D. F. SHEA, *Frequency domain methods for Volterra equations*, Advances in Math., to appear.

[10] D. F. SHEA AND S. WAINGER, *Variants of the Wiener–Lévy theory, with applications to stability problems for some Volterra integral equations*, Amer. J. Math., 97 (1975), pp. 312–343.

SIAM J. MATH. ANAL.
Vol. 8, No. 4, August 1977

ASYMPTOTIC ESTIMATES FOR THE ADIABATIC INVARIANCE OF A SIMPLE OSCILLATOR*

GILBERT STENGLE†

Abstract. Let u be a solution of the differential equation $\ddot{u}+\varphi^2 u=0$ with slowly varying coefficient $\varphi^2(\varepsilon t)$. Let $r^2=\varphi(\varepsilon t)u^2+\varphi^{-1}(\varepsilon t)\dot{u}^2$. Then r is an approximate or "adiabatic" invariant of u in the sense that $\dot{r}=O(\varepsilon)$. J. E. Littlewood [1] has shown that $r^2(+\infty)-r^2(-\infty)=O(\varepsilon^n)$ for all $n>0$ under the hypotheses that $\varphi>0$, φ has positive limits as $t\to\pm\infty$, and $\varphi^{(n)}\in L_1(-\infty,\infty)$ for all $n>0$. The purpose of this paper is to obtain an upper estimate for $r^2(+\infty)-r^2(-\infty)$ under Littlewood's hypotheses. The main result is that the rate of decrease of $r^2(\infty)-r^2(-\infty)$ as $\varepsilon\to 0^+$ is determined by the rate of growth of $\|\varphi^{(m)}\|_1$ as $m\to\infty$. It is shown that if $\|\varphi^{(m)}\|_1=O\{\exp h(m)\}$, where $m\log m=o(h(m))$, then for any $\delta>0$,

$$\frac{r^2(\infty)}{r^2(-\infty)}-1=O\{\exp-h^*(\log\varepsilon^{-1+\delta})\},$$

where h^* is the convex conjugate function of h

$$h^*(x)=\max_y\{xy-h(y)\}.$$

1. Introduction. We consider the differential equation $u''+\varphi^2(\varepsilon\tau)u=0$ with slowly varying coefficient $\varphi^2(\varepsilon\tau)$. Here ε is a small positive parameter and τ ranges over the entire real line. On bounded time intervals or even intervals of length $o(1/\varepsilon)$, this problem can be regarded as a small perturbation of a simple harmonic oscillator $u''+k^2u=0$; for longer durations this resemblance fails. Nevertheless the theory of the limiting behavior of this problem as ε tends to zero is essentially simpler than the theory of the full problem $u''+\varphi^2(\tau)u=0$ and various techniques of asymptotic analysis yield abundant information about equations with slowly varying coefficients. Divergent asymptotic series in powers of ε are usually an important tool here. In this paper we study a subtle facet of this problem which seems to lie beyond these methods and which requires tools not commonly used in asymptotics.

If we associate with a solution u of the differential equation the function $r^2=\varphi(\varepsilon\tau)u^2+\varphi^{-1}(\varepsilon\tau)(u')^2$, then a simple computation shows that $r'=O(\varepsilon)$. Thus the change in r^2 is small on bounded time intervals. This property of r is often indicated by calling r^2 an approximate or "adiabatic" invariant of u. However r^2 has even more remarkable properties. In 1963, J. E. Littlewood [1] showed that $r^2(\infty)-r^2(-\infty)$ is asymptotic to 0 as a function of ε, that is, $r^2(\infty)-r^2(-\infty)=O(\varepsilon^n)$ for each $n\geqq 0$, under the following hypotheses. He assumed that φ is positive valued and has positive limiting values as τ approaches $\pm\infty$ and that φ' is *gentle* in the sense that $\varphi^{(n)}\in L_1(-\infty,\infty)$ for all $n>0$. The purpose of this paper is to give an estimate for $r^2(\infty)-r^2(-\infty)$, the limiting value of the adiabatic invariant $r^2(\tau)-r^2(-\infty)$, precisely under Littlewood's hypotheses. For brevity we shall refer to $r^2(+\infty)-r^2(-\infty)$ as the "adiabatic invariant". The following theorem (proved in § 7) shows that the nature of the zero at $\varepsilon=0^+$ of the adiabatic invariant is

* Received by the editors January 20, 1973, and in revised form March 2, 1976.
† Department of Mathematics, Lehigh University, Bethlehem, Pennsylvania 18015.

determined by the rate of growth of the sequence $\|\varphi^{(m)}\|_1 = \int_{-\infty}^{\infty} \|\varphi^{(m)}(s)\|\, ds$ as $m \to \infty$. In the statement of the theorem and for the balance of the paper we change the time scale by the transformation $t = \varepsilon\tau$ obtaining

$$\varepsilon^2 \ddot{u} + \varphi^2(t)u = 0 \tag{1.1}$$

as the basic differential equation and $r^2 = \varphi u^2 + \varepsilon^2 \varphi^{-1} \dot{u}^2$ as the associated approximate invariant.

THEOREM 1.1. *Let u be a solution of $\varepsilon^2 \ddot{u} + \varphi^2 u = 0$, where φ is a positive function, φ tends to positive limits as t approaches $\pm\infty$ and $\dot{\varphi}$ is gentle. Let $r^2 = \varphi u^2 + \varepsilon^2 \varphi^{-1} \dot{u}^2$. Suppose*

$$\|\varphi^{(m)}\|_1 = O\{\exp h(m)\},$$

where $h(m)/m \log m \to \infty$ as $m \to \infty$. Then for any $\delta > 0$,

$$\frac{r^2(\infty)}{r^2(-\infty)} - 1 = O\{\exp - h^*(\log \varepsilon^{-1+\delta})\},$$

where h^* is the convex conjugate function of h,

$$h^*(x) = \max_y \{xy - h(y)\}.$$

We make the following remarks about the hypotheses of this theorem. We have assumed much more than is required to establish the oscillatory nature of all solutions and the existence of $r(\pm\infty)$. The function r is just the radius in the phase plane associated with the Prüfer transformation $\varphi^{1/2} u = r \sin\theta$, $\varepsilon\varphi^{1/2}\dot{u} = r\cos\theta$. The functions r and θ satisfy $\dot{r} = -\frac{1}{2}\varphi^{-1}\dot{\varphi}\cos 2\theta$ and $\varepsilon\dot{\theta} = \varphi + (\varepsilon/2)\varphi^{-1}\dot{\varphi}\sin 2\theta$. These equations can easily be used to show that $r(\pm\infty)$ exists if $\varphi^{-2}\dot{\varphi}$ has bounded variation (Hille [2]). We also note that the function $h(m)$ describing the growth of the sequence of derivatives of φ is far from arbitrary. This is so because of interpolation properties possessed by derivative norms. However since we employ the operation of convex conjugation which obliterates all distinctions between h and its largest convex minorant, we shall assume without further mention that h is convex and monotonic increasing. We also assume as an inessential convenience that h is defined for real $x \geqq 0$ and that $h(x) = \infty$ for $x < 0$.

Our hypotheses ensure that φ admits the action of the Riemann–Liouville fractional derivative D^x for $x \geqq 0$. This can be defined for $n < x < n+1$ by

$$D^x \varphi(t) = \frac{1}{\Gamma(n+1-x)} \int_0^{\infty} s^{n-x} \varphi^{(n+1)}(t-s)\, ds.$$

This is easily seen to be an element of L_1 for $x \geqq 0$, or we can choose $h_0(x) = \log \int_{-\infty}^{\infty} |s|^x |\varphi(s)|\, ds$, where φ is the Fourier transform of y. The convexity of this function is a simple consequence of the Holder inequality. The function $h(x) = \sup_{0 \leqq s \leqq x} h_0(s)$ is then a monotonic convex measure of derivative growth of the kind we desire.

Finally we assume that $h(m)$ grows more rapidly than $m \log m$, the distinguishing growth rate for holomorphic data.

Recently Wasow [3], [4] and Meyer [5] have investigated the adiabatic invariant obtaining very detailed results in the case that φ is holomorphic in a strip containing the real axis and satisfies various other hypotheses. They find in such cases that the adiabatic invariant is $O(\exp -(C/\varepsilon))$. We illustrate the sense of our result by showing it leads to estimates other than exponential. Suppose $\|\varphi^{(m)}\|_1 = O(\exp \tau m^\rho)$, $\tau > 0$, $\rho > 1$. A simple calculation shows that for $h(x) = \tau x^\rho$, we have $h^*(y) = \tau' y^{\rho'}$, where $\rho' > 1$, $(1/\rho)+(1/\rho') = 1$ and $\tau' = (\rho-1)\tau(\rho\tau)^{-\rho/(\rho-1)}$. In this case the theorem implies that

$$a(\varepsilon) = O\left(\exp -c\left[\log \frac{1}{\varepsilon}\right]^{\rho'}\right)$$

for any $c < \tau'$.

In § 2 we continue our introduction. Section 3 catalogues some changes of variable. In § 4 we obtain a formula for the adiabatic invariant in terms of distinguished solutions of a related nonlinear equation. Section 5 contains analytic results in the special case that φ is holomorphic. Section 6 contains results on the approximation of gentle data by holomorphic data. In § 7 we prove the main theorem.

2. A method of estimation. We estimate the adiabatic invariant by combining some tools of general significance with several very special devices useful only for the second order linear equation. Therefore it seems valuable to us to distinguish the authentic methods from the special tricks with the following remarks.

It frequently happens in analyzing differential equations which depend singularly on a parameter that mere knowledge of the existence of a special solution with derivatives bounded independently of the parameter is a powerful asymptotic tool, more powerful, in fact, than conventional asymptotic results which assert existence of a solution with a given asymptotic expansion. We illustrate this point with a very simple example. Suppose we know that the problem $-\varepsilon y' + y = f$ has a solution y_0 which has derivatives of all orders bounded independently of ε. Then (supposing that f is smooth) it is a rigorous consequence of the differential equation itself that

$$y_0 = f + \varepsilon f' + \varepsilon^2 f'' + \cdots + \varepsilon^n f^{(n)} + \varepsilon^{n+1} y_0^{(n+1)}.$$

But since $y_0^{(n+1)}$ is bounded independently of ε, this formula asserts that the obvious formal series $\sum_{k=0}^{\infty} \varepsilon^k f$ is an asymptotic expansion of the solution y_0. To put it strongly, here the obvious formal calculation, which usually has only heuristic status, becomes a strict deduction. Moreover if actual numerical or order estimates for $y_0^{(m)}$ are known, e.g., $|y_0^{(m)}| \leqq 1000\, e^{m^2}$ or $y_0^{(m)} = O(e^{m^2})$, then we have similar bounds for the error, e.g., $|y_0 - \sum_{k=0}^{m} \varepsilon^k f^{(k)}| \leqq 1000\, \varepsilon^{m+1} e^{(m+1)^2}$. Finally if, as is usual, the error does not tend to zero as we include more terms in the series, then we can assess the theoretical limits of accuracy attainable with the asymptotic formula. Under our concrete illustrative hypotheses we find this to be

$$\min_{m \geqq 0} 1000\, \varepsilon^{m+1} e^{-(m+1)^2} = O\left\{\exp -\frac{1}{4}\left(\log \frac{1}{\varepsilon}\right)^2\right\}.$$

We will apply a subtle variant of the above argument (roughly a case where each term of the expansion is 0) to

$$\dot{Q}^2+\frac{\varepsilon^2}{2}\{Q, t\}=\varphi^2, \tag{2.1}$$

where $\{Q, t\}$ is the Schwarzian derivative

$$\{Q, t\}=\frac{\dddot{Q}}{\dot{Q}}-\frac{3}{2}\left(\frac{\ddot{Q}}{\dot{Q}}\right)^2.$$

Thus our program requires obtaining solutions of this equation together with ε-independent bounds for all derivatives. Plainly such bounds cannot be simply deduced recursively from the differential equation since the highest derivative is multiplied by the small parameter ε. To obtain derivative estimates in a systematic way we exploit a powerful idea that Jacobowitz [6] has used to give remarkably simple proofs for implicit function theorems of the Nash–Kolmogorov–Arnold–Moser type. In problem (2.1) we replace φ by a sequence of function φ_n holomorphic on a sequence of strips converging to the real axis. We do this in such a way that the convergence of φ_n to φ accurately reflects the differentiability properties of φ. We then solve $\dot{Q}_n^2+(\varepsilon^2/2)\{Q_n, t\}=\varphi_n^2$ for holomorphic approximate solutions Q_n. However for the Q_n we have an essential simplification in the problem of obtaining derivative estimates. All derivatives of Q_n can be estimated on a slightly smaller strip by the Cauchy integral formula in terms of an upper bound for the function Q_n itself. We then carry these derivative estimates over to an exact solution Q by a limiting process. As we will see, this problem of obtaining derivative estimates for solutions of the Schwarzian equation (2.1) is the central technical difficulty in this paper.

3. Some special transformations. We introduce the following variant of the Prüfer transformation in which the slowly varying function r^2 appears as a modulus:

$$\begin{gathered} z=\varphi^{1/2}u+i\varepsilon\varphi^{-1/2}\dot{u}, \qquad \hat{z}=\varphi^{1/2}u-i\varepsilon\varphi^{-1/2}\dot{u}, \\ r^2=z\hat{z}, \qquad p=z/\hat{z}. \end{gathered} \tag{3.1}$$

We note that if u is a real solution, then $r=|z|$ and $p=\exp 2i\theta$, where $\theta=\arg z$. However since we are interested in complex solutions we forgo introducing θ. Then a simple computation shows that

$$\frac{\dot{r}}{r}=\frac{\psi}{2}\left(p+\frac{1}{p}\right), \tag{3.2}$$

$$\varepsilon\dot{p}=-2i\varphi p+(\varepsilon\psi/2)(1-p^2), \tag{3.3}$$

where $\psi=\varphi^{-1}\dot{\varphi}$.

We next catalog relations among the original linear equation, the Schwarzian equation (2.1) and the Riccati equation (3.3). It is well known that given a single solution of any one of those equations the complete solution of each can be determined by integrations, differentiations, algebraic operations and exponentiations. We use the Riccati equation, following Wasow [3], [4], because it is easiest

to solve. However for our purposes the Schwarzian equation is simplest because any two of its solutions are related in a direct way. In contrast, simple relations persist only among larger sets of solutions of the other two equations.

In the following we presuppose without further mention the hypotheses of the main theorem. We require the following elementary attributes of the Schwarzian derivative (Hille [2]).

LEMMA 3.1. *The Schwarzian derivative,* $\{f, t\} = (\ddot{f}/\dot{f}) - \frac{3}{2}(\ddot{f}/\dot{f})^2$, *has the following properties*:

(i) *If* y_1, y_2, *satisfy* $\ddot{y} + a(t)y = 0$, *then* $\{y_1/y_2, t\} = 2a(t)$.

(ii) $\{f, t\} = \{g, t\}$ *if and only if* $f = (\alpha g + \beta)/(\gamma g + \delta)$ *where* $\alpha\delta - \beta\gamma = 1$.

(iii) $\{u, w\} = \{u, v\}(dv/dw)^2 + \{v, w\}$.

(iv) $\{u, v\} = -\{v, u\}(du/dv)^2$.

LEMMA 3.2. *Let* p *be a solution of* $\varepsilon\dot{p} + 2i\varphi p = (\varepsilon/2)\psi(1-p^2)$ *satisfying* $|p| < \frac{1}{3}$. *Then* $Q = \int_0^t \varphi(1 - 2\,\mathrm{Re}\,(p/(1+p)))\,ds$ *is a solution of* $(\varepsilon^2/2)\{Q, t\} + \dot{Q}^2 = \varphi^2$ *and the functions* $\dot{Q}^{-1/2} \exp \pm(i/\varepsilon)Q$ *are independent solutions of* $\varepsilon^2\ddot{u} + \varphi^2 u = 0$.

Proof. Suppose $|p| < \frac{1}{3}$. Then our hypotheses on φ imply $\dot{Q} = \varphi(1 - 2\,\mathrm{Re}\,(p/(1+p))) > 0$. A tedious but straightforward computation then shows without further hypotheses that Q satisfies the Schwarzian equation. Instead we give the following argument which uses the preceding properties of the Schwarzian to reveal quickly a direct relationship between p and Q, but which requires that $p \neq 0$. If r and p satisfy (3.2) and (3.3), then $y = \varphi^{-1/2} r^{1/2}(p^{1/2} + p^{-1/2})$ and $\bar{y} = \varphi^{-1/2}\bar{r}^{1/2}(\bar{p}^{1/2} + \bar{p}^{-1/2})$ are solutions of the linear second order equation. Hence by Lemma 3.1(i), $y/\bar{y}$ is a solution of $\{y/\bar{y}, t\} = 2\varphi^2/\varepsilon^2$. Applying Lemma 3.1(iii), the chain rule for the Schwarzian to the (real-valued) function $Q = (\varepsilon/(2i))\log y/\bar{y}$ we find $(2/\varepsilon^2)\dot{Q}^2 + \{Q, t\} = 2\varphi^2$. Since

$$\dot{Q} = \frac{\varepsilon}{2i}\left\{\frac{\dot{r}}{r} - \frac{\dot{\bar{r}}}{\bar{r}} + \frac{d}{dt}\log\frac{p^{1/2} + p^{-1/2}}{\bar{p}^{1/2} + \bar{p}^{-1/2}}\right\},$$

equations (3.2) and (3.3) immediately show that $\dot{Q}$ is given directly in terms of p. Carrying out the differentiation and eliminating derivatives with (3.2) and (3.3) yields the desired relation. Finally the change of variable $\tau = Q(t)$, $v(\tau) = \dot{Q}^{1/2}(t)u(t)$ transforms $\varepsilon^2(d^2v)/(d\tau^2) + v = 0$ into $\varepsilon^2\ddot{u} + (\dot{Q}^2 + (\varepsilon^2/2)\{Q, t\})u = 0$ which is $\varepsilon^2\ddot{u} + \varphi^2 u = 0$. Choosing $v(\tau) = \exp \pm(i\tau)/\varepsilon$ we obtain the indicated solutions of $\varepsilon^2\ddot{u} + \varphi^2 u = 0$.

LEMMA 3.3. *Let* Q_1 *and* Q_2 *be monotonic increasing solutions of* $\dot{Q}^2 + (\varepsilon^2/2)\{Q, t\} = \varphi^2$. *Then*

$$\frac{d}{dt} Q_1 \circ Q_2^{\text{inverse}} = \left\{A(\varepsilon) + B(\varepsilon)\cos\frac{2t}{\varepsilon} + C(\varepsilon)\sin\frac{2t}{\varepsilon}\right\}^{-1},$$

where $A^2 = B^2 + C^2 + 1$.

Proof. Since $\{\tan u/\varepsilon, u\} = 2/\varepsilon^2$ the chain rule Lemma 3.1(iii) shows $\{\tan Qi/\varepsilon, t\} = (2/\varepsilon^2)\varphi^2$. Hence by Lemma 3.1(ii),

$$Q_1 = \varepsilon \tan^{-1}\frac{\alpha \tan (Q_2/\varepsilon) + \beta}{\gamma \tan (Q_2/\varepsilon) + \delta},$$

where $\alpha\delta - \beta\gamma = 1$. Since the indicated linear fractional transformation maps a real segment onto a real segment we can suppose its coefficients are real. However since the functions $\tan Q_k$ have disconnected domains of definition there remains the possibility that different linear fractional transformations occur on each component. Replacing t by $Q_2^{\text{inverse}}(t)$ and differentiating we find after a slight computation that on each component,

$$\frac{d}{dt} Q_1 \circ Q_2^{\text{inverse}} = \left\{ A + B\cos\frac{2t}{\varepsilon} + C\sin\frac{2t}{\varepsilon} \right\}^{-1},$$

where

$$A = \frac{\alpha^2+\beta^2+\gamma^2+\delta^2}{2}, \quad B = \frac{\beta^2+\delta^2-\alpha^2-\gamma^2}{2}, \quad c = \alpha\beta + \gamma\delta.$$

Since the function on the left is smooth, A, B, C are globally constant. Finally a simple computation gives $A^2 - B^2 - C^2 = (\alpha\delta - \beta\gamma)^2 = 1$.

LEMMA 3.4. *Let $\tilde{Q}_1$, $\tilde{Q}_2$ be monotonic increasing solutions of $\dot{Q}^2 + (\varepsilon^2/2)\{Q, t\} = \varphi^2$. Then there are constants c_1, c_2 such that $Q_i = \tilde{Q}_i + c_i$ satisfy*

$$\tan\frac{Q_1}{\varepsilon} = (1+a)\tan\frac{Q_2}{\varepsilon},$$

where $a = A - \sqrt{B^2+C^2} - 1$ and A, B, C are the parameters of Lemma 3.3.

Proof. Lemma 3.3 implies $\dot{\tilde{Q}}_1 = \{A + B\sin(2\tilde{Q}_2/\varepsilon) + C\cos(2\tilde{Q}_2/\varepsilon)\}^{-1}\dot{\tilde{Q}}_2$. This can be written $\dot{\tilde{Q}}_1 = \{A + \sqrt{B^2+C^2}\cos(2/\varepsilon)(\tilde{Q}_2 + C_2)\}^{-1}\dot{\tilde{Q}}_2$ for appropriate C_2. Integrating and using $A^2 - B^2 - C^2 = 1$ we find $\tilde{Q}_1(t) = \varepsilon(\tan^{-1}\{(A - \sqrt{B^2+C^2})\tan((\tilde{Q}_2 + C_2)/\varepsilon)\} + n(t)\pi - C_1)$, where $n(t)$ is an integer and C_1 is a constant. This implies

$$\tan\frac{\tilde{Q}_1 + C_1}{\varepsilon} = (A - \sqrt{B^2+C^2})\tan\frac{\tilde{Q}_2 + C_2}{\varepsilon}.$$

4. A formula for the adiabatic invariant. We now show that for a natural choice of solutions Q_1 and Q_2, the constant a of Lemma 3.4 is precisely the adiabatic invariant. It is not hard to guess that the natural solutions of the Schwarzian equation are solutions possessing simple behavior at $-\infty$ and $+\infty$ respectively.

THEOREM 4.1. *Let $Q_\pm$ be monotonic increasing solutions of $\dot{Q}^2 + (\varepsilon^2/2)\cdot\{Q, t\} = \varphi^2$ related by $\tan Q_-/\varepsilon = (1+a)\tan Q_+/\varepsilon$. Suppose $\dot{Q}_\pm(\pm\infty) = \varphi(\pm\infty)$ and $\ddot{Q}_\pm(\pm\infty) = 0$. Then the adiabatic invariant of the solution $\dot{Q}_-^{-1/2}\sin Q_-/\varepsilon$ of $\varepsilon^2\ddot{u} + \varphi u^2 = 0$ is precisely a.*

Proof. In the following proof (but nowhere else) we use the order symbol $o(1)$ to indicate functions which tend to zero as $t \to +\infty$. The slowly varying function r^2 associated with $u = \dot{Q}_-^{-1/2}\sin Q_-/\varepsilon$ has the form

$$r^2 = \varphi\dot{Q}_-^{-1}\sin^2\frac{Q_-}{\varepsilon} + \varphi^{-1}\left\{\dot{Q}_-^{1/2}\cos\frac{Q_-}{\varepsilon} - \frac{\varepsilon}{2}\ddot{Q}\dot{Q}_-^{-3/2}\sin\frac{Q_-}{\varepsilon}\right\}^2.$$

Our hypotheses imply $r^2(-\infty) = 1$. We now use the transition relation $\tan Q_-/\varepsilon = (1+a)\tan Q_+/\varepsilon$ to express r^2 in terms of Q_+, that is, in a form suitable for

computing $r^2(+\infty)$. We insert

$$\dot{Q}_- = (1+a)\dot{Q}_+\left\{\cos^2\frac{Q_+}{\varepsilon}+(1+a)^2\sin^2\frac{Q_+}{\varepsilon}\right\}^{-1}$$

and

$$\varepsilon\ddot{Q}_- = \dot{Q}_+^2\left(2[1+a][1-(1+a)^2]\sin\frac{Q_+}{\varepsilon}\cos\frac{Q_+}{\varepsilon}\right)$$
$$\cdot\left\{\cos^2\frac{Q_+}{\varepsilon}+(1+a)^2\sin^2\frac{Q_+}{\varepsilon}\right\}^{-2}+o(1)$$

into the previous formula. It then follows by a routine but tedious computation that $r^2(t) = 1+a+o(1)$. Thus a is the adiabatic invariant.

5. Holomorphic data. In Theorem 4.1 we have trapped the adiabatic invariant in the "transition" relation $\tan Q_-/\varepsilon = (1+a)\tan Q_+/\varepsilon$. The balance of our hard work will be to establish the existence and character of the special solutions $Q_\pm$ of $Q^2+(\varepsilon^2/2)\{Q,t\} = \varphi^2$. In the following lemma we accomplish this in the case that φ is holomorphic on a strip about the real axis and has the special form $\varphi = 1+\varepsilon^2 g$.

LEMMA 5.1. *Let g be holomorphic on the strip $|\operatorname{Im} t|<\sigma_0$ and real on the real axis. Suppose $\max_{k=0,1,2}\int_{-\infty}^{\infty}|g^{(k)}(\rho+i\sigma)|\,d\rho<M$ for $|\sigma|<\sigma_0$. Then there exists a positive $\varepsilon_0(M)$ such that for $0<\varepsilon<\varepsilon_0(M)$ the equation $\dot{Q}^2+(\varepsilon^2/2)\{Q,t\} = (1+\varepsilon^2 g)^2$ has holomorphic solutions $Q_\pm$ with the following properties*:

(i) *$Q_\pm$ are real and monotonic on the real axis,*

(ii) *$Q_\pm$ are univalent,*

(iii) $\frac{1}{2}<|\dot{Q}_\pm|<2$,

(iv) $\lim_{\rho\to\pm\infty}\dot{Q}_\pm(\rho+i\sigma)=1$, $\lim_{\rho\to\pm\infty}\ddot{Q}_\pm(\rho+i\sigma)=0$ *uniformly in σ. Moreover if g and $\tilde{g}$ satisfy*

$$\max_{k=0,1,2}\int_{-\infty}^{\infty}|g^{(k)}(\rho+i\sigma)-\tilde{g}^{(k)}(\rho+i\sigma)|\,d\rho<\mu\leqq M,$$

then the corresponding solutions $Q_\pm$ and $\tilde{Q}_\pm$ also satisfy

(v) $|\dot{Q}_\pm-\tilde{Q}_\pm|<\mu$.

Proof. We use the associated Riccati equation

$$\varepsilon\dot{p}+2i(1+\varepsilon^2 g)p = \frac{1}{2}\frac{\varepsilon^3\dot{g}}{1+\varepsilon^2 g}(1-p^2).$$

We find a solution p_- of this equation by solving the integral equation

$$p = \int_{-\infty}^{t}\exp\frac{2i}{\varepsilon}(s-t)\left[\frac{1}{2}\frac{\varepsilon^2\dot{g}}{1+\varepsilon^2 g}(1-p^2)-2i\varepsilon gp\right]ds,$$

where the path of integration is the left horizontal ray terminating at t. This equation has the form $p = F(p)$. Since $\int_{-\infty}^{\infty}|\dot{g}|<M$ implies $|g|<M'$, we find that for $\varepsilon<1/\sqrt{M'}$ the function $\frac{1}{2}(\varepsilon^2\dot{g})/(1+\varepsilon^2 g)$ is integrable. We then easily find that F maps the Banach space of bounded holomorphic functions on the strip (endowed with the supremum norm) into itself and is a contraction on the unit ball if ε is

small. It now follows from a routine application of the principle of contracting mappings that the integral equation has a solution p_- which is $O(\varepsilon^2)$ and therefore satisfies $|p_-|<\frac{1}{6}$ for small ε. Similarly we find that solutions p_- and $\tilde{p}_-$ for data g and $\tilde{g}$ satisfy $|p_- - \tilde{p}_-| = O(\mu\varepsilon^2)$ so that we can suppose $|p-\tilde{p}| \leqq \mu/6$ for small ε. The integral equation itself implies that $p_-(-\infty+i\sigma)=0$. The Riccati equation then implies $\dot{p}_-(-\infty+i\sigma)=0$ (since our hypothesis $\int_{-\infty}^{\infty}|\ddot{g}|<M$ implies $\dot{g}(-\infty+i\sigma)=0$).

To obtain Q_- we use Lemma 3.2. Let $\#$ be the involution on holomorphic functions given by $f^{\#}(t)=\overline{f(\bar{t})}$. Then by Lemma 3.2,

$$Q_- = \int_0^t (1+\varepsilon^2 g)\left(1-\frac{p}{1+p}-\frac{p^{\#}}{1+p^{\#}}\right) ds$$

is a solution of the Schwarzian equation on the real axis and hence, by analytic continuation, on the strip. Univalence for small ε follows from the inequality

$$\begin{aligned}|Q_-(t_2)-Q_-(t_1)| &\geqq \int_{t_1}^{t_2}\left(1-\frac{|p|}{1-|p|}-\frac{|p^{\#}}{1-|p^{\#}|}\right) ds \\ &\quad -\varepsilon^2\int_{t_1}^{t}|g|\left(1+\frac{|p|}{1-|p|}+\frac{|p^{\#}|}{1-|p^{\#}|}\right) ds \\ &\geqq |t_2-t_1|\left(1-2\frac{\frac{1}{6}}{1-\frac{1}{6}}-M\left(1+2\frac{\frac{1}{6}}{1-\frac{1}{6}}\right)\varepsilon^2\right) > \frac{1}{2}|t_2-t_1|.\end{aligned}$$

Plainly $|p_-|<\frac{1}{6}$ also implies $\frac{3}{5}\leqq|\dot{Q}_-|\leqq\frac{7}{5}$. Finally the limiting attributes of p_- imply that Q_- satisfies condition (iv). The inequality (v) follows from

$$\begin{aligned}(\dot{Q}-\dot{\tilde{Q}}) &= \varepsilon^2(g-\tilde{g})\left(1-\frac{\tilde{p}}{1+p}-\frac{\tilde{p}^{\#}}{1+p^{\#}}\right) \\ &\quad +(1+\varepsilon^2 g)\left(\frac{p}{1+p}-\frac{\tilde{p}}{1+\tilde{p}}+\frac{p^{\#}}{1+p^{\#}}-\frac{\tilde{p}^{\#}}{1+\tilde{p}^{\#}}\right),\end{aligned}$$

which combined with $|g-\tilde{g}|\leqq\mu$, $|p-\tilde{p}|\leqq\mu/6$ and $|p|, |\tilde{p}|\leqq\frac{1}{6}$ yields

$$|\dot{Q}-\dot{\tilde{Q}}|\leqq\tfrac{7}{5}\varepsilon^2\mu+(1+\varepsilon^2 M)\tfrac{12}{25}\mu.$$

It follows that for ε small (independently of μ) we have $|\dot{Q}-\dot{\tilde{Q}}|<\mu$.

The solution Q_+ is obtained similarly.

6. Holomorphic approximations to gentle data. Suppose ψ is a gentle function satisfying

$$\|\psi^{(m)}\|_1 \leqq \exp h(m), \qquad m=0,1,\cdots. \tag{6.1}$$

We construct a sequence of holomorphic approximations ψ_N to ψ in the following way. Let $\theta(s)$ be a smooth even real function satisfying $|\theta(t)\leqq 1|$, $\theta=1$ for $|t|\leqq 1$, $\theta=0$ for $|t|\geqq e$. Let $\hat{\psi}$ be the Fourier transform of ψ and let

$$\psi_N = \frac{1}{2\pi}\int_{-\infty}^{\infty}\theta(e^{-N}s)\,e^{its}\hat{\psi}(s)\,ds. \tag{6.2}$$

The following lemma shows the way in which the convergence of ψ_N to ψ is governed by the function h of (6.1). It also contains the first of three successive applications of convex conjugation.

LEMMA 6.1. *Let ψ satisfy* (6.1). *Then the ψ_N given by* (6.2) *are entire functions satisfying the following conditions on the sequence of strips $|\operatorname{Im} t| \leqq e^{-N}$ about the real axis.*

Let $t(\rho) = \rho + i\sigma(\rho)$ be a curve in the strip $|\operatorname{Im} t| \leqq e^{-N}$ parameterized by $\rho = \operatorname{Re}(t)$ satisfying $|dt/d\rho| \leqq C$, where C is independent of N. Then there exists a constant L independent of N such that

$$\text{(i)} \qquad \max_{0 \leqq k \leqq 4} \int_{-\infty}^{\infty} |\psi_N^{(k)}(t(\rho))|\, |dt(\rho)| \leqq L,$$

$$\text{(ii)} \quad \max_{0 \leqq k \leqq 4} \int_{-\infty}^{\infty} |\psi_{N+1}^{(k)}(t(\rho)) - \psi_N^{(k)}(t(\rho))|\, |dt(\rho)| \leqq L \exp -h^*(N-1) + 5N,$$

where h^* is the convex conjugate function of h.

Proof. We first estimate ψ_0. We find

$$\begin{aligned}
\psi_0^{(k)}(t(\rho)) &= \frac{1}{2\pi} \int_{-e}^{e} e^{i(\rho + i\sigma(\rho))s} (is)^k \theta(s) \int_{-\infty}^{\infty} e^{-is\tau} \psi(\tau)\, d\tau\, ds \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi(\tau + \rho) \int_{-e}^{e} e^{-s\sigma(\rho)} (is)^k \theta(s)\, e^{-is\tau}\, ds\, d\tau \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi(\tau + \rho) \frac{1}{1+\tau^2} \int_{-e}^{e} e^{-is\tau} \left(1 - \frac{d^2}{ds^2}\right) \{e^{-s\sigma(\rho)} (is)^k \theta(s)\}\, ds\, d\tau.
\end{aligned}$$

Since for $N = 0$ we assume $|\sigma(\rho)| \leqq 1$, we can estimate

$$\max_{0 \leqq k \leqq 4} |\psi_0^{(k)}(t(\rho))| \leqq L_1 \int_{-\infty}^{\infty} |\psi(\tau + \rho)| \frac{d\tau}{1+\tau^2}$$

which implies

$$\max_{0 \leqq k \leqq 4} \int_{-\infty}^{\infty} |\psi_0^{(k)}(t(\rho))|\, |dt(\rho)| \leqq CL_1 \int_{-\infty}^{\infty} |\psi(\rho)|\, d\rho \int_{-\infty}^{\infty} \frac{d\tau}{1+\tau^2} \leqq L_2.$$

However this argument is not suitable for large N since the length of the s interval enters into our estimate. Instead we estimate $\psi_{N+1} - \psi_N$ by an argument which uses up derivatives of ψ.

We have

$$\begin{aligned}
\psi_{N+1}^{(k)}(t) - \psi_N^{(k)}(t) = \frac{1}{2\pi} &\int_{e^N \leqq |s| \leqq e^{N+2}} (is)^k [\theta(e^{-N-1}s) - \theta(e^{-N}s)]\, e^{ist} \\
&\cdot \int_{-\infty}^{\infty} e^{-is\tau} \psi(\tau)\, d\tau\, ds.
\end{aligned}$$

Because $s = 0$ is excluded from the range of the s-integration, we can integrate by parts with respect to τ obtaining

$$\psi_{N+1}^{(k)}(t) - \psi_N^{(k)}(t) = \frac{1}{2\pi} \int_{e^N \leqq |s| \leqq e^{N+2}} (is)^{k-m} [\theta(e^{-N-1}s) - \theta(e^{-N}s)]\, e^{ist} \cdot \int_{-\infty}^{\infty} e^{-is\tau} \psi^{(m)}(\tau)\, d\tau\, ds.$$

Applying the manipulations of the previous estimation we find

$$\psi_{N+1}^{(k)}(\rho + i\sigma) - \psi_N^{(k)}(\rho + i\sigma) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi^{(m)}(\rho + \tau) \frac{1}{1+\tau^2} \cdot \int_{e^N \leqq |s| \leqq e^{N+2}} e^{-is\tau} \left(1 - \frac{d^2}{ds^2}\right) \left\{ \frac{\theta(e^{-N-1}s) - \theta(e^{-N}s)}{(is)^{m-k}} e^{-s\sigma} \right\} ds\, d\tau.$$

This implies for $m > k$,

$$|\psi_{N+1}^{(k)}(t(\rho)) - \psi_N^{(k)}(t(\rho))| \leqq L_3 m^2 e^{-N(m-k-1)} \int_{-\infty}^{\infty} \psi^{(m)}(\rho + \tau) \frac{d\tau}{1+\tau^2}$$

$$\leqq L_3 m^2 e^{-m} e^{-N(m-k-1)+m} \int_{-\infty}^{\infty} \psi^{(m)}(\rho + \tau) \frac{d\tau}{1+\tau^2}.$$

Hence

$$\max_{0 \leqq k \leqq 4} \int_{-\infty}^{\infty} |\psi_{N+1}^{(k)} - \psi_N^{(k)}|\, |dt(\rho)| \leqq L_4 \exp\{h(m) - (N-1)m + 5N\}$$

$$\leqq L_4 \exp\{-h^*(N-1) + 5N\}.$$

This establishes estimate (ii) of the lemma and also shows

$$\max_{0 \leqq k \leqq 4} \int_{-\infty}^{\infty} |\psi_{N+1}^{(k)} - \psi_N^{(k)}|\, |dt(\rho)| \leqq L_4 e^{-N} \exp\{h(7) - 7(N-1) + 5N + N\}$$

$$\leqq L_5 e^{-N}.$$

Hence for $|\operatorname{Im} t(\rho)| \leqq e^{-N}$,

$$\max_{0 \leqq k \leqq 4} \int_{-\infty}^{\infty} |\psi_N^{(k)}|\, |dt(\rho)| \leqq \sup_{0 \leqq k \leqq 4} \int_{-\infty}^{\infty} \{|\psi_N^{(k)} - \psi_{N-1}^{(k)}| + \cdots + |\psi_0^{(k)}|\} |dt(\rho)|$$

$$\leqq L_2 + L_5 \sum_{N=1}^{\infty} e^{-N}$$

$$\leqq L.$$

7. Proof of the main theorem. We now proceed to attack the equation $\dot{Q}^2 + (\varepsilon^2/2)\{Q, t\} = \varphi^2$ according to the following plan. By hypothesis $\dot{\varphi}$ is a gentle function. Let ψ_N be the sequence of approximations to $\dot{\varphi}$ provided by Lemma 6.1. Let

$$\varphi_N(t) = \varphi(-\infty) + \int_{-\infty}^{t} \psi_N(\tau)\, d\tau \tag{7.1}$$

and let

$$\Phi_N(t) = \int_0^t \varphi_N(\tau)\, d\tau. \tag{7.2}$$

Then we wish to obtain solutions of the problem

$$\dot{Q}^2 + \frac{\varepsilon^2}{2}\{Q, t\} = \varphi_N^2. \tag{7.3}$$

However our existence result, Lemma 5.1, applies only to the small perturbation form of the equation in which φ has the form $1+\varepsilon^2 g$. We therefore normalize (7.3) by the change of variable

$$S_N = \Phi_N(t), \quad Q(t) = R(S_N). \tag{7.4}$$

Formally (using Lemma 3.1) this leads to

$$R'^2 + \frac{\varepsilon^2}{2}\{R, S_N\} = 1 + \frac{\varepsilon^2}{2}\{t, S_N\}, \tag{7.5}$$

which has the required form. However before we can use this transformation we must establish its legitimacy by showing the univalence of Φ_N on a suitable domain.

The reader may well question at this point why we do not introduce such a normalization into the original problem. Our reason is simple and lies at the heart of our method. Such a procedure would leave us the task of carrying derivative estimates through our transformations. The problem of obtaining derivative estimates for $f \circ g$ in terms of estimates for f and g already illustrates this obstacle. We overcome these difficulties by introducing Jacobowitz's method of estimation at the earliest possible stage.

The following elementary lemma will ensure that the functions $\Phi_N(t)$ are univalent if we shrink the strip $|\operatorname{Im} t| \leqq e^{-N}$ by a factor independent of N.

LEMMA 7.1. *Suppose the function γ is holomorphic for $|\operatorname{Im} t| \leqq \sigma_0$, positive on the real axis and satisfies $0 < l \leqq |\gamma| \leqq L$ on the strip. Then there exist $\lambda(l, L)$ and $\mu(l, L)$ such that for $|\operatorname{Im} t| \leqq \sigma_0 e^{-\lambda}$, $\operatorname{Re} \gamma \geqq l/2$, the function $\int_0^t \gamma(s)\, ds$ is univalent for $|\operatorname{Im} t| \leqq \sigma_0 e^{-\lambda}$, and the image of this strip contains a strip $|\operatorname{Im} \int_0^t \gamma(s)\, ds| \leqq \sigma_0 e^{-\mu}$.*

Proof. We prove that on a narrower strip $|\operatorname{Im} t| \leqq \sigma_0 e^{-\lambda}$, $|\int_{t_1}^{t_2} \gamma(s)\, ds| \geqq (l/2)|t_2 - t_1|$. This inequality implies univalence and shows that the image contains a strip $|\operatorname{Im} \int_0^t \gamma(s)\, ds| \leqq (l/2)\sigma_0 e^{-\lambda} = \sigma_0 e^{-\mu}$. To establish the inequality we estimate $\dot{\gamma}$ by the Cauchy integral formula on a strip of width $\sigma_0/2$ by $|\dot{\gamma}| \leqq 2L/\sigma_0$. We next express $\int_{\rho_1+i\sigma_1}^{\rho_2+i\sigma_2} \gamma(\tau)\, d\tau$ using integrals parameterized with real variables as follows:

$$\int_{\rho_1+i\sigma_1}^{\rho_2+i\sigma_2} \gamma(\tau)\, ds = \int_{\rho_1}^{\rho_2} \gamma(s)\, ds + i\int_{\sigma_1}^{\sigma_2} \gamma(\rho_2)\, ds$$
$$+ \int_{\rho_1}^{\rho_2} [\gamma(s+i\sigma_1) - \gamma(s)]\, ds + i\int_{\sigma_1}^{\sigma_2} [\gamma(\rho_2+is) - \gamma(\rho_2)]\, ds.$$

Since γ is real on the real axis and $|\gamma| \geqq l$, we can estimate

$$\left|\int_{\rho_1+i\sigma_1}^{\rho_2+i\sigma_2} \gamma(\tau)\, d\tau\right| \geqq l[(\rho_2-\rho_1)^2+(\sigma_1-\sigma_2)^2]^{1/2}$$

$$-\frac{2}{\sigma_0}L|\sigma_1|\,|\rho_2-\rho_1|-\frac{2}{\sigma_0}L \max(|\sigma_1|, |\sigma_2|)|\sigma_2-\sigma_1|.$$

If we require $|\operatorname{Im} t| \leqq (l/(8L))\sigma_0$, this implies $\left|\int_{t_1}^{t_2} \gamma(\tau)\, d\tau\right| \geqq (l/2)|t_2-t_1|$. On this strip

$$\operatorname{Re} \gamma(\rho+i\sigma) \geqq \gamma(\rho)+\int_0^\sigma |\dot\gamma(\rho+is)|\, ds \geqq l-\frac{2L}{\sigma_0}\frac{l\sigma_0}{8L} \geqq \frac{3}{4}l.$$

COROLLARY 7.2. *There are constants l, L, λ, μ such that $0<l<\operatorname{Re}\varphi_N \leqq |\varphi_N| \leqq L$ and the Φ_N are univalent on a strip $|\operatorname{Im} t| \leqq e^{-N-\lambda}$. The image of this strip contains a strip $|\operatorname{Im}\Phi_N| \leqq e^{-N-\mu}$.*

Proof. By hypothesis $0<l' \leqq \varphi \leqq (-\infty)+\int_{-\infty}^{\infty}|\dot\varphi|\, dt = L'$. Moreover $\lim_{N\to\infty}|\varphi-\varphi_N| \leqq \lim_{N\to\infty}\int_{-\infty}^{\infty}|\dot\varphi-\psi_N|\, dt = 0$. Hence for large N, $0<l'/2 \leqq |\varphi_N| \leqq 2L'$. Since the conclusion of the previous lemma depends only on upper and lower bounds for $|\gamma|$, the corollary follows.

With this corollary we have justified the change of variables (7.4) transforming (7.3) into (7.5). Our aim is now to apply Lemma 5.1 to (7.5) which has the form $R'^2+(\varepsilon^2/2)\{R, s\} = (1+\varepsilon^2 g_N)^2$, where

$$g_N(s, \varepsilon) = \varepsilon^{-2}[(1+(\varepsilon^2/2)\{t, \Phi_N\})^{1/2}-1]. \tag{7.6}$$

We therefore must establish that g_N possesses the integrability conditions along horizontal lines required by the hypotheses of Lemma 5.1.

LEMMA 7.3. *There exist constants $M_1, \mu_1, N_1, \varepsilon_1$ such that for $0<\varepsilon \leqq \varepsilon_1$, $N \geqq N_1$ and $|\operatorname{Im} s| \leqq e^{-N-\mu_1}$, the functions $g_N(s, \varepsilon)$ defined by* (7.6) *satisfy*

$$\text{(i)} \qquad \sup_{k=0,1,2}\int_{-\infty}^{\infty}|g_N^{(k)}(\rho+i\sigma)|\, d\rho \leqq M_1,$$

$$\text{(ii)} \qquad \sup_{k=0,1,2}\int_{-\infty}^{\infty}|g_{N+1}^{(k)}(\rho+i\sigma)-g_N^{(k)}(\rho+i\sigma)|\, d\rho \leqq M_1 \exp(-h^*(N-1)+6N).$$

Proof. The integrability of g_N and its first two derivatives along the level curves of $\operatorname{Im} s$ will follow from corresponding integrability conditions for $\psi_N = \dot\varphi_N$ and its first four derivatives along the level curves of $\operatorname{Im}\Phi_N(t)$ in the t-plane. To obtain the latter integrability condition from the conclusion of Lemma 6.1, we must investigate these level curves. By Corollary 7.2 we can suppose that these curves lie in a strip $|\operatorname{Im} t| \leqq e^{-N-\lambda}$. Also since ds is real along these curves,

$$\left|\frac{dt}{d\operatorname{Re} t}\right| = \left|\frac{dt}{ds}\right|\left|\operatorname{Re}\frac{dt}{ds}\right|^{-1} = |\varphi_N|\,|\operatorname{Re}\varphi_N|^{-1}.$$

Hence by Corollary 7.2, $|dt/d\operatorname{Re} t| \leqq L/l$. The integrability of ψ_N and its first four derivatives then follows from Lemma 6.1. Specifically, since the derivatives of $\dot\varphi$ are estimated by $\|\dot\varphi^{(m)}\|_1 \leqq \exp h(m+1)$, we have the existence of a constant M

such that

$$\int_{\operatorname{Im}\Phi_N=c} |\psi_N^{(k)}(t)|\,dt| \leqq M, \qquad k=0,1,\cdots,4,$$

$$\int_{\operatorname{Im}\Phi_N=c} |\psi_{N+1}^{(k)}(t)-\psi_N^{(k)}(t)|\,d|t| \leqq M\exp -(h(m+1))^*(N-1)+5N \tag{7.7}$$

$$\leqq M\exp\{-h^*(N-1)+6N\}.$$

We are now in position to estimate the integrals of $g_N^{(k)}(s)$ by parametizing them in the t-plane. By Lemma 3.1(iv),

$$g_N(s(t)) = \varepsilon^{-2}\left[\left(1-\frac{\varepsilon^2}{2}\{\Phi_N, t\}\varphi_N^{-2}\right)^{1/2}-1\right]$$

$$= \varepsilon^{2}\left[\left(1-\frac{\varepsilon^2}{2}\left[\frac{\dot\psi_N}{\varphi_N^3}-\frac{\psi_N^2}{\varphi_N^4}\right]\right)^{1/2}-1\right].$$

Since for large N we have a lower bound on $|\varphi_N|$ and upper bounds on ψ_N, $\dot\psi_N$, for ε sufficiently small $g_N(s(t))$ is holomorphic in t for $[\operatorname{Im} t| \leqq e^{-N-\lambda}$. Now

$$g_N^{(k)}(s)\,ds = \left(\varphi_N(t)\frac{d}{dt}\right)^k g_N(s(t))\varphi_N(t)\,dt. \tag{7.8}$$

For $k=0,1,2$ this expresses $g_N^{(k)}(s)\,ds$ in terms of φ_N, ψ_N and its first three derivatives. We have also provided that $\psi_N^{(4)}$ is integrable, so that the lower derivatives are also bounded. Combining (7.7) and (7.8) we obtain, by elementary but tedious calculations, the required estimates.

We next appeal to Lemma 5.1 to obtain solutions of $R'^2+(\varepsilon^2/2)\{R,s\}=(1+\varepsilon^2 g_N)^2$ (and hence solutions $Q=R\circ\Phi_N$ of $\dot Q^2+(\varepsilon^2/2)\{Q,t\}=\varphi_N^2$).

LEMMA 7.4. *Equation* (7.5) *has solutions* $R_{N\pm}$ *satisfying the following conditions for* $0<\varepsilon<\varepsilon_2$ *and* $N>N_2$:

(i) $R_{N\pm}$ *are real and monotonic increasing on the real axis.*

(ii) $R_{N\pm}$ *is univalent for* $|\operatorname{Im} s| \leqq e^{-N-\mu_2}$.

(iii) $\frac{1}{2}<|R_{N\pm}|<2$ *for* $|\operatorname{Im} s| \leqq e^{-N-\mu_2}$.

(iv) $\lim_{\rho\to\pm\infty} R'_{N\pm}(\rho+i\sigma)=1$, $\lim_{\rho\to\pm\infty} R''_{N\pm}(\rho+i\sigma)=0$ *uniformly in* σ.

(v) $|R'_{N+1,\pm}-R'_{N\pm}|<M_2\exp(-h^*(N-1)+6N)$ *for* $|\operatorname{Im} s| \leqq e^{-N-1-\mu_2}$.

Proof. This follows directly from the previous lemma and Lemma 5.1.

The existence of solutions of the Schwarzian equation together with derivative estimates follows almost immediately.

LEMMA 7.5. *Let* $\Phi(t)=\int_0^t \varphi(s)\,ds$. *There exist solutions* $R_\pm$ *of* $R'^2+(\varepsilon^2/2)\{R,s\}=1+(\varepsilon^2/2)\{\Phi^{-1},s\}$ *satisfying* $R'_\pm(\pm\infty)=1$, $R''_\pm(\pm\infty)=0$ *and*

$$|R_\pm^{(k)}| \leqq \exp\{h(k+7)+k\log k+ck+d\}.$$

Proof. By conclusion (v) of Lemma 7.4 we can estimate $|R_{N+1,\pm}^{\prime(m)}-R_{N\pm}^{\prime(m)}|$ on the narrower strip $|\operatorname{Im} s| \leqq e^{-N-2-\mu_2}$ by the Cauchy integral formula

$$f^{(m)}(s)=\frac{m!}{2\pi i}\int\frac{f(\sigma)}{(s-\sigma)^{m+1}}\,ds,$$

where we choose the path of integration to be a square with horizontal sides on $|\operatorname{Im} s| = e^{-N-1-\mu_2}$ with s at its center. Then on the narrower strip we estimate $|s-\sigma| > e^{-N-2-\mu_2}$ and

$$|R'^{(m)}_{N+1} - R'^{(m)}_N| \leqq M_4 \exp -h^*(N-1) + 6N + m(N+2+\mu_2) + m \log m$$

which can be expressed

$$|R'^{(m)}_{N+1,\pm} - R'^{(m)}_{N\pm}| \leqq M_5 e^{-N} \exp\{-h^*(N-1) + (m+7)(N-1) + m \log m + (3+\mu_2)m\mu_2\}.$$

We now appeal to a basic property of convex conjugation [7]: if h is convex, then $h^{**} = h$. Since $h^*(N-1) - (m+7)(N-1) \geqq -h^{**}(m+7)$, we find $|R'^{(m)}_{N+1,\pm} - R'^{(m)}_{N\pm}| \leqq (M_6/e^N) \exp h(m+7) + m \log m + cm$. It follows that the sequence $R'^{(m)}_{N+1,\pm}$ converges uniformly to a function $R'_\pm$ satisfying

$$|R'^{(m)}_\pm| \leqq M_3 \exp\{h(m+7) + m \log m + cm\}.$$

Moreover the convergence of φ_N to φ and $R'_{N\pm}$ to $R'_\pm$ together with convergence of all derivatives ensures that $R_\pm$ satisfy the differential equation and limiting conditions.

From this point on we require derivative estimates for the combination $\xi(t) = \{(d/dt)Q_- \circ Q_+^{\text{inverse}}\}^{-1}$. Since $Q_\pm = R_\pm \circ \Phi$ we have $Q_- \circ Q_+^{\text{inverse}} = R_- \circ R_+^{\text{inverse}}$. Therefore we have the pleasant simplification that we can estimate this function directly in terms of $R_\pm$. This we accomplish by again following our basic method of assessing the approximations $\{(d/dt)R_{N-} \circ R_{N+}^{\text{inverse}}\}^{-1} = \xi_N$.

LEMMA 7.6. *For $0 < \varepsilon < \varepsilon_3$ and $N > N_3$, the functions ξ_N satisfy an estimate*

$$|\xi_{N+1}(t) - \xi_N(t)| < \exp(-h^*[N-1] + 6N + c_3)$$

for $|\operatorname{Im} t| \leqq e^{-N-1-\mu_3}$ and $|\operatorname{Re} t| \leqq 1$.

Proof. This is a consequence of estimate (v) of Lemma 7.4 and the fact that we can estimate $R_{N+1} - R_N$ by the estimates for $R'_{N+1} - R'_N$ on the bounded t-domain under consideration.

LEMMA 7.7. *The function $\xi(t) = \{(d/dt)R_- \circ R_+^{\text{inverse}}(t)\}^{-1}$ satisfies the estimate*

$$\max_{|t| \leqq 1} |\xi^{(m)}| \leqq \exp\{h(m+7) + m \log m + cm + d\}.$$

Proof. The proof is analogous to that of Lemma 7.5.

We now have the resources to prove the main theorem. By Lemma 3.4 we can suppose $\tan(R_-/\varepsilon) = (1+a)\tan(R_+/\varepsilon)$. This implies

$$\xi(t) = \left\{\frac{d}{dt}\varepsilon \tan^{-1}(1+a)\tan\frac{t}{\varepsilon}\right\}^{-1}$$

$$= \frac{1}{2}\left(1 + a + \frac{1}{1+a}\right) + \frac{1}{2}\left(\frac{1}{1+a} - 1 - a\right)\cos\frac{2t}{\varepsilon}.$$

Hence, computing $\xi^{(m)}(0)$ if m is even or $\xi^{(m)}(\varepsilon(\pi/4))$ if m is odd, we find using Lemma 7.7 that

$$\frac{1}{2}\left|1 + a - \frac{1}{1+a}\right| \leqq \left(\frac{\varepsilon}{2}\right)^m \exp\{h(m+7) + m \log m + cm + d\}.$$

Since $(m \log m)/(h(m+7)) < \delta/(1-\delta)$ for large m, this implies

$$\begin{aligned}\frac{1}{2}\left|1+a-\frac{1}{1+a}\right| &\leqq M \exp \frac{1}{1-\delta} h(m+7) - m \log \frac{1}{\varepsilon} \\ &\leqq M \exp\left\{-\frac{1}{1-\delta} h^*(\log \varepsilon^{+\delta-1}) + 7 \log \frac{1}{\varepsilon}\right\} \\ &\leqq M \exp\left\{-h^*(\log \varepsilon^{\delta-1})\right\} \exp\left\{\frac{-\delta}{1-\delta} h^*\left((1-\delta)\log\frac{1}{\varepsilon}\right) + 7 \log \frac{1}{\varepsilon}\right\}.\end{aligned}$$

But $\sup\{(-\delta/(1-\delta))h^*[(1-\delta)x]+7x\} < \infty$ since otherwise $h^{**}(7/\delta) = h(7/\delta) = \infty$. Hence

$$\left|1+a-\frac{1}{1+a}\right| = O\left\{\exp -h^*(\log \varepsilon^{\delta-1})\right\},$$

which implies

$$a(\varepsilon) = O\{\exp -h^*(\log \varepsilon^{\delta-1})\}.$$

We have thus estimated $a(\varepsilon)$ for the special solution $u = \dot{Q}_-^{-1/2} \sin Q_-/\varepsilon$. Since translates $\varphi(t+t_0)$ of $\varphi(t)$ obey the same estimates, this suffices to estimate $a(\varepsilon)$ for the general solution $u = C\dot{Q}_-^{-1/2} \sin((Q_-/\varepsilon)+t_0)$.

REFERENCES

[1] J. E. LITTLEWOOD, *Lorentz's pendulum problem*, Ann. Physics, 21 (1963), pp. 233–242.
[2] E. HILLE, *Lectures on Ordinary Differential Equations*, Addison-Wesley, Reading, Mass., 1969.
[3] W. WASOW, *Adiabatic invariance of a simple oscillator*, this Journal, 1 (1970), pp. 153–170.
[4] ———, *Calculation of an adiabatic invariant by turning point theory*, MRC Tech. Summary Rep. 1217, Univ. of Wisconsin, Madison, 1973.
[5] R. E. MEYER, *Adiabatic variation, Part I. Exponential property for the simple oscillator*, Rep. 39, Fluid Mechanics Research Institute, Univ. of Essex, England, 1973.
[6] H. JACOBOWITZ, *Implicit function theorems and isometric embeddings*, Ann. of Math., 95 (1972), pp. 191–225.
[7] J. STOER AND C. WITZGALL, *Convexity and Optimization in Finite Dimensions I*, Springer-Verlag, New York, 1970.

SIAM J. MATH. ANAL.
Vol. 8, No. 4, August 1977

ASYMPTOTIC EXPANSIONS OF INTEGRAL TRANSFORMS OF FUNCTIONS WITH LOGARITHMIC SINGULARITIES*

NORMAN BLEISTEIN†

Abstract. Asymptotic expansions of integral transforms of functions with logarithmic singularities are defined. The method is based on the Mellin transform technique developed by Handelsman and Lew. Examples included are Laplace, Airy, Weber and Stieltjes transforms.

1. Introduction. We shall develop a technique for the asymptotic expansions of integrals of the form

$$I(\lambda)=\int_0^\infty h(\lambda t)f(t)\,dt, \qquad \lambda\to\infty. \tag{1.1}$$

We are concerned here with functions $f(t)$ which have logarithmic singularities at $t=0^+$.

A quite general class of such functions $I(\lambda)$ has been treated in the literature in a series of papers by Handelsman and Lew (1969), (1971). This work is also discussed in detail in Bleistein and Handelsman (1975). In that work, $f(t)$ is assumed to have an asymptotic expansion of the form

$$f(t)\sim\sum_{m=0}^{\infty}\sum_{n=0}^{N(m)}c_{mn}t^{\alpha_m}(\log t)^n, \qquad t\to 0^+. \tag{1.2}$$

Here, Re $\alpha_m\uparrow+\infty$ and $N(m)$ is finite for each m.

We note that the powers of $\log t$ are nonnegative integers here, and there are finitely many of them associated with each power of t.

We shall extend these results in two ways. First, we shall consider rational functions of $\log t$, for which a prototype is the function

$$f(t)=\frac{t^\alpha}{a-\log t}, \qquad 0<|\arg a|<\pi. \tag{1.3}$$

Bouwkamp (1971) has treated a special case of this type in which h is the exponential function ($I(\lambda)$ is then the Laplace transform) and

$$f(t)=\frac{t^{\alpha-1}}{a^2+(\log t)^2}, \tag{1.4}$$

which is a sum of functions of the type (1.3).

We remark that the latter case is not subsumed under (1.2) because the series expansion for (1.4) has an infinite set of powers of $\log t$, including some negative powers, associated with one single power of t.

Secondly, we shall consider functions $f(t)$ which have asymptotic expansions

* Received by the editors January 27, 1975, and in final revised form November 3, 1975.

† Department of Mathematics, College of Arts and Science, Division of Mathematical Sciences, Denver Research Institute, University of Denver, Denver, Colorado 80210. This research was supported in part by the Office of Naval Research under Contracts N00014-67-A-0394-0005 and N00014-76-C-0039.

near the origin of the form

$$f(t)\sim \sum_{m=0}^{\infty}\sum_{n=0}^{N(m)} c_{mn}t^{\alpha_m}(\log t)^{\beta_{mn}}, \qquad t\to 0^+. \tag{1.5}$$

Here α_m and $N(m)$ are as in (1.2), but the β_{mn}'s may be any complex numbers.

Recently, Olmstead and Handelsman (1976) have obtained a leading term of the form $O((\log t)^{-1})$ as the solution of a nonlinear Volterra integral equation. From the formal nature of their result, it is not clear whether or not their expansion is one of the two types cited here. Nonetheless, it was a discussion with Olmstead which, in part, motivated this paper.

To derive the asymptotic expansion of (1.1), the asymptotic technique employed by Handelsman and Lew (1969), (1971) will be used. In this method, the Mellin transforms of the functions f and h play a crucial role. For proofs of the results about Mellin transforms stated here, we refer the reader to Titchmarsh (1948) and the aforementioned papers by Handelsman and Lew.

The Mellin transform of a function $f(t)$ evaluated at a point z is defined by the integral

$$M[f;z]=\int_0^{\infty} f(t)t^{z-1}\,dt, \qquad z=x+iy. \tag{1.6}$$

When the transform exists, it does so and is analytic in an open vertical strip (perhaps semi-infinite or infinite horizontally) in the complex z-plane.

The deleted argument of f on the left side in (1.6) is understood to be t itself. Other arguments will be shown explicitly. When employing Mellin transforms to analyze (1.1) the result

$$M[h(\lambda t);z]=\lambda^{-z}M[h;z] \tag{1.7}$$

is needed, as well as the Mellin–Parseval theorem. In terms of the integral (1.1), we may state this theorem as

$$I(1)=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} G(z)\,dz, \tag{1.8}$$

$$G(z)=M[h;z]M[f;1-z]. \tag{1.9}$$

Here, the Bromwich contour, $c-i\infty$ to $c+i\infty$, is in the common strip of analyticity (whose existence we assume) of the two Mellin transforms appearing in the integrand. Indeed, the existence of the two Mellin transforms and of $I(1)$ assure the existence of such a number c.

When we combine (1.7), (1.8) and (1.9), we find

$$I(\lambda)=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \lambda^{-z}G(z)\,dz. \tag{1.10}$$

The integral is observed to be $O(\lambda^{-c})$ and even $o(\lambda^{-c})$ if $G(z)$ is absolutely integrable. The latter estimate follows from writing λ^{-z} as $\lambda^{-c}\exp\{-iy\log\lambda\}$ and invoking the Riemann–Lebesgue lemma for $\log\lambda\to\infty$.

To obtain the asymptotic expansion of $I(\lambda)$, we attempt to replace the Bromwich contour by another, further to the right, say with c replaced by $c+k$.

This new integral would be $O(\lambda^{-c-k})$. Thus the asymptotic expansion to this order must arise from the deformations of the Bromwich contour imposed by avoidance of the singularities of the analytic continuation of the function $G(z)$.

In specific examples, this method has been used by others prior to Handelsman and Lew. The major contribution of these two authors was to show that for an extremely broad class of functions, f and h, the analytic continuation of $G(z)$ was, at worst, a meromorphic function. Consequently, the asymptotic expansion of $I(\lambda)$ to any order arises as a series of residues at the poles of $G(z)$. For the functions to be considered here, the analytic continuation is no longer meromorphic. Consequently, the asymptotic expansion of $I(\lambda)$ will be considerably more complicated.

Erdélyi (1956) extended the method of stationary phase to the case in which the amplitude contained the first power of $\log t$. (A minor error in the result was corrected by McKenna (1965).) In another paper, Erdélyi (1961) derived asymptotic expansions of Laplace transforms of functions of the type (1.5). Both papers rely on direct analysis of the Laplace transform of the function $(\log t)^{\alpha}t^{\beta-1}$. Wong (1970) presents an alternative proof of a major theorem in the Erdélyi paper.

D. S. Jones (1969) gives two examples (Fourier and Hankel transforms) with first power of $\log t$ in the amplitude function. His book on generalized functions (1966) contains some examples of Fourier transforms of functions of this type.

Wong and Wyman (1972) consider Laplace transforms of functions of the type (1.5). They concern themselves with a technical problem to which we only allude, below Theorem 5.

Riekstins (1974) considers Laplace transforms of functions of $\log t$ alone (a special case of (1.5)). He also considers more general transforms of these functions for small argument, but with the domain of integration being $(1, \infty)$. Finally, he considers kernels which are functions of $\log t$ alone. Thus there is some overlap between his results and ours, although our methods are quite different.

Wong (1975) considers Laplace transforms near the origin of functions whose asymptotic expansions near $t = \infty$ involve an infinite set of negative integer powers of $\log t$, but only one algebraic power of t. For small values of the transform variable, the Laplace transform may be viewed as an algebraic kernel. Thus Wong's results are closely related to the results on algebraically dominated kernels in § 4, below.

2. Rational functions and exponentially decaying kernels. We consider first the integral (1.1) with f given by (1.3). We shall assume that h is an "exponentially decaying kernel", i.e.,

$$h(t) = O(t^{-r}(\log t)^{N} \exp(-st^{\nu})), \qquad t \to \infty. \tag{2.1}$$

Here, s and ν are positive constants. The algebraic and logarithmic factors play no crucial role in the discussion below, but are inserted here only to emphasize the general nature of the kernels being considered. Furthermore,

$$h(t) = O(t^{\gamma}), \qquad t \to 0^{+}, \quad \operatorname{Re}(\alpha + \gamma) > -1, \tag{2.2}$$

and $h(t)$ is locally integrable on $(0, \infty)$.

Let us set

$$I(\lambda) = I_1(\lambda) + I_2(\lambda), \tag{2.3}$$

$$I_j(\lambda) = \int_0^\infty h(\lambda t) f_j(t)\, dt, \qquad j = 1, 2, \tag{2.4}$$

with the functions f_j defined by

$$f_1(t) = \begin{cases} f(t), & 0 < t < 1, \\ 0, & 1 \leqq t < \infty, \end{cases}$$
$$f_2(t) = \begin{cases} 0, & 0 < t < 1, \\ f(t), & 1 \leqq t < \infty. \end{cases} \tag{2.5}$$

As a consequence of (2.1),

$$I_2(\lambda) = O(\exp(-\lambda^\nu(1-\varepsilon))), \tag{2.6}$$

any $\varepsilon > 0$. Thus we proceed by focusing our attention on the integral $I_1(\lambda)$. By applying the Mellin–Parseval theorem (1.8), (1.9) to $I_1(\lambda)$, (2.4), we find

$$I_1(\lambda) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \lambda^{-z} M[h; z] M[f_1; 1-z]\, dz, \qquad -\operatorname{Re}\gamma < c < \operatorname{Re}\alpha + 1. \tag{2.7}$$

Here c is chosen in the strip determined by the overlapping half-plane of analyticity of the Mellin transforms and

$$M[f_1; 1-z] = e^{a(\alpha+1-z)} E_1(a(\alpha+1-z)), \qquad x < \operatorname{Re}\alpha + 1, \tag{2.8}$$

with $E_1(z)$ the exponential integral (Erdélyi et al. (1953, p. 143 ff.)),

$$e^z E_1(z) = \int_0^\infty \frac{e^{-\sigma}}{z+\sigma}\, d\sigma, \qquad |\arg z| < \pi. \tag{2.9}$$

The analytic continuation of $M[f_1; 1-z]$ to the right is explicit in (2.9). Its properties follow from well-known properties of the exponential integral. We list them below.

(i) $M[f_1; 1-z]$ has a logarithmic branch point at $z = 1+\alpha$. (We take the branch cut from $\alpha+1$ to extend horizontally to $+\infty$.) Indeed,

$$M[f_1; 1-z] = -e^{a(\alpha+1-z)} \log(\alpha+1-z) + F(z), \tag{2.10}$$

with $F(z)$ analytic in the right half-plane.

(ii) In the right half-plane, $x > \operatorname{Re}\alpha + 1$,

$$M[f_1; 1-z] = O(|z|^{-1}), \qquad |z| \to \infty. \tag{2.11}$$

The function $M[h; z]$ is known to be analytic in the right half-plane $x > -\operatorname{Re}\gamma$ and also

$$|M[h; z]| \to 0 \quad \text{as } |y| \to \infty. \tag{2.12}$$

These estimates suffice to justify the replacement of $I_1(\lambda)$ by a contour integral over the path $P_1+P_2+P_3$ in Fig. 1 and

$$x = \operatorname{Re}\alpha + 1 + k \quad \text{on } P_2, P_3, \quad \text{any } k > 0. \tag{2.13}$$

Consequently,

$$I_1(\lambda) = -\frac{1}{2\pi i}\int_{P_1} e^{-\mu z + a(\alpha+1-z)} \log(\alpha+1-z) M[h; z]\, dz + O(\lambda^{-\alpha-1-k}). \tag{2.14}$$

Here, the error estimate arises from the integrals over P_2 and P_3, and we have set

$$\mu = \log \lambda. \tag{2.15}$$

The integral over P_1 can be replaced by an integral from $\alpha+1$ to $\alpha+1+k$. With $\alpha+1$ as the origin of coordinates, the result is

$$I_1(\lambda) = \lambda^{-\alpha-1}\int_0^k e^{-(\mu+a)\zeta} M[h; \alpha+1+\zeta]\, d\zeta + O(\lambda^{-\alpha-1-k}). \tag{2.16}$$

The asymptotic expansion of this integral for large μ may be obtained by Watson's lemma. This yields the following result for $I(\lambda)$:

$$I(\lambda) \sim \lambda^{-\alpha-1} \sum_{m=0}^{\infty} c_m (\log \lambda)^{-m-1} + O(\lambda^{-\alpha-1-k}) + O(\exp(-\lambda^{\nu}(1-\varepsilon))). \tag{2.17}$$

Here the coefficients are defined by

$$e^{-a\zeta} M[h; \alpha+1+\zeta] = \sum_{m=0}^{\infty} \frac{c_m}{m!}\zeta^m, \tag{2.18}$$

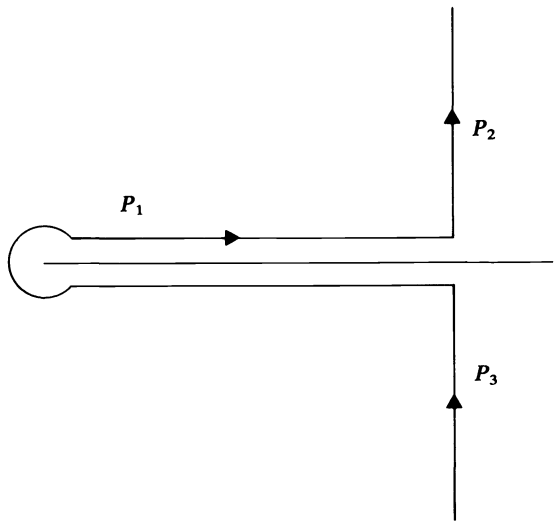


FIG. 1

and the series in (2.17) is an asymptotic series. Alternatively, we may view $\mu + a$ as the large parameter in (2.16) and write

$$I(\lambda) \sim \lambda^{-\alpha-1} \sum_{m=0}^{\infty} \tilde{c}_m (a + \log \lambda)^{-m-1} + O(\lambda^{-\alpha-1-k}) + O(\exp(-\lambda^{\nu}(1-\varepsilon))), \tag{2.19}$$

with $\tilde{c}_m$'s defined by

$$M[h; \alpha + 1 + \zeta] = \sum_{m=0}^{\infty} \frac{\tilde{c}_m}{m!} \zeta^m. \tag{2.20}$$

The latter set of coefficients may be easier to calculate than the former.

We have now proved the following.

THEOREM 1. *Let*

$$I(\lambda) = \int_0^{\infty} \frac{h(\lambda t) t^{\alpha}}{a - \log t}\, dt, \tag{2.21}$$

with $h(t)$ being locally integrable and satisfying (2.1) *and* (2.2). *Then the asymptotic expansion of $I(\lambda)$ with respect to the sequence $\{\lambda^{-\alpha-1}(\log \lambda)^{-m-1}\}$ is given by* (2.17), (2.18); *the asymptotic expansion with respect to the sequence $\{\lambda^{-\alpha-1}(a + \log \lambda)^{-m-1}\}$ is given by* (2.19), (2.20).

As examples of this result, we present the following. The relevant Mellin transforms appear in Erdélyi et al. (1953) or can be derived from results appearing there.

Example 1. *The Laplace transform.* Here

$$h(t) = \exp(-t), \quad M[h; z] = \Gamma(z), \qquad x > 0 = -\gamma. \tag{2.22}$$

Thus the asymptotic expansion of $I(\lambda)$ is given by (2.17) with

$$e^{-a\zeta}\Gamma(\alpha + 1 + \zeta) = \sum_{m=0}^{\infty} \frac{c_m}{m!} \zeta^m, \tag{2.23}$$

or by (2.19) with

$$\Gamma(\alpha + 1 + \zeta) = \sum_{m=0}^{\infty} \frac{\tilde{c}_m \zeta^m}{m!}. \tag{2.24}$$

Returning to Bouwkamp's example now, we set

$$\frac{t^{\sigma-1}}{a^2 + (\log t)^2} = -\frac{t^{\sigma-1}}{2ia}\left[\frac{1}{ia - \log t} - \frac{1}{-ia - \log t}\right]. \tag{2.25}$$

Thus the asymptotic expansion is a sum of integrals of the type considered here, with a replaced by $\pm ia$ and α replaced by $\sigma - 1$. The result is

$$\int_0^{\infty} \frac{t^{\sigma-1} e^{-\lambda t}}{a^2 + (\log t)^2}\, dt \sim \lambda^{-\sigma} \sum_{m=0}^{\infty} c_m (\log \lambda)^{-m-1}. \tag{2.26}$$

Here, by combining the two generating functions, we find that the c_m's are defined by

$$\frac{\sin a\zeta}{a}\Gamma(\sigma+\zeta)=\sum_{m=0}^{\infty}\frac{c_m}{m!}\zeta^m, \tag{2.27}$$

and this agrees with Bouwkamp's result. Alternatively,

$$\int_0^{\infty}\frac{t^{\sigma-1}e^{-\lambda t}}{a^2+(\log t)^2}dt\sim-\lambda^{-\sigma}a^{-1}\sum_{m=0}^{\infty}\tilde{c}_m\,\mathrm{Im}\,(ai+\log\lambda)^{-m-1} \tag{2.28}$$

with

$$\Gamma(\sigma+\zeta)=\sum_{m=0}^{\infty}\frac{\tilde{c}_m\zeta^m}{m!}. \tag{2.29}$$

We remark that Bouwkamp's result is valid for $\sigma\geqq 0$ ($\alpha\geqq -1$) while ours requires strict inequality. This follows from the fact that $x^{-1}(\log x)^{-2}$ has an integrable singularity at $x=0$ while $x^{-1}(\log x)^{-1}$ does not. Thus the flaw arises because we choose to derive the result by combining two fractions (2.25) to obtain one. We may remedy this in the following way. We observe in (2.27) that the expansion on the right remains valid even when $\sigma=0$ because

$$\frac{\sin a\zeta}{a}\Gamma(\zeta)=\frac{\pi\sin a\zeta}{a\sin\pi\zeta}\frac{1}{\Gamma(1-\zeta)} \tag{2.30}$$

has a removable singularity at $\zeta=0$. Thus we extend our results to $\sigma=0$ by analytic continuation.

Alternatively, we could repeat the entire derivation for such functions. However, the extension here also depends on the nature of the Mellin transform of the kernel (2.22) and the fact that this function has a simple pole at the left boundary of its domain of analyticity in ζ. Thus the extension to $\sigma=0$ follows from the happy coincidence of a zero of one Mellin transform and the pole of the other. This would not seem to be sufficiently likely, in general, to justify an extension of our theory.

Example 2. *K-transform.* Here, we set

$$h(t)=K_{\mu}(t), \tag{2.31}$$

with μ real. For this function,

$$M[h;z]=2^{z-2}\Gamma\left(\frac{z-\mu}{2}\right)\Gamma\left(\frac{z+\mu}{2}\right),\qquad x>|\mu|. \tag{2.32}$$

Thus

$$\int_0^{\infty}\frac{t^{\alpha}K_{\mu}(\lambda t)}{a-\log t}dt\sim\lambda^{-\alpha-1}\sum_{m=0}^{\infty}d_m(\log\lambda)^{-m-1},\qquad \mathrm{Re}\,\alpha+1>|\mu|, \tag{2.33}$$

$$e^{-a\zeta}2^{\alpha+\zeta-1}\Gamma\left(\frac{\alpha+1+\zeta-\mu}{2}\right)\Gamma\left(\frac{\alpha+1+\zeta+\mu}{2}\right)=\sum_{m=0}^{\infty}\frac{d_m}{m!}\zeta^m. \tag{2.34}$$

For the next two integrals, we only list the results.

Example 3. "*Airy transform*".

$$\int_0^\infty \frac{t^\alpha \operatorname{Ai}(\lambda t)}{a-\log t}\,dt \sim \lambda^{-\alpha-1} \sum_{m=0}^{\infty} e_m (a+\log\lambda)^{-m-1}, \qquad \operatorname{Re}\lambda > -1, \tag{2.35}$$

$$(2\pi)^{-1} 3^{2z/3-(7/6)} \Gamma\left(\frac{z}{3}\right)\Gamma\left(\frac{z+1}{3}\right) = \sum_{m=0}^{\infty} \frac{e_m}{m!}(z-\alpha-1)^m. \tag{2.36}$$

Example 4. "*Weber transform*".

$$\int_0^\infty \frac{t^\alpha D_\mu(\lambda t)}{a-\log t}\,dt \sim \lambda^{-\alpha-1} \sum_{m=0}^{\infty} f_m (a+\log\lambda)^{-m-1}, \qquad \operatorname{Re}\alpha > -1, \tag{2.37}$$

$$\frac{\Gamma(z)\sqrt{\pi}\,2^{(z+1+\mu)/2}}{\Gamma((z+1-\mu)/2)} F\left(\frac{z+1}{2}, \frac{1-\mu}{2}; \frac{z+1-\mu}{2}; -1\right) = \sum_{m=0}^{\infty} \frac{f_m}{m!}(z-\alpha-1)^m. \tag{2.38}$$

Here $F(a, b; c; x)$ is the hypergeometric function.

We can formally calculate the asymptotic expansion of the integral

$$I^{(n)}(\lambda) = \int_0^\infty \frac{t^\alpha h(\lambda t)}{(a-\log t)^{n+1}}\,dt, \qquad n = 1, 2, \cdots, \tag{2.39}$$

by observing that, for I given by (1.1) and f given by (1.3),

$$I^{(n)}(\lambda) = \left(-\frac{d}{da}\right)^n I(\lambda). \tag{2.40}$$

For this calculation, the form (2.19) is more useful, since the dependence of the result on a is explicit. To justify this result, we define

$$f_{n+1}(t) = \frac{t^\alpha}{(a-\log t)^{n+1}}. \tag{2.41}$$

Again, using Erdélyi et al. (1953), we find

$$M[f_{n+1}; 1-z] = -(z-\alpha-1)^n e^{a(\alpha+1-z)} \log(\alpha+1-z) + F(z), \tag{2.42}$$

with $F(z)$ analytic in the right half-plane $x > \operatorname{Re}\alpha + 1$. Thus the analysis through (2.15) proceeds as above, except that now

$$I(\lambda) = (-1)^n \lambda^{-\alpha-1} \int_0^k e^{-(\mu+a)\zeta} \zeta^n M[h; \alpha+1+\zeta]\,d\zeta + O(\lambda^{-\alpha-1-k}), \tag{2.43}$$

for which the asymptotic expansion via Watson's lemma leads to the result formally conjectured above. Thus we have proved

LEMMA 1. *Let $I(\lambda)$ be given by* (2.39). *Then, as $\lambda \to \infty$, under the conditions of Theorem* 1,

$$I(\lambda) \sim \lambda^{-\alpha-1} \sum_{m=0}^{\infty} \tilde{c}_m m_n (-1)^n (a+\log\lambda)^{-m-n-1}. \tag{2.44}$$

Here the $\tilde{c}_m$'s are defined in (2.20) *and*

$$m_n = (m+n)!/m! \tag{2.45}$$

Let us suppose now that $f(t)$ is a rational function of $\log t$. Then a partial fraction expansion leads to a sum of integrals with integrands of the type (2.41). To leading order, then, we obtain a sum of expansions of the type (2.44). We retain in this sum only those series for which the real parts of the α's agree.

3. Complex λ. We shall consider here the extension of the results of § 2 to complex λ. First, we note from (2.6) that $I_2(\lambda)$ is asymptotically zero with respect to either of the asymptotic sequences of Theorem 1, only so long as $|\arg \lambda| < \pi/2\nu$. However, that result may be further restricted by the sector of validity of the asymptotic expansion of $I_1(\lambda)$. Thus we consider now the integral (2.7).

If

$$\lambda = |\lambda| \exp(i\phi), \tag{3.1}$$

then

$$|\lambda^{-z}| = |\lambda|^{-x} e^{\phi y}. \tag{3.2}$$

It therefore follows that the integral (2.7) will continue to make sense for the complex λ, (3.3), only if the integrand is $o(e^{-\phi y})$, $y \to \infty$. Since $M[f_1; 1-z]$ decays only algebraically, the burden for this behavior must be placed on $M[h; z]$. We state the following result proved in Bleistein and Handelsman (1975, § 4.7).

THEOREM 2. *Suppose that in the open sector of the complex t-plane defined by* $|t| > 0$, $|\arg t| < \theta_0$, $h(t)$ *is analytic and satisfies* (2.1) *and* (2.2). *Then*

$$M[h; z] = O(\exp[-(\theta - \varepsilon)|y|]), \qquad |y| \to \infty \tag{3.3}$$

for $x > -\operatorname{Re} \gamma$ *and any* $\varepsilon > 0$. *Here*

$$\theta = \min(\theta_0, \pi/2\nu). \tag{3.4}$$

As a consequence of this theorem, we obtain the following.

THEOREM 3. *If* $h(t)$ *satisfies the conditions of Theorem* 2, *then the results* (2.17), (2.18) *or* (2.19), (2.20) *are valid for* $|\arg \lambda| < \theta$, *with* θ *defined by* (3.4).

Proof. In the prescribed sector the integral $I_2(\lambda)$, which is defined by (2.4) and estimated in (2.6), is asymptotically zero with respect to the asymptotic sequences appearing in (2.17) or (2.19). Thus we must consider $I_1(\lambda)$. For any ϕ, choose $\varepsilon = |\theta - \phi|/2$. Then on any vertical contour, the integrand is $O(e^{-\varepsilon|y|})$ as $|y| \to \infty$. Thus this is true for the integrals over the contours P_2 and P_3 in Fig. 1, and (2.16) is true for this complex λ. The asymptotic expansion of the integral in (3.6) is obtained by Watson's lemma and is valid for $|\arg \mu| < \pi/2$. But this introduces no new restriction, since $\arg \mu \to 0$ as $|\lambda| \to \infty$ in the given sector, as does $\arg(\mu + a)$. Watson's lemma applied to (2.16) yields (2.17) or (2.19), depending on the choice of asymptotic sequence, but now for $|\arg \lambda| < \theta$. This completes the proof.

COROLLARY. *For the kernels considered in the previous section, we obtain the following sectors of validity*:

$$\begin{aligned} e^{-t}; &\quad \theta = \pi/2, \\ K_\mu(t); &\quad \theta = \pi/2, \\ \operatorname{Ai}(t); &\quad \theta = \pi/3, \\ D_\mu(t); &\quad \theta = \pi/4. \end{aligned} \tag{3.5}$$

4. Algebraically dominated kernels. We consider (1.1) now with f given by (1.3), and we suppose now that, in addition to (2.2),

$$h(t) \sim \sum_{m=0}^{\infty} \sum_{m=0}^{\tilde{N}(m)} d_{mn} t^{-r_m} (\log t)^n, \qquad t \to \infty. \tag{4.1}$$

Here $\operatorname{Re} r_m \uparrow \infty$, and $\tilde{N}(m)$ is finite for each m. For convergence, it is necessary that

$$\operatorname{Re} \alpha + 1 < \operatorname{Re} r_0. \tag{4.2}$$

This condition and (2.2) assure the convergence of $I(\lambda)$ and further that $M[h; z]$ exists and is analytic in the nonvacuous strip

$$-\operatorname{Re} \gamma < x < \operatorname{Re} r_0.$$

We use the decomposition (2.3), (2.4), (2.5). Then with the aid of the estimate (4.1), we immediately conclude that

$$I_2(\lambda) = O(\lambda^{-r_0+\varepsilon} (\log \lambda)^{\tilde{N}(0)}), \quad \text{any } \varepsilon > 0. \tag{4.3}$$

Furthermore, the integral $I_1(\lambda)$ has the same asymptotic expansion (2.17), except that now the error is $O(\lambda^{-r_0})$. It is shown in Handelsman and Lew (1971) that for $h(t)$ having the expansion (4.1), $M[h; z]$ has a pole of order $\tilde{N}(0)+1$ at r_0, and further, poles of order $\tilde{N}(m)+1$ at r_m. Thus one might proceed to find correction terms as a residue series. However, the actual value of these residues depend upon the nonunique analytic continuation of $M[f_1; 1-z]$; in particular, it depends on whether we extend the logarithmic branch cuts above or below each of these poles.

At first glance, this nonuniqueness might seem to make our results suspect. However, there is really no contradiction, since the consequences of this nonuniqueness are asymptotically zero with respect to the asymptotic sequences of Theorem 1.

The extension of this result to complex λ depends again on the nature of $h(t)$. We state the following corollary to Theorem 2, also proved in Bleistein and Handelsman (1975).

COROLLARY 1. *If, in Theorem* 2, *we replace* (2.1) *by* (4.1), *then* $\theta = \theta_0$.

As an example of this result, we take

$$h(t) = (1+t)^{-1} \tag{4.4}$$

for which $I(\lambda)$ is the Stieltjes transform of $f(t)$ evaluated at λ^{-1} and multiplied by λ^{-1}. Thus we are effectively calculating the Stieltjes transform for small argument. We use the result

$$M[(1+t)^{-1}; z] = \pi \csc \pi z \tag{4.5}$$

to find that for $-1 < \operatorname{Re} \alpha < 0$,

$$\int_0^{\infty} \frac{t^{\alpha}}{a - \log t} \frac{dt}{1+\lambda t} \sim \lambda^{-\alpha-1} \sum_{m=0}^{\infty} \tilde{C}_m (a + \log \lambda)^{-m-1}, \tag{4.6}$$

$$|\arg \lambda| < \pi.$$

Here, we determine the coefficients $\tilde{C}_m$ from

$$\pi \operatorname{cosec} \pi z = \sum_{m=0}^{\infty} \frac{\tilde{C}_m}{m!}(z-\alpha-1)^m, \tag{4.7}$$

and, in particular,

$$\tilde{C}_0 = -\pi \operatorname{cosec} \pi\alpha. \tag{4.8}$$

5. Functions with fractional powers of logarithms. We consider now integrals of the form (1.1) under the assumption that $f(t)$ has an asymptotic expansion of the form (1.5). We define I_1, I_2, f_1 and f_2 as in (2.3)–(2.5). We take h to satisfy (2.2) and (2.1) or (4.1). Then, as above, I_2 is respectively exponentially small or $O(\lambda^{-r_0})$ with $\operatorname{Re} r_0 > \operatorname{Re} \alpha_0 + 1$.

We begin first by assuming that (2.1) is true and thus proceed to study $I_1(\lambda)$ only. The asymptotic analysis of I_1 depends on the proof of the following.

THEOREM 4. *Suppose that $f(t) \equiv 0$ for $t \geqq 1$, locally integrable on $(0, 1)$ and has the asymptotic expansion (1.5) with $N(m)$ finite for each m and $\operatorname{Re} \alpha_m \uparrow \infty$. Then*

(i) *$M[f; 1-z]$ is analytic for $x < \operatorname{Re} \alpha_0 + 1$;*

(ii) *the analytic continuation of $M[f; 1-z]$ to the right takes the form*

$$M[f; 1-z] = \sum_{\operatorname{Re}(\alpha_m - \alpha_0) < k} \left\{ \sum_{n=0}^{N(m)}{}^{*} \frac{c_{mn}\, e^{i\pi\beta_{mn}} \Gamma(\beta_{mn}+1)}{(\alpha_m + 1 - z)^{\beta_{mn}+1}} + \sum_{\beta_{mn} = -l}{}'' \frac{c_{mn}(\alpha_m + 1 - z)^{l-1}}{(l-1)!} \log(z - \alpha_m - 1) \right\} + M_k(z). \tag{5.1}$$

Here, in $\sum^{*}$, we exclude the terms with β_{mn} a negative integer, while, in $\sum''$, we include only terms with β_{mn} a negative integer. The function $M_k(z)$ is analytic for $x < \operatorname{Re} \alpha_0 + k$, and the result is correct for any k.

The proof of this result is given in the Appendix. We see here that negative integer powers, β_{mn}, lead to logarithmic branch points in the Mellin-transform-plane, nonnegative integer powers, β_{mn}, lead to poles, and all other powers lead to algebraic branch points in the transform-plane.

Equation (2.7) is again true for $I_1(\lambda)$ and

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \lambda^{-z} M[h; z] M_k(z)\, dz = O(\lambda^{-(\operatorname{Re}\alpha_0 + k - \varepsilon)}), \tag{5.2}$$

any $\varepsilon > 0$. Thus we need only consider the asymptotic expansion of each term of the sums in (5.1). To this end, we first define

$$J(\alpha, \beta, \lambda) = \frac{e^{i\pi\beta}}{2\pi i} \int_{c-i\infty}^{c+i\infty} \lambda^{-z} (\alpha + 1 - z)^{-\beta-1} M[h; z]\, dz. \tag{5.3}$$

We must separate the analysis here into distinct cases.

Case 1. $\beta = l$, a nonnegative integer. In this case, $J(\alpha, l, \lambda)$ is given as a residue of the integrand at $\alpha + 1$. This is the case treated by Handelsman and Lew

in the papers cited above:

$$J(\alpha, l, \lambda) = \frac{1}{l!}\left(\frac{d}{dz}\right)^l \{\lambda^{-z} M[h; z]\}\Big|_{z=\alpha+1}. \tag{5.4}$$

In particular,

$$J(\alpha, 0, \lambda) = \lambda^{-\alpha-1} M[h; \alpha+1], \tag{5.5}$$

while for larger values of l, the differentiation of λ^{-z} with respect to z produces powers of $\log \lambda$, as well.

Case 2. β is not an integer. Here, we set

$$J(\alpha, \beta, \lambda) = \frac{e^{i\pi\beta}}{2\pi i} \int_{P_1} e^{-\mu z} (\alpha + 1 - z)^{-\beta-1} M[h; z]\, dz + O(\lambda^{-\mathrm{Re}\alpha - k}), \tag{5.6}$$

wth $\mu = \log \lambda$ as above, P_1 as defined in § 2, and k any positive number.

The asymptotic expansion of this integral for large μ can be obtained by a generalization of Watson's lemma to "loop" integrals. This expansion is discussed in Bleistein and Handelsman (1975). The result is

$$J(\alpha, \beta, \lambda) \sim \sum_{j=0}^{\infty} \frac{C_j e^{i\pi\beta} (\log \lambda)^{\beta-j}}{\lambda^{\alpha+1} \Gamma(1+\beta-j)}, \qquad \beta \text{ not an integer.} \tag{5.7}$$

Here, the coefficients C_j are defined by

$$M[h; z] = \sum_{j=0}^{\infty} C_j (z - \alpha - 1)^j, \qquad C_j = \frac{M^{(j)}[h; \alpha+1]}{j!}. \tag{5.8}$$

We must also deal with integrands arising from terms of the second sum in (5.1). Thus we define

$$K(\alpha, l, \lambda) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \lambda^{-z} (\alpha + 1 - z)^{l-1} \log (z - \alpha - 1) M[h; z]\, dz. \tag{5.9}$$

Again, we can deform the contour onto P_1. Furthermore, we explicitly use the jump in the logarithm across the cut to reduce this integral to

$$K(\alpha, l, \lambda) = (-1)^l \lambda^{-\alpha-1} \int_0^k e^{-\mu z} z^{l-1} M[h; z + \alpha + 1]\, dz + O(\lambda^{-\alpha-1-k}), \tag{5.10}$$

with $\mu = \log \lambda$ as above.

The asymptotic expansion of this integral is obtained by Watson's lemma; the result is

$$K(\alpha, l, \lambda) \sim \lambda^{-\alpha-1} \sum_{j=0}^{\infty} C_j \binom{l+j}{j} (\log \lambda)^{-l-j}. \tag{5.11}$$

The coefficients, C_j, are determined by

$$z^{l-1} M[h; z + \alpha + 1] = \sum_{j=0}^{\infty} \frac{C_j}{j!} z^{j+l-1}. \tag{5.12}$$

We use the above defined functions J, (5.3) and K, (5.9), and Theorem 4 to set

$$I_1(\lambda) = \sum_{\mathrm{Re}(\alpha_m-\alpha_0)<k} \left\{ \sum_{n=0}^{N(m)}{}^{*} c_{mn}\Gamma(\beta_{mn}+1)J(\alpha_m, \beta_{mn}, \lambda) + \sum_{\beta_{mn}=-l} \frac{c_{mn}}{(l-1)!} K(\alpha_m, l, \lambda) \right\} + O(\lambda^{-\alpha_0-1-k+\varepsilon}), \tag{5.13}$$

and $\varepsilon > 0$. The results of the discussion above are now summarized in the following.

THEOREM 5. *Let $I(\lambda)$, (1.1) be an absolutely convergent (perhaps improper) integral with f and h locally integrable on $(0, \infty)$ and h satisfying* (2.1), (2.2). *Then, with I_1 and I_2 defined by* (2.3)–(2.5), *I_2 satisfies* (2.6), *and $I_1(\lambda)$ has the asymptotic expansion* (5.13). *The functions appearing in the sums in* (5.13) *have the asymptotic expansions* (5.4) *or* (5.7) *and* (5.11). *The notation $\sum^{*}$ and $\sum''$ is defined below* (5.1).

We remark that, unless β_{0n} are all nonnegative integers, it makes no sense to proceed beyond $m = 0$. This is so, because the alternative leads to a nonterminating expansion in powers of $\log \lambda$ arising from α.

If all β_{0n} are nonnegative integers, we proceed to contributions for $m = 1$. If all β_{1n} are nonnegative integers, we calculate that finite series and proceed to consider contributions from $m = 2$. In this manner, we arrive at the following special case, which is really a result of Handelsman and Lew (1969):

$$I_1(\lambda) = \sum_{\mathrm{Re}(\alpha_m-\alpha_0)<k} \sum_{n=0}^{N(m)} c_{mn}\left(-\frac{d}{dz}\right)^n \{\lambda^{-z} M[h; z]\}\Big|_{z=\alpha_m+1}, \tag{5.14}$$

$\beta_{mn} = n$, all n.

We turn now to the consideration of algebraic kernels, so that (2.1) is to be replaced by (4.1). We then would not carry the sum in (5.13) past λ^{-r_0} because the sum would not take account of poles of $M[h; z]$ at r_0, r_1, etc. Of course, we can still use (5.13) up to $O(\lambda^{-r_0})$; i.e., if $h(t) = O(t^{-r_0})$, $t \to \infty$, then $I(\lambda)$ has the asymptotic expansion (5.13) with

$$k = \mathrm{Re}\,(r_0 - \alpha_0 - 1). \tag{5.15}$$

Hence the series contains contributions from all singularities of $M[f; 1-z]$ to the left of r_0, and the error is $O(\lambda^{-r_0+\varepsilon})$, any $\varepsilon > 0$.

To extend this expansion further, we must know the nature of the analytic continuation of $M[h; z]$ to the right. This result was derived by Handelsman and Lew (1969) and will not be repeated here.

The extension of the results of this section to complex λ are straightforward. The constraint on complex λ depends upon $h(t)$ as in Theorem 2 but also on the nature of $f(t)$. Indeed, further information about $f(t)$ can only improve the results of Theorem 2 (§ 3) and Corollary 1 (§ 4). Here "improve" means extend the sector of validity. The reader is referred to Bleistein and Handelsman (1975) for results of this type.

We shall close this section with an example. We consider the integral

$$I(\lambda) = \int_0^1 |\ln t|^{-1/2} e^{-\lambda t}\, dt. \tag{5.16}$$

Here $f(t)$ is a single term of the form (1.5) with

$$\alpha_0 = 0, \quad \beta_{00} = -\tfrac{1}{2}, \quad c_{00} = e^{i\pi/2}, \tag{5.17}$$

the last being chosen so that

$$c_{00}(\ln t)^{-1/2} = |\ln t|^{-1/2} \tag{5.18}$$

is real and positive for $0 < t < 1$. For this case, $M[f_1; 1-z]$ in (5.1) has no sum $\sum''$ and only a single term in the sum $\sum^*$. Indeed, $M_k(z)$ in (5.1) is zero in this case and

$$M[f_1; 1-z] = \frac{\sqrt{\pi}}{(1-z)^{1/2}}. \tag{5.19}$$

Thus one must consider J in (5.3) with α, β as given in (5.17), h the exponential function and

$$M[h; z] = M[e^{-t}; z] = \Gamma(z). \tag{5.20}$$

Now, for the asymptotic expansion of J, in this case, we obtain (5.7) with the appropriate values substituted. Thus we find

$$\begin{aligned} e^{i\pi/2} J(0, -\tfrac{1}{2}, \lambda) &\sim \sum_{j=0}^{\infty} \frac{C_j (\log \lambda)^{(1/2)-j}}{\lambda \Gamma(\frac{1}{2} - j)}, \\ C_j &= \frac{1}{j!} \left(\frac{d}{dz}\right)^j \Gamma(z) \Big|_{z=1}. \end{aligned} \tag{5.21}$$

We remark that

$$C_0 = \Gamma(1) = 1, \qquad C_1 = \Gamma'(1) = -\gamma = -0.57721 \cdots. \tag{5.22}$$

Here γ is the Euler–Mascheroni constant.

We use this result in (5.13) now to obtain the following one and two term expansions of $I(\lambda)$:

$$\begin{aligned} I(\lambda) &= I_0[1 + O((\ln \lambda)^{-1})], \qquad I_0 = \lambda^{-1} (\ln \lambda)^{-1/2}, \\ I(\lambda) &= I_1[1 + O((\ln \lambda)^{-1})], \qquad I_1 = I_0[1 - \gamma/(2 \ln \lambda)]. \end{aligned} \tag{5.23}$$

In Table 1 we compare I_0 and I_1 with a 10-point Laguerre integration of $I(\lambda)$ for three values of λ, 10, 50, 100. That is, we make λt a new variable of integration and then use the formula

$$\int_0^{\infty} e^{-x} f(x)\, dx \approx \sum_{i=1}^{n} \omega_i f(x_i),$$

with the ω_i and x_i values taken from Abramowitz and Stegun (1965, p. 26). We tabulate $\log \lambda$, as well, because this is the "large" parameter in the asymptotic expansion. Percentage errors are shown in the last two lines of the table. The percentage error for the one-term expansion when $\lambda = 10$ is viewed by this author as surprisingly good, in light of the small value of $\log \lambda$. The large error of the two-term expansion, here, indicates that one term is the optimum number for this case. In contrast, when $\lambda = 50$ or 100, the correction term in I_1 improves the result and the optimum number of terms is at least two.

TABLE 1

λ	10	50	100
$\log\lambda$	2.3025	3.9120	4.6051
$I\times 10^2$	63.76	9.733	4.497
$I_0\times 10^3$	65.90	10.111	4.659
$I_1\times 10^3$	57.64	9.365	4.368
$100(I_0-I)/I$	3.35	3.88	3.61
$100(I-I_1)/I$	9.59	3.78	2.87

Appendix. We shall prove Theorem 4 (§ 5) in this Appendix. Actually, (i) is proved in Titchmarsh (1948). Thus we turn to (ii). We define

$$S_k(t)=S_k^{(1)}(t)+S_k^{(2)}(t)+S_k^{(3)}(t). \tag{A.1}$$

Here, each of these functions satisfies

$$S_k^{(j)}(t)\equiv 0, \qquad t\geqq 1, \tag{A.2}$$

and for $0<t<1$,

$$S_k^{(1)}(t)=\sum_{\mathrm{Re}(\alpha_m-\alpha_0)<k}\ \sum_{\substack{n=0\\ \mathrm{Re}\beta_{mn}>-1}}^{N(m)} c_{mn}(\log t)^{\beta_{mn}}t^{\alpha_m}, \tag{A.3}$$

$$S_k^{(2)}(t)=\sum_{\mathrm{Re}(\alpha_m-\alpha_0)<k}\ \sum_{\mathrm{Re}\beta_{mn}\leqq -1}{}' c_{mn}(\log t)^{\beta_{mn}}t^{\alpha_m}(1-t^k)^L, \tag{A.4}$$

$$S_k^{(3)}(t)=\sum_{\mathrm{Re}(\alpha_m-\alpha_0)<k}\ \sum_{\substack{l=1\\ \beta_{mn}=-l}}^{L}{}'' c_{mn}(\log t)^{\beta_{mn}}t^{\alpha_m}(1-t^k)^L. \tag{A.5}$$

Here L is chosen so large that $L+\mathrm{Re}\,\beta_{mn}>0$ for all β_{mn} appearing in the sums $S_k^{(2)}$ and $S_k^{(3)}$. In $\sum'$, β_{mn} is never a negative integer while in $\sum''$, β_{mn} is always a negative integer.

We remark that $S_k^{(1)}$ contains all terms with powers of $\log t$ integrable at $t=1$. The remaining sums contain powers of $\log t$ nonintegrable at $t=1$, but separated according to integer versus noninteger powers. The choice of L makes $S_k^{(2)}$ and $S_k^{(3)}$ integrable near $t=1$, while the factor $1-t^k$ assures that $f(t)$ and $S_k(t)$ have the same asymptotic expansion to order

$$\mathrm{Re}\,\alpha_0+k-\varepsilon, \quad \text{any } \varepsilon>0.$$

We set

$$f_k(t)=f(t)-S_k(t)=o(t^{\alpha_0+k-\varepsilon}), \quad \text{any } \varepsilon>0. \tag{A.6}$$

Thus the Mellin transform of f_k,

$$M[f_k;1-z]=F_k^{(0)}(z), \tag{A.7}$$

is analytic for

$$x = \operatorname{Re} z < \operatorname{Re} \alpha_0 + k + 1, \tag{A.8}$$

again, following Titchmarsh (1948).

We now introduce

$$M^{(j)}(z) = M[S_k^{(j)}; 1-z], \qquad j = 1, 2, 3. \tag{A.9}$$

For $j = 1$, we consider a single term of the sum on the right, namely,

$$M_{mn}^{(1)}(z) = \int_0^1 (\log t)^{\beta_{mn}} t^{\alpha_m - z}\, dt, \qquad \operatorname{Re} \beta_{mn} > -1, \quad x < \operatorname{Re} \alpha_m + 1. \tag{A.10}$$

With the change of variable of integration,

$$\log t = -\tau = \tau e^{i\pi}, \tag{A.11}$$

the integral is readily recognized as

$$M_{mn}^{(1)}(z) = \frac{e^{i\pi\beta_{mn}} \Gamma(\beta_{mn}+1)}{(\alpha_m + 1 - z)^{\beta_{mn}+1}}. \tag{A.12}$$

Here, for $\alpha_m + 1 - z > 0$, we take its argument $= 0$ and require that the branch cut go to infinity to the right of $\alpha_m + 1$. In our specific calculation, we take the cut horizontal. The analytic continuation of $M_{mn}^{(1)}(z)$ is explicit in (A.12). Since $M^{(1)}(z)$ is a finite sum of such terms, its analytic continuation is explicit and contributes to (5.1) that part of $\sum^*$ for which $\operatorname{Re} \beta_{mn} > -1$. We turn now to $j = 2$ in (A.9) and again consider a single term of the sum of transforms arising from (A.4):

$$M_{mn}^{(2)}(z) = \int_0^1 (\log t)^{\beta_{mn}} t^{\alpha_m - z} (1 - t^k)^L\, dt, \qquad x < \operatorname{Re} \alpha_m + 1. \tag{A.13}$$

We first use the transformation (A.11) to obtain

$$M_{mn}^{(2)}(z) = e^{i\pi\beta_{mn}} \int_0^\infty \tau^{\beta_{mn}} e^{-\tau(\alpha_m + 1 - z)} (1 - e^{-k\tau})^L\, d\tau. \tag{A.14}$$

We can replace the path of integration $(0+)$ which is the contour $-P_1$ of Fig. 1 extended to $+\infty$. The integrand is the same as (A.14) except that the multiplicative factor $\exp[i\pi\beta_{mn}]$ is replaced by $(2i \sin \pi\beta_{mn})^{-1}$. Now that the path of integration does not touch the origin, we can apply the binomial theorem to the last factor ot obtain

$$M_{mn}^{(2)}(z) = \frac{1}{2i \sin \pi\beta_{mn}} \sum_{j=0}^{\infty} \binom{L}{j} (-1)^j \int_{0+} \tau^{\beta_{mn}} e^{-\tau(\alpha_m + 1 + jk - z)}\, d\tau. \tag{A.15}$$

All integrals with $j \geqq 1$ are analytic for $x < \operatorname{Re} \alpha_0 + k + 1$. The integral with $j = 0$ can be calculated from a contour integral definition of $\Gamma(z)$ (see Erdélyi et al. (1953, p. 14, (4)). The result is

$$M_{mn}^{(2)}(z) = \frac{e^{i\pi\beta_{mn}} \Gamma(\beta_{mn}+1)}{(\alpha_m + 1 - z)^{\beta_{mn}+1}} + F_{mn}^{(2)}(z) \tag{A.16}$$

with $F_{mn}^{(2)}$ analytic for $x < \operatorname{Re}\alpha_0 + k + 1$. The comment below (A.12) applies here, as well.

Now, from (A.4), (A.9), (A.13), (A.14), (A.16), we find that $M^{(2)}(z)$ is a triple sum over m, n and j which yields the remainder of $\sum^*$ in (5.1) (from terms with $j = 0$) and a further contribution to $M_k(z)$ (from terms with $j \geqq 1$ or equivalently, from the sums over the functions $F_{mn}^{(2)}(z)$).

We turn now to $j = 3$ in (A.9). To examine a single term in the sum representing $M^{(3)}$, we again must consider (A.13), but with $\beta_{mn} = -l$, a negative integer. Again we use (A.14), but now set

$$M_{mn}^{(2)}(z) = \frac{(-1)}{2\pi i}\int_{0_+} \tau^{\beta_{mn}} \log\tau\, e^{-\tau(\alpha_m+1-z)}(1-e^{-k\tau})^L\, d\tau. \tag{A.17}$$

Now, application of the binomial theorem yields

$$M_{mn}^{(2)}(z) = \frac{1}{2\pi i}\sum (-1)^{l+j}\binom{L}{j}\int_{0_+} \tau^{-l} \log\tau\, e^{-\tau(\alpha_m+1+jk-z)}\, d\tau. \tag{A.18}$$

Again, for $j \geqq 1$, the integral is analytic for $x < \operatorname{Re}\alpha_0 + k + 1$. Thus

$$M_{mn}^{(2)}(z) = \frac{(-1)^l}{2\pi i}\int_{0_+} \tau^{-l} \log\tau\, e^{-\tau(\alpha_m+1-z)}\, d\tau + F^{(3)}(z) \tag{A.19}$$

with $F^{(3)}(z)$ analytic for $x < \operatorname{Re}\alpha_0 + k + 1$. "Stretching" by $\alpha_m + 1 - z$ yields

$$\begin{aligned} M_{mn}^{(2)}(z) = &-(z-\alpha_m-1)^{l-1}\int_{0_+} \tau^{-l}[\log\tau - \log(\alpha_m+1-z)]\, e^{-\tau}\, d\tau \\ &+ F_{mn}^{(3)}(z). \end{aligned} \tag{A.20}$$

The integral with factor $\log\tau$ is analytic for all z since l is a positive integer. Thus

$$M_{mn}^{(2)}(z) = (z-\alpha_m-1)^{l-1}\log(\alpha_m+1-z)\int_{0_+} \tau^{-l} e^{-\tau}\, d\tau + F_{mn}^{(4)}(z), \tag{A.21}$$

$F^{(4)}$ analytic for $x < \operatorname{Re}\alpha_0 + k + 1$. The integral here is simply a residue. Also, we introduce only another analytic function by reversing the sign of the argument of the log. Thus

$$M_{mn}^{(2)}(z) = \frac{(z-\alpha_m-1)^{l-1}}{(l-1)!}\log(z-\alpha_m-1) + F_{mn}^{(5)}(z). \tag{A.22}$$

We now sum these Mellin transforms to obtain $M^{(3)}(z)$ in (A.9), obtaining the coefficients in the sums from (A.5). This sum yields $\sum''$ in (5.1) and an additional contribution to $M_k(z)$.

To recapitulate: $M^{(1)}(z)$ yielded a part of the sum $\sum^*$ in (5.1); $M^{(2)}(z)$ yielded the remainder of $\sum^*$ and a contribution to $M_k(z)$; $M^{(3)}(z)$ yielded all of $\sum''$ and a contribution to $M_k(z)$; $M[f; 1-z] - M^{(1)} - M^{(2)} - M^{(3)}$ contributes only to $M_k(z)$.

This completes the proof.

REFERENCES

M. ABRAMOWITZ AND I. A. STEGUN (1965), *Handbook of Mathematical Functions*, 7th ed., Dover, New York.

N. BLEISTEIN AND R. A. HANDELSMAN (1975), *Asymptotic Expansions of Integrals*, Holt, Rinehart and Winston, New York.

C. J. BOUWKAMP (1971), *Note on an asymptotic expansion*, Indiana Univ. Math. J., 21, pp. 547–549.

A. ERDÉLYI, W. MAGNUS, F. OBERHETTINGER AND F. G. TRICOMI (1953), *Higher Transcendental Functions*, vol. II, McGraw-Hill, New York.

A. ERDÉLYI (1956), *Asymptotic expansions of Fourier integrals involving logarithmic singularities*, SIAM J. Appl. Math., 4, pp. 38–47.

———, (1961), *General asymptotic expansions of Laplace integrals*, Arch. Rational Mech. Anal., 7, no. 1, pp. 1–20.

R. A. HANDELSMAN AND J. S. LEW (1969), *Asymptotic expansion of a class of integral transforms via Mellin transforms*, Ibid., pp. 382–396.

——— (1971), *Asymptotic expansion of integral transforms having algebraically dominated kernels*, J. Math. Anal. Appl., 35, pp. 405–433.

D. S. JONES (1966), *Generalised Functions*, McGraw-Hill, New York.

——— (1969), *Generalised transforms and their asymptotic behavior*, Phil. Trans. Roy. Soc. London Ser. A, 265, no. 1158, pp. 1–43.

J. MCKENNA (1965), *Note on asymptotic expansions of Fourier integrals involving logarithmic singularities*, SIAM J. Appl. Math., 15, pp. 810–812.

W. E. OLMSTEAD AND R. A. HANDELSMAN (1976), *Asymptotic solution to a class of nonlinear Volterra integral equations. II*, Ibid., 30, pp. 180–189.

E. RIEKSTINS (1974), *Asymptotic expansions for some type of integrals involving logarithms*, Latvian Math. Yearbook, 15, pp. 113–130.

E. C. TITCHMARSH (1948), *An Introduction to the Theory of the Fourier Integral*, 2nd ed., Oxford University Press, London.

R. WONG (1970), *On a Laplace integral involving logarithms*, this Journal, 1, pp. 360–364.

R. WONG AND M. WYMAN (1972), *A generalization of Watson's lemma*, Canad. J. Math., 24, no. 2, pp. 185–208.

R. WONG (1975), *On Laplace transforms near the origin*, Math. Comp., 29, pp. 573–576.

SIAM J. MATH. ANAL.
Vol. 8, No. 4, August 1977

CONNECTION FORMULAS FOR SECOND-ORDER DIFFERENTIAL EQUATIONS HAVING AN ARBITRARY NUMBER OF TURNING POINTS OF ARBITRARY MULTIPLICITIES*

F. W. J. OLVER†

Abstract. Consider the differential equation

$$d^2w/dx^2 = \{u^2 f(x) + g(x)\}w, \qquad x \in (a, b),$$

in which (a, b) is a finite or infinite open interval, u is a positive parameter, $f(x)$ is real and twice continuously differentiable, and $g(x)$ is continuous. It is well known that in any subinterval of (a, b) not containing a turning point, that is, a zero of $f(x)$, uniform asymptotic solutions for large u can be constructed in terms of the so-called Liouville–Green or WKBJ functions:

$$f^{-1/4}(x) \exp\left\{\pm u \int f^{1/2}(x)\, dx\right\}.$$

If (a, b) contains turning points, then differing combinations of the Liouville–Green functions have to be used in subintervals that are separated by a turning point in order to represent the same solution.

This paper solves the general problem of connecting the Liouville–Green approximations throughout the interval (a, b) for any number of turning points of arbitrary multiplicities. Several illustrative examples are given, including an arbitrary number of turning points of even multiplicity, an arbitrary number of turning points of odd multiplicity, an eigenvalue problem involving four turning points of multiplicities 1, 2, 3, and 4, and a problem with two simple turning points and one multiple turning point that is solvable exactly in terms of Whittaker functions.

1. Introduction and summary.

In this paper we study the differential equation

$$d^2w/dx^2 = \{u^2 f(x) + g(x)\}w, \tag{1.01}$$

in which the independent variable x ranges over an open, possibly infinite, interval (a, b), and u is a large positive parameter. Our primary assumptions are that within (a, b) the function $f(x)$ is real and twice continuously differentiable, and $g(x)$ is continuous. For simplicity of exposition we shall also suppose that $f(x)$ and $g(x)$ are independent of u, but the analysis that we shall give is extendible in a straightforward manner to a wide variety of cases in which $f(x)$ and $g(x)$ depend on u.[1]

Let $[a_1, b_1]$ be any compact interval lying within (a, b) and not containing any zeros of $f(x)$. Then the theory of the Liouville–Green (LG) approximation, given for example in [24, Chap. 6] shows that when $u \to \infty$ the functions

$$f^{-1/4}(x) \exp\left\{\pm u \int f^{1/2}(x)\, dx\right\} \tag{1.02}$$

furnish asymptotic representations of a pair of linearly independent solutions of (1.01) with relative errors that are uniformly $O(u^{-1})$ in $[a_1, b_1]$. Zeros of $f(x)$ are known as turning points or transition points of equation (1.01), and their exclusion

* Received by the editors December 19, 1975, and in revised form March 15, 1976.

† Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742, and National Bureau of Standards, Washington, D.C. 20234. This research was supported by the U.S. Army Research Office, Durham under Contract DA ARO D 31 124 73 G204, and the National Science Foundation under Grant GP 32841X2.

[1] Compare [24, Chap. 13], and [27].

from the intervals of validity of the LG approximation is essential. This is because for each value of u the solutions of (1.01) are twice continuously differentiable throughout (a, b) and therefore bounded in any interior compact interval, whereas the LG functions (1.02) become infinite at a turning point.

Now suppose that $[a_2, b_2]$ is another compact interval that lies within (a, b) and is free from turning points. Suppose also that $[a_1, b_1]$ and $[a_2, b_2]$ are disjoint. For large u any chosen solution of (1.01) may be represented asymptotically in $[a_1, b_1]$ as a certain linear combination of the LG functions (1.02). It is natural to enquire whether this solution is represented by the same combination in $[a_2, b_2]$. If the interval I, say, that separates $[a_2, b_2]$ from $[a_1, b_1]$ is also free from turning points, then the answer is in the affirmative, because the theory of the LG approximation can be applied to the single interval comprising the union of $[a_1, b_1]$, $[a_2, b_2]$, and I. On the other hand, if I contains one or more turning points, then different combinations of LG functions have to be employed in the intervals $[a_1, b_1]$ and $[a_2, b_2]$, in general, in order to represent the same solution of (1.01). As we shall see in later sections, the changes in these combinations depend on the number of turning points in I and also on their multiplicities, that is, the orders of the zeros of $f(x)$. The general problem of finding the manner in which the combinations of the LG functions change is called the *connection formula problem*, and is of considerable importance in quantum mechanics; see, for example, [1], [5], [7].

In the present paper we apply recent results of the present writer [27] concerning single turning points of arbitrary multiplicity to solve the most general connection formula problem for real variables. That is, we show how to calculate the approximate changes in the combinations of the LG functions in $[a_1, b_1]$ and $[a_2, b_2]$ when the intervening interval I contains any number of turning points, each having arbitrary multiplicity.

The paper is arranged as follows. In § 2 we state theorems that supply rules for connecting the LG approximations across a single turning point of any multiplicity. Four distinct cases arise, depending on the sign of $f(x)$ and the parity of the multiplicity of the turning point. An important feature of the rules is that they are readily combinable with each other, thereby facilitating passage through any number of turning points in succession. The rules are given explicitly only for passage from left to right through each turning point; corresponding rules for passage in the other direction are easily found by changing the sign of x in the original differential equation. The theorems of § 2 are proved in § 3, and then discussed briefly in § 4.

In § 5, which is easily the longest section of the paper, the rules are illustrated by six representative examples. Examples 1 and 2 treat the two cases that involve an arbitrary number of turning points of even multiplicity. In Example 1 the sign of $f(x)$ between consecutive pairs of turning points is positive, causing the solutions of (1.01) to be monotonic; in Example 2 this sign is negative, causing the solutions to be oscillatory. In a similar way, Examples 3 and 4 treat the cases involving an arbitrary number of turning points of odd multiplicity, including the solution of the eigenvalue problem that arises in one of these cases. Example 5 gives the solution of an eigenvalue problem involving four turning points whose multiplicities are, in order, 2, 1, 4, and 3. This example serves to illustrate the use

of all four rules in the same problem, and the results were checked by direct numerical methods. The final illustration, Example 6, is a problem due to Heading [11] that involves three turning points. Two of the turning points are simple; the third has any even multiplicity and is located midway between the other two. This example is solvable exactly in terms of Whittaker functions, thereby furnishing another useful check on the new asymptotic theory.

In the concluding section, § 6, the work of other writers on connection formulas for second-order differential equations having two or more turning points is described and related to the results of the present investigation.

2. Basic rules. In this section we consider the differential equation (1.01) in the case when the independent variable x ranges over a finite or infinite open interval (a_0, b_0). The function $f(x)$ is real, and $g(x)$ is real or complex. Furthermore, we assume that:

(i) *$f(x)/(x-x_0)^l$ is nonvanishing and twice continuously differentiable within (a_0, b_0), where x_0 is an interior point of (a_0, b_0) and l is a nonnegative integer.*

(ii) *$g(x)$ is continuous within (a_0, b_0).*

(iii) *The integral*

$$\int \left\{ \frac{1}{|f(x)|^{1/4}} \frac{d^2}{dx^2}\left(\frac{1}{|f(x)|^{1/4}}\right) - \frac{g(x)}{|f(x)|^{1/2}} \right\} dx \tag{2.01}$$

is absolutely convergent as $x \to a_0+$ or b_0-.

The following points should be noted. First, in consequence of condition (i) $f(x)$ is a twice continuously differentiable function whose only possible zero within (a_0, b_0) is a zero of multiplicity l at x_0. Secondly, the integral (2.01) diverges as $x \to x_0$ (but this is immaterial). Thirdly, the functions $f(x)$ and $g(x)$ are permitted to become infinite as x tends to either endpoint a_0 or b_0, provided that condition (iii) is satisfied. However, if (a_0, b_0) is finite and conditions (i) and (ii) are satisfied in the closure of (a_0, b_0), then at both endpoints $f(x)$ is finite and nonvanishing, $g(x)$ is finite, and condition (iii) is automatically fulfilled.

Let us denote by $\hat{a}_0$ and $\hat{b}_0$ two arbitrary fixed points such that

$$a_0 < \hat{a}_0 < x_0 < \hat{b}_0 < b_0. \tag{2.02}$$

Then in the intervals $(a_0, \hat{a}_0)$ and $(\hat{b}_0, b_0)$ uniform asymptotic solutions of (1.01) for large u can be constructed in terms of LG functions. In order to formulate concise rules for connecting these asymptotic solutions we introduce the following notations.

First, we define

$$\xi \equiv \xi(x) = \int^x |f(t)|^{1/2}\, dt. \tag{2.03}$$

Any convenient value may be assigned to the lower limit of integration, provided that the same choice is adhered to in a given context. In consequence of condition (i), $\xi(x)$ is continuously differentiable in the interval (a_0, b_0). Moreover, because $\xi(x)$ is increasing the relation between ξ and x is one to one.

Secondly, we suppose that the symbols $\theta(u)$, $\phi(u)$, and $\chi(u)$ denote three conveniently chosen real functions of the positive parameter u that are independent of x. We require $\phi(u)$ and $\chi(u)$ to have the following properties as $u \to \infty$:

$$\phi(u) = O(1), \quad \chi(u) \to 0, \quad \frac{1}{\chi(u)} = O(u). \tag{2.04}$$

We also stipulate that $\chi(u)$ has to be positive, and we shall often write χ for $\chi(u)$. An admissible choice, for example, is $\chi = u^{-1/2}$.

Thirdly, we define $\chi_l \equiv \chi_l(u)$ and $\hat{\chi}_l \equiv \hat{\chi}_l(u)$ by the equations

$$\chi_l(u) = \left\{ \begin{array}{ll} u^{-1}, & l = 0, 1 \\ u^{-1} \ln u, & l = 2 \\ u^{-4/(l+2)}, & l \geqq 3 \end{array} \right\}, \qquad \hat{\chi}_l(u) = \max\{\chi(u), \chi_l(u)\}. \tag{2.05}$$

Case I. In this case we assume l to be even and the sign of $f(x)/(x - x_0)^l$ to be positive.

THEOREM 1. *Let $w(u, x)$ be a solution of* (1.01) *having the following properties when u is large and $x \in (a_0, \hat{a}_0)$:*

$$f^{1/4}(x) w(u, x) = \{\phi(u) + O(\chi)\} e^{u\xi(x)}, \tag{2.06}$$

and

$$\frac{d}{dx}\{f^{1/4}(x) w(u, x)\} = \{\phi(u) + O(\chi)\} u f^{1/2}(x) e^{u\xi(x)}, \tag{2.07}$$

the O-terms being uniform with respect to x. Then in the interval $(\hat{b}_0, b_0)$ this solution has the properties

$$f^{1/4}(x) w(u, x) = \csc\left(\frac{\pi}{l+2}\right)\{\phi(u) + O(\hat{\chi}_l)\} e^{u\xi(x)}, \tag{2.08}$$

and

$$\frac{d}{dx}\{f^{1/4}(x) w(u, x)\} = \csc\left(\frac{\pi}{l+2}\right)\{\phi(u) + O(\hat{\chi}_l)\} u f^{1/2}(x) e^{u\xi(x)}, \tag{2.09}$$

as $u \to \infty$, the O-terms again being uniform with respect to x.

The proof of this result, and also of Theorems 2, 3, and 4 stated below, forms the subject of the next section.

It will be observed that because $d\xi/dx = |f(x)|^{1/2}$, equations (2.07) and (2.09) may be obtained by formal differentiation of (2.06) and (2.08) respectively, ignoring the differentiation of the error terms $O(\chi)$ and $O(\hat{\chi}_l)$. This kind of pairing occurs frequently in the rest of this paper, and in order to avoid excessive repetition *we shall use the symbol $\underset{(2)}{=}$ in place of $=$ to signify that a given equation is valid, and also the corresponding equation obtained by formal differentiation with respect to x ignoring the differentiation of all O-terms.* We also adopt the convention that *whenever an O-term appears in an equation it is uniformly valid with respect to all values of x associated with that equation.* With these understandings Theorem 1

is expressed concisely by the following statement:

$$f^{1/4}(x)w(u,x) \underset{(2)}{=} \{\phi(u)+O(\chi)\}\, e^{u\xi(x)}, \qquad x \in (a_0, \hat{a}_0), \tag{2.10}$$

implies

$$f^{1/4}(x)w(u,x) \underset{(2)}{=} \csc\left(\frac{\pi}{l+2}\right)\{\phi(u)+O(\hat{\chi}_l)\}\, e^{u\xi(x)}, \qquad x \in (\hat{b}_0, b_0). \tag{2.11}$$

Case II. Here l is even and the sign of $f(x)/(x-x_0)^l$ is negative.

THEOREM 2.

$$\begin{aligned} |f(x)|^{1/4}w(u,x) \underset{(2)}{=} &\{\phi(u)+O(\chi)\}\cos\{u\xi(x)-u\xi(x_0)+\theta(u)\} \\ &+O(\chi)\sin\{u\xi(x)-u\xi(x_0)+\theta(u)\}, \qquad x \in (a_0, \hat{a}_0), \end{aligned} \tag{2.12}$$

implies

$$\begin{aligned} |f(x)|^{1/4}w(u,x) \underset{(2)}{=} &\{\Lambda(u)+O(\hat{\chi}_l)\}\cos\{u\xi(x)-u\xi(x_0)+\lambda(u)\} \\ &+O(\hat{\chi}_l)\sin\{u\xi(x)-u\xi(x_0)+\lambda(u)\}, \qquad x \in (\hat{b}_0, b_0), \end{aligned} \tag{2.13}$$

where

$$\begin{aligned} \Lambda(u) = \phi(u)\Big[&\cot^2\left(\frac{\pi}{2l+4}\right)\cos^2\left\{\theta(u)+\frac{1}{4}\pi\right\} \\ &+\tan^2\left(\frac{\pi}{2l+4}\right)\sin^2\left\{\theta(u)+\frac{1}{4}\pi\right\}\Big]^{1/2}, \end{aligned} \tag{2.14}$$

and

$$\lambda(u) = -\frac{1}{4}\pi + \tan^{-1}\left[\tan^2\left(\frac{\pi}{2l+4}\right)\tan\left\{\theta(u)+\frac{1}{4}\pi\right\}\right]. \tag{2.15}$$

In the last equation the branch of the inverse tangent is taken to be zero when $\theta(u) = -\frac{1}{4}\pi$, and defined by continuity for other values of $\theta(u)$.

It should be noted that the two terms $O(\chi)$ in (2.12) denote different error terms, in general. The undifferentiated form of (2.12) is equivalent to

$$|f(x)|^{1/4}w(u,x) = \phi(u)\cos\{u\xi(x)-u\xi(x_0)+\theta(u)\}+O(\chi), \qquad x \in (a_0, \hat{a}_0),$$

and the differentiated form is equivalent to

$$\frac{d}{dx}\{|f(x)|^{1/4}w(u,x)\} = -u|f(x)|^{1/2}[\phi(u)\sin\{u\xi(x)-u\xi(x_0)+\theta(u)\}+O(\chi)],$$

$$x \in (a_0, \hat{a}_0).$$

Similarly for (2.13), and also (2.16) and (2.19) below.

Case III. Here l is odd and the sign of $f(x)/(x-x_0)^l$ is positive.

THEOREM 3.

$$|f(x)|^{1/4}w(u,x) \underset{(2)}{=} \{\phi(u)+O(\chi)\}\cos\{u\xi(x)-u\xi(x_0)+\theta(u)\} + O(\chi)\sin\{u\xi(x)-u\xi(x_0)+\theta(u)\}, \qquad x\in(a_0,\hat{a}_0), \tag{2.16}$$

implies

$$f^{1/4}(x)w(u,x) \underset{(2)}{=} \frac{1}{2}\csc\left(\frac{\pi}{2l+4}\right)\left[\phi(u)\cos\left\{\theta(u)+\frac{1}{4}\pi\right\}+O(\hat{\chi}_l)\right] \times e^{u\xi(x)-u\xi(x_0)}, \qquad x\in(\hat{b}_0,b_0). \tag{2.17}$$

Case IV. Here l is odd and the sign of $f(x)/(x-x_0)^l$ is negative.

THEOREM 4.

$$f^{1/4}(x)w(u,x) \underset{(2)}{=} \{\phi(u)+O(\chi)\}\,e^{u\xi(x)}, \qquad x\in(a_0,\hat{a}_0), \tag{2.18}$$

implies

$$|f(x)|^{1/4}w(u,x) \underset{(2)}{=} \csc\left(\frac{\pi}{2l+4}\right)e^{u\xi(x_0)}\left[\{\phi(u)+O(\hat{\chi}_l)\}\cos\left\{u\xi(x)-u\xi(x_0)-\frac{1}{4}\pi\right\} + O(\hat{\chi}_l)\sin\left\{u\xi(x)-u\xi(x_0)-\frac{1}{4}\pi\right\}\right], \qquad x\in(\hat{b}_0,b_0). \tag{2.19}$$

3. Proof of the theorems of § 2. We begin with the proof of Theorem 1. Since $f(x)/(x-x_0)^l$ is positive and l is even, the function $f(x)$ is positive throughout (a_0,b_0), except at x_0. In consequence of conditions (i), (ii), and (iii) stated at the beginning of § 2, there exists a solution $\hat{w}(u,x)$, say, of (1.01) such that for large u

$$f^{1/4}(x)\hat{w}(u,x) \underset{(2)}{=} \{1+O(u^{-1})\}\,e^{u\xi(x)}, \qquad x\in(a_0,\hat{a}_0); \tag{3.01}$$

see [24, p. 203]. If we now define

$$w(u,x)=\{\phi(u)+O(\chi)\}\hat{w}(u,x), \tag{3.02}$$

where the term $O(\chi)$ in this equation is independent of x, then $w(u,x)$ is another solution of (1.01). Hypothesis (2.04) requires $|\phi(u)|$ to be bounded and $O(u^{-1})\subseteq O(\chi)$. Hence from (3.01) and (3.02) it follows that

$$f^{1/4}(x)w(u,x) \underset{(2)}{=} \{\phi(u)+O(\chi)\}\,e^{u\xi(x)}, \qquad x\in(a_0,\hat{a}_0), \tag{3.03}$$

where the term $O(\chi)$ may now depend on x. In other words, with the given conditions a solution $w(u,x)$ having the properties (2.06) and (2.07) certainly exists, but because there is freedom of choice of the term $O(\chi)$ in (3.02) this solution is not unique.

The problem of finding the correct LG approximation to $w(u,x)$ in the interval $(\hat{b}_0,b_0)$ is very similar to the connection formula problem solved in [27,

§ 4.2]. An important difference in the present case, however, is that the endpoints a_0 and b_0 need not be singularities of the differential equation. The effect of this modification is to render the matching of the solutions somewhat more complicated.

Let us define $\zeta \equiv \zeta(x)$ by

$$\zeta(x) = -\left\{\int_x^{x_0} f^{1/2}(t)\,dt\right\}^{2/(l+2)}, \qquad a_0 < x \leqq x_0, \tag{3.04}$$

and

$$\zeta(x) = \left\{\int_{x_0}^{x} f^{1/2}(t)\,dt\right\}^{2/(l+2)}, \qquad x_0 \leqq x < b_0. \tag{3.05}$$

Then in terms of the function $\xi(x)$ defined by (2.03), we have

$$|\zeta(x)|^{(l+2)/2} = |\xi(x) - \xi(x_0)|, \qquad x \in (a_0, b_0). \tag{3.06}$$

From (3.04) and (3.05) it is clear that $\zeta(x)$ is an increasing function of x, and from the lemma of [27, § 3.2] we also know that $\zeta(x)$ is twice continuously differentiable. Furthermore, by taking $m = l+2$ and $\varpi = 0$ in §§ 3.2 and 5.2 of the same reference, we see that (1.01) has twice continuously differentiable solutions $w_-(u, x)$ and $w_+(u, x)$, say, such that for large u and $x \in (a_0, b_0)$

$$\hat{f}^{1/4}(x) w_-(u, x) \underset{(2)}{=} 2^{1/2}(l+2)^{-1/2} u^{l/(2l+4)}\{1 + O(\chi_l)\} U_{l+2}(-u^{2/(l+2)}\zeta), \tag{3.07}$$

and

$$\hat{f}^{1/4}(x) w_+(u, x) \underset{(2)}{=} 2^{1/2}(l+2)^{-1/2} u^{l/(2l+4)}\{1 + O(\chi_l)\} U_{l+2}(u^{2/(l+2)}\zeta). \tag{3.08}$$

In these relations $\hat{f}(x)$ is defined by

$$\hat{f}(x) \equiv \frac{4f(x)}{(l+2)^2|\zeta|^l} = \left(\frac{d\zeta}{dx}\right)^2, \tag{3.09}$$

the normalizing factors $2^{1/2}(l+2)^{-1/2}u^{l/(2l+4)}$ have been introduced for later convenience, the function $\chi_l \equiv \chi_l(u)$ is defined by (2.05), and the function $U_{l+2}(t)$ is defined in terms of modified Bessel functions by the equations

$$U_{l+2}(t) = (2t/\pi)^{1/2} K_{1/(l+2)}(t^{(l+2)/2}), \qquad t > 0,$$

$$U_{l+2}(t) = (2|t|/\pi)^{1/2}[\pi \csc\{\pi/(l+2)\} I_{1/(l+2)}(|t|^{(l+2)/2}) + K_{1/(l+2)}(|t|^{(l+2)/2})], \qquad t < 0.$$

Properties of $U_{l+2}(t)$ are discussed in [27, §2.1].

Suppose now that $x \in (a_0, \hat{a}_0)$. Then from (2.02) and (3.04) it follows that ζ is negative and bounded away from zero. Consequently the functions U_{l+2} in (3.07) and (3.08) may be replaced by their respective asymptotic approximations for large positive and negative arguments, given in [27, § 2.1]; similarly for the derivatives of these functions. Aided by (3.06) and (3.09) and the fact that

$O(u^{-1}) \subseteq O(\chi_l)$, we arrive at

$$f^{1/4}(x)w_-(u,x) \underset{(2)}{=} \{1+O(\chi_l)\}\exp\{u\xi(x)-u\xi(x_0)\}, \qquad x\in(a_0,\hat{a}_0), \tag{3.10}$$

and

$$f^{1/4}(x)w_+(u,x) \underset{(2)}{=} \{1+O(\chi_l)\}\csc\left(\frac{\pi}{l+2}\right)\exp\{u\xi(x_0)-u\xi(x)\}, \qquad x\in(a_0,\hat{a}_0). \tag{3.11}$$

Secondly, suppose that $x\in(\hat{b}_0, b_0)$. Then by similar analysis we find that

$$f^{1/4}(x)w_-(u,x) \underset{(2)}{=} \{1+O(\chi_l)\}\csc\left(\frac{\pi}{l+2}\right)\exp\{u\xi(x)-u\xi(x_0)\}, \qquad x\in(\hat{b}_0,b_0), \tag{3.12}$$

and

$$f^{1/4}(x)w_+(u,x) \underset{(2)}{=} \{1+O(\chi_l)\}\exp\{u\xi(x_0)-u\xi(x)\}, \qquad x\in(\hat{b}_0,b_0). \tag{3.13}$$

Let the linear relation that holds between the wanted solution $w(u,x)$ and the known solutions $w_-(u,x)$ and $w_+(u,x)$ be denoted by

$$w(u,x) = Aw_-(u,x)+Bw_+(u,x), \tag{3.14}$$

where the coefficients A and B are independent of x (but may depend on u). To find A and B, we let $x\to a_0+$ in the equation

$$f^{1/4}(x)w(u,x) = Af^{1/4}(x)w_-(u,x)+Bf^{1/4}(x)w_+(u,x), \tag{3.15}$$

and its differentiated form. Assuming for the moment that $\xi(a_0)$ is finite, and referring to (2.06), (2.07), (3.10), and (3.11), we obtain the equations

$$\{\phi(u)+O(\chi)\}\exp\{u\xi(a_0)\} = A\{1+O(\chi_l)\}\exp\{u\xi(a_0)-u\xi(x_0)\} + B\{1+O(\chi_l)\}\csc\left(\frac{\pi}{l+2}\right)\exp\{u\xi(x_0)-u\xi(a_0)\},$$

and

$$\{\phi(u)+O(\chi)\}\exp\{u\xi(a_0)\} = A\{1+O(\chi_l)\}\exp\{u\xi(a_0)-u\xi(x_0)\} - B\{1+O(\chi_l)\}\csc\left(\frac{\pi}{l+2}\right)\exp\{u\xi(x_0)-u\xi(a_0)\}.$$

Solving for A and B, remembering that $|\phi(u)|$ is assumed to be bounded, we derive

$$A = \exp\{u\xi(x_0)\}\{\phi(u)+O(\hat{\chi}_l)\}, \tag{3.16}$$

$$B = \exp\{2u\xi(a_0)-u\xi(x_0)\}O(\hat{\chi}_l), \tag{3.17}$$

where $\hat{\chi}_l$ is defined by (2.05).

Alternatively, if $\xi(a_0) = -\infty$ then by similar analysis we find that A is again given by (3.16) and B vanishes. In other words, it is legitimate to replace the right-hand side of (3.17) by its limiting value as $\xi(a_0) \to -\infty$.

Uniform asymptotic approximations for $f^{1/4}(x)w(u, x)$ in the interval $(\hat{b}_0, b_0)$ may now be found from (3.15) by substituting (3.12), (3.13), (3.16), and (3.17). When $\xi(a_0)$ is finite it is easily seen that the whole of the contribution from the second solution on the right-hand side of (3.15) is absorbable in the uniform error term associated with the contribution from the first solution. Alternatively, when $\xi(a_0) = -\infty$ there is no contribution from the second solution. Similarly for the derivatives. The final results are given by (2.08) and (2.09) or, equivalently, by (2.11). This completes the proof.

To prove Theorem 2, we begin with the results of §§ 3.3 and 5.2 of [27]. Using methods similar to those just employed for Theorem 1, we find that there exist solutions $w_-(u, x)$ and $w_+(u, x)$, say, of equation (1.01) with the properties

$$|f(x)|^{1/4}w_-(u, x) \underset{(2)}{=} \{1 + O(\chi_l)\} \sec\left(\frac{\pi}{2l+4}\right) \cos\left\{u\xi(x) - u\xi(x_0) - \frac{1}{4}\pi\right\} + O(\chi_l) \sin\left\{u\xi(x) - u\xi(x_0) - \frac{1}{4}\pi\right\}, \qquad x \in (a_0, \hat{a}_0), \tag{3.18}$$

$$|f(x)|^{1/4}w_+(u, x) \underset{(2)}{=} \{1 + O(\chi_l)\} \csc\left(\frac{\pi}{2l+4}\right) \cos\left\{u\xi(x) - u\xi(x_0) + \frac{1}{4}\pi\right\} + O(\chi_l) \sin\left\{u\xi(x) - u\xi(x_0) + \frac{1}{4}\pi\right\}, \qquad x \in (a_0, \hat{a}_0), \tag{3.19}$$

and

$$|f(x)|^{1/4}w_-(u, x) \underset{(2)}{=} \{1 + O(\chi_l)\} \csc\left(\frac{\pi}{2l+4}\right) \cos\left\{u\xi(x) - u\xi(x_0) - \frac{1}{4}\pi\right\} + O(\chi_l) \sin\left\{u\xi(x) - u\xi(x_0) - \frac{1}{4}\pi\right\}, \qquad x \in (\hat{b}_0, b_0), \tag{3.20}$$

$$|f(x)|^{1/4}w_+(u, x) \underset{(2)}{=} \{1 + O(\chi_l)\} \sec\left(\frac{\pi}{2l+4}\right) \cos\left\{u\xi(x) - u\xi(x_0) + \frac{1}{4}\pi\right\} + O(\chi_l) \sin\left\{u\xi(x) - u\xi(x_0) + \frac{1}{4}\pi\right\}, \qquad x \in (\hat{b}_0, b_0). \tag{3.21}$$

The wanted solution $w(u, x)$ has the property (2.12), and is easily seen to exist by the theory of the Liouville–Green approximation. From (3.18), (3.19), and the fact that $|\phi(u)|$ is assumed to be bounded, we find that the linear relation holding between the three solutions is representable in the form

$$|f(x)|^{1/4}w(u, x) = \left[\phi(u) \cos\left\{\theta(u) + \frac{1}{4}\pi\right\} + O(\hat{\chi}_l)\right] \cos\left(\frac{\pi}{2l+4}\right) |f(x)|^{1/4}w_-(u, x) + \left[\phi(u) \sin\left\{\theta(u) + \frac{1}{4}\pi\right\} + O(\hat{\chi}_l)\right] \sin\left(\frac{\pi}{2l+4}\right) |f(x)|^{1/4}w_+(u, x);$$

compare (3.15). The required results (2.13), (2.14), and (2.15) follow on substituting in the right-hand side of this equation, and its differentiated form, by means of (3.20) and (3.21).

The proofs of Theorems 3 and 4 are similar to those of Theorems 1 and 2 and it is unnecessary to record details.

4. Remarks on the theorems of § 2. Theorems 1, 2, 3, and 4 describe the approximate changes in the form of the LG approximations to the solutions of the differential equation (1.01) on passing through the various types of turning point. For example, in Case I the net effect is to multiply the asymptotic forms (2.06) and (2.07) by the factor $\csc\{\pi/(l+2)\}$. In particular, when $l=0$, that is, when there is no turning point, there is no change in the asymptotic forms—as is to be expected.

In comparing the four theorems, we note that there are differences in the number of solutions of the differential equation that are supplied. We discuss here briefly the underlying reasons, taking each case in turn.

Case I. The effect of changing the lower limit in the definition (2.03) of $\xi(x)$ is to multiply the solution $w(u,x)$ by a factor that is independent of x. In consequence, Theorem 1 supplies direct information for only one independent solution of the differential equation. However, the theory of the LG approximation shows that for large u there exists a second solution $w_2(u,x)$, say, having the property

$$f^{1/4}(x)w_2(u,x) \underset{(2)}{\equiv} \{\phi(u)+O(\chi)\}\,e^{-u\xi(x)}, \qquad x\in(a_0,\hat{a}_0);$$

compare (2.10). Like $w(u,x)$ this solution is expressible in the form

$$w_2(u,x)=Aw_-(u,x)+Bw_+(u,x), \tag{4.01}$$

where $w_-(u,x)$ and $w_+(u,x)$ are as in § 3, but A and B have new values. Carrying through the analysis in a similar manner to § 3, we find that

$$A=\exp\{u\xi(x_0)-2u\xi(a_0)\}O(\hat{\chi}_l),$$

$$B=\sin\left(\frac{\pi}{l+2}\right)\exp\{-u\xi(x_0)\}\{\phi(u)+O(\hat{\chi}_l)\},$$

provided that $\xi(a_0)$ is finite. However, on substituting these results in (4.01) we obtain little information concerning $f^{1/4}(x)w_2(u,x)$ and its derivative in $(\hat{b}_0,{}^{\circ}b_0)$, because in this interval $w_-(u,x)$ dominates $w_+(u,x)$ when u is large (compare (3.12) and (3.13)), and the coefficient A of $w_-(u,x)$ is available only as an O-estimate.

Noting that we are able to obtain a satisfactory connection formula for the solution that is growing in magnitude as we pass through the turning point, but not for the solution that is decaying, we perceive that the way to obtain a second connection formula is to reverse the roles of the original solutions by passing through the turning point in the opposite direction, that is, from right to left. Thus the required formula is found by replacing x by $-x$ in (1.01), applying Theorem 1, and then changing the sign of x again. The result is expressible in the form:

$$f^{1/4}(x)w(u,x) \underset{(2)}{\equiv} \{\phi(u)+O(\chi)\}\,e^{-u\xi(x)}, \qquad x\in(\hat{b}_0,b_0),$$

implies

$$f^{1/4}(x)w(u,x) \underset{(2)}{=} \csc\left(\frac{\pi}{l+2}\right)\{\phi(u)+O(\hat{\chi}_l)\}\, e^{-u\xi(x)}, \qquad x\in(a_0,\hat{a}_0).$$

Case II. Since there is no restriction on the choice of the function $\theta(u)$, and the only restriction on $\phi(u)$ is that this function is bounded in absolute value, Theorem 2 informs us directly how any solution of the differential equation continues through the turning point. In particular, if we select $\theta(u)=-\frac{1}{4}\pi$ and $\frac{1}{4}\pi$, then we obtain a pair of solutions that are asymptotically out of phase by $\frac{1}{2}\pi$, both for $x\in(a_0,\hat{a}_0)$ and $x\in(\hat{b}_0,b_0)$; compare (2.15). They therefore comprise a numerically satisfactory pair of solutions in the sense of Miller [19].

The reason for the difference between Theorems 1 and 2 concerning the number of connection formulas that each theorem supplies is traceable to the fact that in Case II no solution grows exponentially in magnitude compared with other solutions as we pass through the turning point.

Case III. Since $\theta(u)$ and $\phi(u)$ again may be freely chosen, subject to $|\phi(u)|$ being bounded, the situation resembles Case II superficially. However, if $\theta(u)-\frac{1}{4}\pi$ is zero or an integer multiple of π, then the factor contained in the square brackets in (2.17) reduces to the error term $O(\hat{\chi}_l)$. Moreover, if we exclude these values of $\theta(u)$, then it is impossible to construct a pair of solutions that are numerically satisfactory in the interval $(\hat{b}_0,b_0)$. To obtain a satisfactory companion connection formula we have to pass through the turning point in the opposite direction, in the manner of Case I. This is effected by reversing the sign of x in the given differential equation and applying Theorem 4.

Case IV. The situation here is analogous to Case I. Theorem 4 supplies direct information for only one independent solution of the differential equation. To obtain a second connection formula we pass through the turning point from right to left with the aid of Theorem 3.

5. Examples. In this section we consider the differential equation (1.01), that is,

$$d^2w/dx^2=\{u^2f(x)+g(x)\}w, \qquad x\in(a,b), \tag{5.01}$$

in various cases for which there is more than one turning point in the finite or infinite open interval (a,b).

As before, we suppose that within (a,b) the function $f(x)$ is real, $f''(x)$ is continuous, $g(x)$ is real or complex and continuous, and the integral (2.01) is absolutely convergent as x approaches the endpoints a and b. We again define $\xi(x)$ by the integral (2.03), with the understanding that we adhere to the same lower limit of integration in each example.

Example 1. *Arbitrary number of even turning points: first case.* As our first example we suppose that there are n turning points x_r, $r=1,2,\cdots,n$, arranged as follows:

$$a<x_1<x_2<\cdots<x_n<b. \tag{5.02}$$

We also suppose that the function

$$(x_1-x)^{-2l_1}(x_2-x)^{-2l_2}\cdots(x_n-x)^{-2l_n}f(x) \tag{5.03}$$

is positive and twice continuously differentiable within (a, b), where $l_1, l_2, \cdots, l_n$ are positive integers. Thus the multiplicity of the turning point x_r is $2l_r$, and $f(x)$ is positive within (a, b) except at the turning points.

We first introduce a set of arbitrary fixed points $a_r, b_r, r = 1, 2, \cdots, n$, that satisfy

$$a < a_1 < x_1 < b_1 < a_2 < x_2 < b_2 < \cdots < a_n < x_n < b_n < b. \tag{5.04}$$

With the given conditions, we know from the theory of the LG approximation that there is a solution of (5.01) with the property

$$f^{1/4}(x)w(u, x) \underset{(2)}{=} \{1 + O(u^{-1})\}\, e^{u\xi(x)}, \qquad x \in (a, a_1), \tag{5.05}$$

as $u \to \infty$. Applying Theorem 1 of § 2 with $x_0 = x_1$, $l = 2l_1$, $\phi(u) = 1$, and $\chi(u) = 1/u$, we find that

$$f^{1/4}(x)w(u, x) \underset{(2)}{=} \csc\left(\frac{\pi}{2l_1 + 2}\right)\{1 + O(\chi_{2l_1})\}\, e^{u\xi(x)}, \qquad x \in (b_1, a_2).$$

Theorem 1 may now be applied a second time to continue $w(u, x)$ across the turning point x_2 to the interval (b_2, a_3), and so on. The final result is evidently given by

$$f^{1/4}(x)w(u, x) \underset{(2)}{=} \left\{\prod_{r=1}^{n} \csc\left(\frac{\pi}{2l_r + 2}\right)\right\}\{1 + O(\chi_{2l})\}\, e^{u\xi(x)}, \qquad x \in (b_n, b), \tag{5.06}$$

where

$$l = \max(l_1, l_2, \cdots, l_n). \tag{5.07}$$

Remarks. It is interesting to observe that the result (5.06) is independent of the actual location of the turning points in (a, b) and also of the order in which the differing multiplicities occur. It needs to be stressed, however, that we have supposed that the turning points are fixed. In some problems the function $f(x)$ and the turning points may depend continuously on a second parameter ρ, say. Assuming that the given conditions are satisfied uniformly with respect to ρ, we may easily verify that the result (5.06) holds uniformly with respect to ρ, provided that none of the turning points coalesce with each other, or with the endpoints, as ρ varies. If, however, the turning points x_r and x_{r+1}, say, were to coalesce for a certain value of ρ, then in (5.06) the factors

$$\csc\left(\frac{\pi}{2l_r + 2}\right) \csc\left(\frac{\pi}{2l_{r+1} + 2}\right) \tag{5.08}$$

would have to be replaced by the single factor

$$\csc\left(\frac{\pi}{2l_r + 2l_{r+1} + 2}\right). \tag{5.09}$$

Unless the expressions (5.08) and (5.09) are equal—as they are when $l_r = l_{r+1} = 1$—such an abrupt change cannot possibly be uniform with respect to ρ.

Example 2. *Arbitrary number of even turning points: second case.* We assume the same conditions and notation as in Example 1, except that we now suppose

that the sign of the function (5.03) is negative. Thus $f(x)$ is negative, except at the turning points.

From the theory of the LG approximation, there exist solutions $w_{-1}(u, x)$ and $w_1(u, x)$ such that when u is large and $x \in (a, a_1)$

$$|f(x)|^{1/4} w_{-1}(u, x) \underset{(2)}{=} \{1 + O(u^{-1})\} \cos\{u\xi(x) - \tfrac{1}{4}\pi\} + O(u^{-1}) \sin\{u\xi(x) - \tfrac{1}{4}\pi\}, \tag{5.10}$$

and

$$|f(x)|^{1/4} w_1(u, x) \underset{(2)}{=} \{1 + O(u^{-1})\} \cos\{u\xi(x) + \tfrac{1}{4}\pi\} + O(u^{-1}) \sin\{u\xi(x) + \tfrac{1}{4}\pi\}. \tag{5.11}$$

The asymptotic forms of these solutions in the interval (b_1, a_2) can be found by direct application of Theorem 2 of § 2 with $x_0 = x_1$. Subsequent analysis is considerably simplified, however, by taking $\theta(u) = \mp\frac{1}{4}\pi$ in Theorem 2, for from (2.14) and (2.15) we then obtain

$$\Lambda(u) = \phi(u) \cot\left(\frac{\pi}{2l+4}\right), \qquad \lambda(u) = -\tfrac{1}{4}\pi, \quad \text{when } \theta(u) = -\tfrac{1}{4}\pi;$$

or

$$\Lambda(u) = \phi(u) \tan\left(\frac{\pi}{2l+4}\right), \qquad \lambda(u) = \tfrac{1}{4}\pi, \quad \text{when } \theta(u) = \tfrac{1}{4}\pi.$$

Accordingly, we first convert the arguments of the trigonometric functions in (5.10) and (5.11) into the forms $u\xi(x) - u\xi(x_1) - \frac{1}{4}\pi$ and $u\xi(x) - u\xi(x_1) + \frac{1}{4}\pi$, respectively, by setting

$$u\xi(x) \mp \tfrac{1}{4}\pi = u\xi(x_1) + \{u\xi(x) - u\xi(x_1) \mp \tfrac{1}{4}\pi\},$$

and expanding by means of the addition rules. The result is expressed most concisely in matrix form. Let $\mathbf{R}_r$ denote the rotation matrix

$$\mathbf{R}_r = \begin{bmatrix} \cos u\xi(x_r) & \sin u\xi(x_r) \\ -\sin u\xi(x_r) & \cos u\xi(x_r) \end{bmatrix}, \qquad r = 1, 2, \cdots, n.$$

Then we have

$$\begin{bmatrix} |f(x)|^{1/4} w_{-1}(u, x) \\ |f(x)|^{1/4} w_1(u, x) \end{bmatrix} \underset{(2)}{=} \mathbf{R}_1 \begin{bmatrix} \{1 + O(u^{-1})\} \cos\{u\xi(x) - u\xi(x_1) - \frac{1}{4}\pi\} \\ \quad + O(u^{-1}) \sin\{u\xi(x) - u\xi(x_1) - \frac{1}{4}\pi\} \\ \{1 + O(u^{-1})\} \cos\{u\xi(x) - u\xi(x_1) + \frac{1}{4}\pi\} \\ \quad + O(u^{-1}) \sin\{u\xi(x) - u\xi(x_1) + \frac{1}{4}\pi\} \end{bmatrix},$$

valid when $x \in (a, a_1)$.

We now apply Theorem 2 with $x_0 = x_1$, $l = 2l_1$, $\phi(u) = 1$, $\chi(u) = 1/u$, and $\theta(u) = -\frac{1}{4}\pi$ and $\frac{1}{4}\pi$ in turn. The result is expressible in the form

$$\begin{bmatrix} |f(x)|^{1/4} w_{-1}(u,x) \\ |f(x)|^{1/4} w_1(u,x) \end{bmatrix}$$

$$\underset{(2)}{=} \mathbf{R}_1 \mathbf{D}_1 \begin{bmatrix} \{1+O(\chi_{2l_1})\}\cos\{u\xi(x)-u\xi(x_1)-\frac{1}{4}\pi\} \\ \quad +O(\chi_{2l_1})\sin\{u\xi(x)-u\xi(x_1)-\frac{1}{4}\pi\} \\ \{1+O(\chi_{2l_1})\}\cos\{u\xi(x)-u\xi(x_1)+\frac{1}{4}\pi\} \\ \quad +O(\chi_{2l_1})\sin\{u\xi(x)-u\xi(x_1)+\frac{1}{4}\pi\} \end{bmatrix}, \tag{5.12}$$

valid when $x \in (b_1, a_2)$, where

$$\mathbf{D}_r = \begin{bmatrix} \cot\{\pi/(4l_r+4)\} & 0 \\ 0 & \tan\{\pi/(4l_r+4)\} \end{bmatrix}, \qquad r = 1, 2, \cdots, n.$$

To prepare for passage through the next turning point x_2 we convert the arguments of the cosine and sine functions appearing in (5.12) into the forms $u\xi(x)-u\xi(x_2)-\frac{1}{4}\pi$ and $u\xi(x)-u\xi(x_2)+\frac{1}{4}\pi$, respectively, by application of the addition formulas. Thus we have

$$\begin{bmatrix} \{1+O(\chi_{2l_1})\}\cos\{u\xi(x)-u\xi(x_1)-\frac{1}{4}\pi\}+O(\chi_{2l_1})\sin\{u\xi(x)-u\xi(x_1)-\frac{1}{4}\pi\} \\ \{1+O(\chi_{2l_1})\}\cos\{u\xi(x)-u\xi(x_1)+\frac{1}{4}\pi\}+O(\chi_{2l_1})\sin\{u\xi(x)-u\xi(x_1)+\frac{1}{4}\pi\} \end{bmatrix}$$

$$\underset{(2)}{=} \begin{bmatrix} \cos\{u\xi(x_2)-u\xi(x_1)\} & \sin\{u\xi(x_2)-u\xi(x_1)\} \\ -\sin\{u\xi(x_2)-u\xi(x_1)\} & \cos\{u\xi(x_2)-u\xi(x_1)\} \end{bmatrix}$$

$$\times \begin{bmatrix} \{1+O(\chi_{2l_1})\}\cos\{u\xi(x)-u\xi(x_2)-\frac{1}{4}\pi\}+O(\chi_{2l_1})\sin\{u\xi(x)-u\xi(x_2)-\frac{1}{4}\pi\} \\ \{1+O(\chi_{2l_1})\}\cos\{u\xi(x)-u\xi(x_2)+\frac{1}{4}\pi\}+O(\chi_{2l_1})\sin\{u\xi(x)-u\xi(x_2)+\frac{1}{4}\pi\} \end{bmatrix}$$

$$\underset{(2)}{=} \mathbf{R}_1^{-1}\mathbf{R}_2 \begin{bmatrix} \{1+O(\chi_{2l_1})\}\cos\{u\xi(x)-u\xi(x_2)-\frac{1}{4}\pi\} \\ \quad +O(\chi_{2l_1})\sin\{u\xi(x)-u\xi(x_2)-\frac{1}{4}\pi\} \\ \{1+O(\chi_{2l_1})\}\cos\{u\xi(x)-u\xi(x_2)+\frac{1}{4}\pi\} \\ \quad +O(\chi_{2l_1})\sin\{u\xi(x)-u\xi(x_2)+\frac{1}{4}\pi\} \end{bmatrix},$$

since

$$\mathbf{R}_1^{-1} = \begin{bmatrix} \cos u\xi(x_1) & -\sin u\xi(x_1) \\ \sin u\xi(x_1) & \cos u\xi(x_1) \end{bmatrix}.$$

We may now apply Theorem 2 with $x_0 = x_2$, $l = 2l_2$, $\phi(u) = 1$, $\chi(u) = \chi_{2l_1}(u)$, and $\theta(u) = \mp\frac{1}{4}\pi$, to continue the solutions to the interval (b_2, a_3). The process may then be repeated for successive turning points, and the final result is easily seen to be

$$\begin{bmatrix} |f(x)|^{1/4} w_{-1}(u,x) \\ |f(x)|^{1/4} w_1(u,x) \end{bmatrix}$$

$$\underset{(2)}{=} \mathbf{M}_1\mathbf{M}_2\cdots\mathbf{M}_n \begin{bmatrix} \{1+O(\chi_{2l})\}\cos\{u\xi(x)-\frac{1}{4}\pi\}+O(\chi_{2l})\sin\{u\xi(x)-\frac{1}{4}\pi\} \\ \{1+O(\chi_{2l})\}\cos\{u\xi(x)+\frac{1}{4}\pi\}+O(\chi_{2l})\sin\{u\xi(x)+\frac{1}{4}\pi\} \end{bmatrix}, \tag{5.13}$$

valid when $x \in (b_n, b)$, where l is defined by (5.07), and

$$\mathbf{M}_r = \mathbf{R}_r \mathbf{D}_r \mathbf{R}_r^{-1}, \qquad r = 1, 2, \cdots, n.$$

Equation (5.13) is the required connection formula.

Remarks. It will be observed that each matrix $\mathbf{M}_r$ depends on x_r as well as l_r; thus $\mathbf{M}_r \equiv \mathbf{M}(x_r, l_r)$. In contrast to Example 1, a change in the value of x_r alters the final result (5.13). Furthermore, it is readily verified that

$$\mathbf{M}(x_r, l_r)\mathbf{M}(x_{r+1}, l_{r+1}) \neq \mathbf{M}(x_r, l_{r+1})\mathbf{M}(x_{r+1}, l_r),$$

in general. Hence the final result is also affected by the order in which differing multiplicities occur, again in contrast to Example 1.

Example 3. *Arbitrary number of odd turning points: first case.* In this example we suppose that there are n turning points x_r, $r = 1, 2, \cdots, n$, again enumerated to satisfy (5.02), and that the multiplicity of the turning point at x_r is $2l_r + 1$, where l_r now denotes a nonnegative integer. We also assume that the function

$$(x_1 - x)^{-2l_1 - 1}(x_2 - x)^{-2l_2 - 1} \cdots (x_n - x)^{-2l_n - 1} f(x) \tag{5.14}$$

is negative and twice continuously differentiable within (a, b).

As in Examples 1 and 2, we begin by introducing arbitrary fixed points a_r, b_r, $r = 1, 2, \cdots, n$, that satisfy (5.04).

When $x \in (a, a_1)$ the function $f(x)$ is negative, and with the assumed conditions the theory of the LG approximation shows that there exists a solution $w(u, x)$ having the property

$$\begin{aligned} |f(x)|^{1/4} w(u, x) &\underset{(2)}{=} \{1 + O(u^{-1})\} \cos\{u\xi(x) - u\xi(x_1) + \theta(u) - \tfrac{1}{4}\pi\} \\ &\quad + O(u^{-1}) \sin\{u\xi(x) - u\xi(x_1) + \theta(u) - \tfrac{1}{4}\pi\}, \qquad x \in (a, a_1), \end{aligned} \tag{5.15}$$

where, as before, $\theta(u)$ is any prescribed real function of the large positive parameter u.

Applying Theorem 3 of § 2 with $x_0 = x_1$, $l = 2l_1 + 1$, $\phi(u) = 1$, $\chi(u) = 1/u$, and $\theta(u)$ replaced by $\theta(u) - \frac{1}{4}\pi$, we see that

$$f^{1/4}(x) w(u, x) \underset{(2)}{=} \frac{1}{2} \csc\left(\frac{\pi}{4l_1 + 6}\right)\{\cos\theta(u) + O(\chi_{2l_1+1})\}\, e^{u\xi(x) - u\xi(x_1)}, \qquad x \in (b_1, a_2).$$

Then applying Theorem 4 of § 2 with $x_0 = x_2$, $l = 2l_2 + 1$, $\phi(u) = \cos\theta(u)$, and $\chi(u) = \chi_{2l_1+1}(u)$, we obtain

$$\begin{aligned} |f(x)|^{1/4} w(u, x) &\underset{(2)}{=} \frac{1}{2} \csc\left(\frac{\pi}{4l_1 + 6}\right) \csc\left(\frac{\pi}{4l_2 + 6}\right) e^{u\xi(x_2) - u\xi(x_1)} \\ &\quad \times [\{\cos\theta(u) + O(\chi_{l_{1,2}})\} \cos\{u\xi(x) - u\xi(x_2) - \tfrac{1}{4}\pi\} \\ &\quad + O(\chi_{l_{1,2}}) \sin\{u\xi(x) - u\xi(x_2) - \tfrac{1}{4}\pi\}], \end{aligned}$$

valid when $x \in (b_2, a_3)$, where $l_{1,2} = 2 \max(l_1, l_2) + 1$.

The last formula describes the net effect of passing through the turning points x_1 and x_2, and the process is repeatable in a straightforward manner any number of times. If the total number of turning points n is even, then the final result is given by

$$
\begin{aligned}
|f(x)|^{1/4}w(u,x) \underset{(2)}{=} \frac{1}{2^{n/2}}\left\{\prod_{r=1}^{n}\csc\left(\frac{\pi}{4l_r+6}\right)\right\}\exp\left\{u\sum_{r=1}^{n}(-)^r\xi(x_r)\right\} \\
\times[\{M(u)+O(\chi_{2l+1})\}\cos\{u\xi(x)-u\xi(x_n)-\tfrac{1}{4}\pi\} \\
+O(\chi_{2l+1})\sin\{u\xi(x)-u\xi(x_n)-\tfrac{1}{4}\pi\}],
\end{aligned}
\tag{5.16}
$$

valid when $x\in(b_n, b)$, where l is defined by (5.07), and[2]

$$
M(u)=\cos\theta(u)\prod_{r=1}^{[(n-1)/2]}\cos\{u\xi(x_{2r+1})-u\xi(x_{2r})\}. \tag{5.17}
$$

On the other hand, if n is odd, then with the same definitions of l and $M(u)$, we have

$$
\begin{aligned}
f^{1/4}(x)w(u,x) \underset{(2)}{=} \frac{1}{2^{(n+1)/2}}\left\{\prod_{r=1}^{n}\csc\left(\frac{\pi}{4l_r+6}\right)\right\}\exp\left\{u\sum_{r=1}^{n}(-)^r\xi(x_r)\right\} \\
\times\{M(u)+O(\chi_{2l+1})\}\,e^{u\xi(x)}, \qquad x\in(b_n, b).
\end{aligned}
\tag{5.18}
$$

Example 4. *Arbitrary number of odd turning points: second case.* We make the same assumptions as in Example 3, except that the sign of the function (5.14) is now assumed to be positive. Using the same notation, we know that there exists a solution $w(u, x)$ having the property

$$
f^{1/4}(x)w(u,x) \underset{(2)}{=} \{1+O(u^{-1})\}\,e^{u\xi(x)}, \qquad x\in(a, a_1). \tag{5.19}
$$

The continuation of this solution to the interval (b_n, b) may be achieved by successive application of Theorems 4 and 3, in a similar manner to Example 3. The final result is readily verified to be

$$
\begin{aligned}
f^{1/4}(x)w(u,x) \underset{(2)}{=} \frac{1}{2^{n/2}}\left\{\prod_{r=1}^{n}\csc\left(\frac{\pi}{4l_r+6}\right)\right\}\exp\left\{u\sum_{r=1}^{n}(-)^{r-1}\xi(x_r)\right\} \\
\times\{N(u)+O(\chi_{2l+1})\}\,e^{u\xi(x)}, \qquad x\in(b_n, b),
\end{aligned}
\tag{5.20}
$$

when n is even, or

$$
\begin{aligned}
|f(x)|^{1/4}w(u,x) \underset{(2)}{=} \frac{1}{2^{(n-1)/2}}\left\{\prod_{r=1}^{n}\csc\left(\frac{\pi}{4l_r+6}\right)\right\}\exp\left\{u\sum_{r=1}^{n}(-)^{r-1}\xi(x_r)\right\} \\
\times[\{N(u)+O(\chi_{2l+1})\}\cos\{u\xi(x)-u\xi(x_n)-\tfrac{1}{4}\pi\} \\
+O(\chi_{2l+1})\sin\{u\xi(x)-u\xi(x_n)-\tfrac{1}{4}\pi\}], \quad x\in(b_n, b),
\end{aligned}
\tag{5.21}
$$

[2] As usual, $[(n-1)/2]$ denotes the integer part of $(n-1)/2$. When n is even this is, of course $(n/2)-1$ but in the following equation we use the same definition of $M(u)$ when n is odd.

when n is odd. Here l is again defined by (5.07), and

$$N(u)=\prod_{r=1}^{[n/2]}\cos\{u\xi(x_{2r})-u\xi(x_{2r-1})\}. \tag{5.22}$$

A subsidiary problem of physical interest [1, Chap. 13], is the following eigenvalue problem. The number of turning points n is even, $\xi(x)\to-\infty$ as $x\to a+$, $\xi(x)\to+\infty$ as $x\to b-$, and we seek those values of u for which there is a solution that is recessive at both endpoints. From (5.20) it follows that a necessary and sufficient condition for the existence of such a solution is given by[3]

$$N(u)+O(\chi_{2l+1})=0. \tag{5.23}$$

Using (2.03) and (5.22) we derive

$$u=\left\{\int_{x_{2r-1}}^{x_{2r}}|f(x)|^{1/2}\,dx\right\}^{-1}(k+\tfrac{1}{2})\pi+O\{\chi_{2l+1}(k)\}, \tag{5.24}$$

where k is a large positive integer, r has any of the values $1, 2, \cdots, \frac{1}{2}n$, l is defined by (5.07), and $\chi_l(u)$ is defined by (2.05).

It may be noted in passing that as an immediate consequence of Theorem 1, the result (5.24) continues to hold when any (finite) number of even turning points are interposed in the intervals (a, x_1), (x_2, x_3), (x_4, x_5), $\cdots$, (x_{n-2}, x_{n-1}), (x_n, b), provided that $2l+1$ now denotes the highest multiplicity of all the turning points in (a, b).

Example 5. *An eigenvalue problem with four mixed turning points.* In this example, we seek the large eigenvalues of the system

$$d^2w/dx^2=u^2h(x)(x+c_1)^2x(x-c_2)^4(x-c_3)^3w, \qquad -\infty<x<\infty, \tag{5.25}$$

where c_1, c_2, and c_3 are positive constants such that $c_2<c_3$, and $h(x)$ is a positive, twice continuously differentiable function of x with the properties

$$h(x)\sim(\text{constant})x^{\gamma_1}, \quad x\to-\infty; \quad h(x)\sim(\text{constant})x^{\gamma_2}, \quad x\to+\infty, \tag{5.26}$$

where γ_1 and γ_2 are constants such that $\gamma_1>-12$ and $\gamma_2>-12$. We also assume that the relations (5.26) are twice differentiable.[4] This problem has been selected because all the theorems of § 2 are used in its solution.

---

[3] This may be proved as follows. With the given conditions we know from §§ 2, 3, and 5 of [24, Chap. 6], that (5.01) has solutions of the form $f^{-1/4}(x)\,e^{u\xi(x)}\{1+\varepsilon_1(u,x)\}$ and $f^{-1/4}(x)\,e^{-u\xi(x)}\times\{1+\varepsilon_2(u,x)\}$, such that as $x\to b-$, $\varepsilon_2(u,x)$ vanishes and $\varepsilon_1(u,x)$ tends to a constant value $\varepsilon_1(u,b)$, say. Furthermore $\varepsilon_1(u,b)$ is $O(u^{-1})$ as $u\to\infty$. Denote the error term $O(\chi_{2l+1})$ in (5.20) by $\eta(u,x)$. Then

$$e^{u\xi(x)}\{N(u)+\eta(u,x)\}=A(u)\,e^{u\xi(x)}\{1+\varepsilon_1(u,x)\}+B(u)\,e^{-u\xi(x)}\{1+\varepsilon_2(u,x)\},$$

where $A(u)$ and $B(u)$ are independent of x. Dividing throughout by $e^{u\xi(x)}$ and letting $x\to b-$, we see that $\eta(u,x)$ tends to a constant $\eta(u,b)$, say, and also that

$$A(u)=\{N(u)+\eta(u,b)\}/\{1+\varepsilon_1(u,b)\}=N(u)+O(\chi_{2l+1}).$$

The condition that the solution (5.20) be recessive at b is $A(u)=0$; which yields (5.23).

[4] The interpretation of this condition in the cases in which γ_1 or γ_2 is 0 or 1 is the same as in § 4.2 of [24, Chap. 6].

We begin by fixing the arbitrary constant in the definition (2.03) of $\xi(x)$ by the condition $\xi(0)=0$; thus

$$\xi(x)=\int_0^x \{h(t)\}^{1/2}|(t+c_1)t^{1/2}(t-c_2)^2(t-c_3)^{3/2}|\,dt. \tag{5.27}$$

There are two reasons for introducing the conditions (5.26). First they ensure that $\xi(x)\to-\infty$ as $x\to-\infty$ and $\xi(x)\to+\infty$ as $x\to+\infty$. Secondly, with

$$f(x)=h(x)(x+c_1)^2x(x-c_2)^4(x-c_3)^3, \qquad g(x)=0,$$

they ensure that the integral (2.01) converges absolutely as $x\to\pm\infty$.

Let a be any fixed point such that $a<-c_1$. Then the theory of the LG approximation shows that (5.25) has a solution $w(u,x)$ with the property

$$f^{1/4}(x)w(u,x)\underset{(2)}{=}\{1+O(u^{-1})\}\,e^{u\xi(x)}, \qquad x\in(-\infty,a).$$

Furthermore, this solution is recessive as $x\to-\infty$.

Applying Theorem 1, with $x_0=-c_1$, $l=2$, $\phi(u)=1$, and $\chi(u)=1/u$, we find that in any fixed closed interval within $(-c_1,0)$, we have

$$f^{1/4}(x)w(u,x)\underset{(2)}{=}\csc(\tfrac{1}{4}\pi)\{1+O(u^{-1}\ln u)\}\,e^{u\xi(x)}.$$

The next step is to apply Theorem 4 with $x_0=0$, $l=1$, $\phi(u)=1$, and $\chi(u)=u^{-1}\ln u$. Since $\xi(0)=0$, we find that in any fixed closed interval within $(0,c_2)$

$$|f(x)|^{1/4}w(u,x)\underset{(2)}{=}\csc(\tfrac{1}{4}\pi)\csc(\tfrac{1}{6}\pi)[\{1+O(u^{-1}\ln u)\}\cos\{u\xi(x)-\tfrac{1}{4}\pi\} +O(u^{-1}\ln u)\sin\{u\xi(x)-\tfrac{1}{4}\pi\}].$$

Thirdly, we apply Theorem 2 with $x_0=c_2$ and $l=4$ to obtain

$$|f(x)|^{1/4}w(u,x)\underset{(2)}{=}\csc(\tfrac{1}{4}\pi)\csc(\tfrac{1}{6}\pi)[\{\Lambda(u)+O(u^{-2/3})\}\cos\{u\xi(x)-u\xi(c_2)+\lambda(u)\} +O(u^{-2/3})\sin\{u\xi(x)-u\xi(c_2)+\lambda(u)\}],$$

valid in any fixed closed interval within (c_2,c_3), where $\Lambda(u)$ and $\lambda(u)$ are given by (2.14) and (2.15) with $\phi(u)=1$ and

$$\theta(u)=u\xi(c_2)-\tfrac{1}{4}\pi. \tag{5.28}$$

The final continuation is to apply Theorem 3 with $x_0=c_3$, $l=3$, $\phi(u)=\Lambda(u)$, and $\chi(u)=u^{-2/3}$, to obtain

$$f^{1/4}(x)w(u,x)\underset{(2)}{=}\tfrac{1}{2}\csc(\tfrac{1}{4}\pi)\csc(\tfrac{1}{6}\pi)\csc(\tfrac{1}{10}\pi)\,e^{-u\xi(c_3)} \times[\Lambda(u)\cos\{u\xi(c_3)-u\xi(c_2)+\lambda(u)+\tfrac{1}{4}\pi\}+O(u^{-2/3})]\,e^{u\xi(x)},$$

valid when $x\in(b,\infty)$, where b is any fixed point such that $b>c_3$.

Since $1/\Lambda(u)$ is bounded it follows, as in Example 4, that the eigenvalues are given by

$$\cos\{u\xi(c_3)-u\xi(c_2)+\lambda(u)+\tfrac{1}{4}\pi\}=O(u^{-2/3}),$$

that is,

$$\{\xi(c_3)-\xi(c_2)\}u+\lambda(u)+\tfrac{1}{4}\pi=(k+\tfrac{1}{2})\pi+O(u^{-2/3}),$$

where k is an integer. On using (2.15), with $l=4$, and (5.28), the last equation becomes

$$\{\xi(c_3)-\xi(c_2)\}u+\tan^{-1}[\tan^2(\tfrac{1}{12}\pi)\tan\{\xi(c_2)u\}]=(k+\tfrac{1}{2})\pi+O(u^{-2/3}). \tag{5.29}$$

To solve (5.29) we first discuss the properties of the function $\omega(u)$ defined by

$$\tan^{-1}\{\tan^2(\tfrac{1}{12}\pi)\tan u\}=u-\omega(u). \tag{5.30}$$

Consistent with § 2, we are using the continuous branch of the left-hand side that vanishes at $u=0$. The following properties are easily verified:

$$\omega(0)=\omega(\tfrac{1}{2}\pi)=\omega(\pi)=0, \qquad \omega(u)=\omega(u+\pi)=-\omega(-u).$$

The graph of $\omega(u)$ is indicated in Figure 5.1 for the interval $(0, \pi)$ and continues by periodicity. By differentiating (5.30) we find that

$$\omega'(u)=1-\frac{\tan^2(\tfrac{1}{12}\pi)}{\cos^2 u+\tan^4(\tfrac{1}{12}\pi)\sin^2 u}.$$

Therefore for all u

$$-\{\cot^2(\tfrac{1}{12}\pi)-1\}\leqq\omega'(u)\leqq 1-\tan^2(\tfrac{1}{12}\pi).$$

Returning to (5.29) and substituting by means of (5.30), we obtain

$$\xi(c_3)u-\omega\{\xi(c_2)u\}=(k+\tfrac{1}{2})\pi+O(u^{-2/3}). \tag{5.31}$$

Now consider the equation

$$\xi(c_3)u-\omega\{\xi(c_2)u\}=(k+\tfrac{1}{2})\pi. \tag{5.32}$$

The u-derivative of the left-hand side is

$$\xi(c_3)-\xi(c_2)\omega'\{\xi(c_2)u\}. \tag{5.33}$$

Since $\xi(c_3)>\xi(c_2)>0$ and the (algebraically) largest value of $\omega'\{\xi(c_2)u\}$ is $1-\tan^2(\frac{1}{12}\pi)$, we conclude that (5.33) is always positive and bounded away from zero. Accordingly, for each value of k equation (5.32) has exactly one root $u=u_k$,

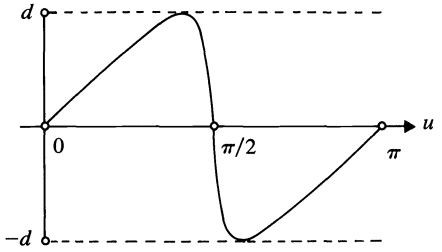


FIG. 5.1. *Graph of* $\omega(u)$. $d=1.047\cdots$

say. Hence from (5.31) and the mean-value theorem we conclude that the required eigenvalues are given by

$$u = u_k + O(k^{-2/3}), \qquad k \to +\infty. \tag{5.34}$$

Remarks. (i) The need to solve a transcendental equation, namely (5.32), in order to obtain adequate approximations to the eigenvalues stems from the presence of a turning point of Case II type. Suppose that instead of (5.25) we had chosen the system

$$d^2w/dx^2 = u^2 h(x)(x+c_1)^2 x(x-c_3)^3 w, \qquad -\infty < x < \infty,$$

with the conditions (5.26) modified by requiring $\gamma_1 > -8$ and $\gamma_2 > -8$. Then by similar analysis we find that the eigenvalues are explicitly given by

$$u = \{\xi(c_3)\}^{-1}(k+\tfrac{1}{2})\pi + O(k^{-4/5}),$$

where k again denotes a large positive integer, and $\xi(x)$ is now defined by

$$\xi(x) = \int_0^x \{h(t)\}^{1/2} |(t+c_1)t^{1/2}(t-c_3)^{3/2}|\, dt.$$

(ii) As in Example 4, the presence of the turning point of Case I type at $x = -c_1$ essentially has no effect on the final result (5.34). Indeed, this turning point may be replaced by an arbitrary finite number of even turning points in the intervals $(-\infty, 0)$ and (c_3, ∞), provided that none is of multiplicity exceeding 4. If the last condition is not fulfilled, then the only change needed is to increase the estimate for the error term in (5.34) from $O(k^{-2/3})$ to $O(k^{-4/(l+2)})$, where l is the multiplicity of the turning point of highest order.

(iii) It is of interest to compare the results yielded by the asymptotic formula (5.34) with those calculated by direct numerical methods, especially as an error bound has not been constructed for the asymptotic estimate. For this purpose we select the following equation:

$$d^2w/dx^2 = u^2(x+1)^2 x(x-1)^4(x-2)^3 w. \tag{5.35}$$

Thus in the notation of (5.25) we have $h(x) = 1$, $c_1 = c_2 = 1$, and $c_3 = 2$. From (5.27) we find that[5]

$$\xi(1) = \tfrac{3}{32}\pi + \tfrac{2}{15}, \qquad \xi(2) = \tfrac{3}{16}\pi. \tag{5.36}$$

Substituting these values in (5.30) and (5.32) we see that the equation for the estimated eigenvalues u_k is given by

$$(\tfrac{3}{32}\pi - \tfrac{2}{15})u + \tan^{-1}[\tan^2(\tfrac{1}{12}\pi)\tan\{(\tfrac{3}{32}\pi + \tfrac{2}{15})u\}] = (k+\tfrac{1}{2})\pi, \tag{5.37}$$

where, as noted previously, the branch of the inverse tangent vanishes at $u = 0$ and is a continuous function of u.

Corresponding to any prescribed value of k, the desired approximation $u = u_k$ to an eigenvalue of (5.35) is obtained by solving (5.37) by successive

[5] These values were verified by use of the MACSYMA System of the Massachusetts Institute of Technology.

approximation. Taking $k = 5$, 6, 7, and 8, we find that

$$u_5 = 29.257, \quad u_6 = 33.262, \quad u_7 = 40.347, \quad u_8 = 47.111, \tag{5.38}$$

correct to three places of decimals.

To calculate the eigenvalues by numerical analysis the following method was used. Equation (5.35) is integrated numerically from a sufficiently large negative value $x = X_1$, say, to $x = 0$, using arbitrary initial values of w and dw/dx at $x = X_1$. The same equation is also integrated numerically from a sufficiently large positive value $x = X_2$, say, to $x = 0$ using arbitrary initial values at $x = X_2$. The value of u is then adjusted until the Wronskian of the two solutions vanishes at $x = 0$. (The actual values of X_1 and X_2 are not crucial, nor are the initial values at these points; for the present range of values of u and required accuracy it was found that $X_1 = -1.1$ and $X_2 = 2.6$ are adequately large.) This method was used to search the interval $28 \leqq u \leqq 48$ systematically for eigenvalues, and the only ones were found to be

$$u = 28.822, \quad 33.185, \quad 40.265, \quad 46.909, \tag{5.39}$$

again, correct to three places of decimals. The agreement with the asymptotic estimates (5.38) is clearly quite satisfactory.

Example 6. *A problem with three turning points solvable in terms of Whittaker functions.* Our final example is furnished by the equation

$$d^2w/dx^2 = u^2x^{2p-2}(c^{2p} - x^{2p})w, \tag{5.40}$$

in which u is a positive parameter, p is a positive integer, and c is a fixed positive constant. On the real axis there are simple turning points at $x = \pm c$, and a turning point of multiplicity $2p-2$ at $x = 0$. The theory of the LG approximation shows that there are unique solutions $w_1(u, x)$, $w_2(u, x)$, and $w_3(u, x)$ with the properties

$$|f(x)|^{1/4} w_1(u, x) \sim e^{iu\xi(x)}, \qquad x \to -\infty, \tag{5.41}$$

$$|f(x)|^{1/4} w_2(u, x) \sim e^{iu\xi(x)}, \qquad x \to +\infty, \tag{5.42}$$

$$|f(x)|^{1/4} w_3(u, x) \sim e^{-iu\xi(x)}, \qquad x \to +\infty, \tag{5.43}$$

where

$$f(x) = x^{2p-2}(c^{2p} - x^{2p}), \tag{5.44}$$

and

$$\xi(x) = \int_0^x |t^{p-1}(c^{2p} - t^{2p})^{1/2}| \, dt. \tag{5.45}$$

Our intention is to find the coefficients A and B in the connection formula

$$w_1(u, x) = Aw_2(u, x) + Bw_3(u, x). \tag{5.46}$$

This problem is of interest in the theory of barrier penetration. In this context the solutions $w_1(u, x)$, $w_2(u, x)$, and $w_3(u, x)$ represent respectively the transmitted, incident, and reflected waves. Heading [11] has solved the problem exactly by transforming (5.40) into Whittaker's equation. A valuable check on the analysis of the present paper is to apply the rules of § 2 to solve the problem approximately

for large u, and then compare the answers with the asymptotic forms of the Whittaker functions.

The theory of the LG approximation shows that corresponding to any assigned values of real constants a and θ, with $a > c$, there is a solution $w(\theta, u, x)$ of (5.40) with the properties

$$|f(x)|^{1/4} w(\theta, u, x) = \cos\{u\xi(x) + \theta\} + o(1), \qquad x \to -\infty, \tag{5.47}$$

when u is fixed, and

$$|f(x)|^{1/4} w(\theta, u, x) \underset{(2)}{=} \{1 + O(u^{-1})\} \cos\{u\xi(x) + \theta\} + O(u^{-1}) \sin\{u\xi(x) + \theta\}, \qquad -\infty < x < -a, \tag{5.48}$$

uniformly when u is large. Applying Theorems 3, 1, and 4 in succession in the manner of previous examples, we find that

$$\begin{aligned} &\sin\left(\frac{\pi}{2p}\right) e^{u\xi(-c) - u\xi(c)} |f(x)|^{1/4} w(\theta, u, x) \\ &\quad \underset{(2)}{=} [2\cos\{u\xi(-c) + \theta + \tfrac{1}{4}\pi\} + O(\chi_{2p-2})] \cos\{u\xi(x) - u\xi(c) - \tfrac{1}{4}\pi\} \\ &\qquad + O(\chi_{2p-2}) \sin\{u\xi(x) - u\xi(c) - \tfrac{1}{4}\pi\}, \end{aligned}$$

valid when $x \in (a, \infty)$. From (5.45) it is clear that $\xi(x)$ is an odd function of x; hence by rearrangement of the last relation we deduce that

$$\begin{aligned} |f(x)|^{1/4} w(\theta, u, x) = \csc\left(\frac{\pi}{2p}\right) e^{2u\xi(c)} &[\cos\{u\xi(x) - 2u\xi(c) + \theta\} \\ &+ \sin\{u\xi(x) - \theta\} + O(\chi_{2p-2})]. \end{aligned} \tag{5.49}$$

From (5.47) it follows that the solution $w_1(u, x)$ having the property (5.41) is given by the linear combination

$$w_1(u, x) = w(0, u, x) + iw(-\tfrac{1}{2}\pi, u, x).$$

Hence from (5.49) we derive

$$|f(x)|^{1/4} w_1(u, x) = \csc\left(\frac{\pi}{2p}\right) e^{2u\xi(c)} \{e^{iu\xi(x) - 2iu\xi(c)} + i\, e^{-iu\xi(x)} + O(\chi_{2p-2})\},$$

valid when $x \in (a, \infty)$. The values of the coefficients A and B in (5.46) may now be found by letting $x \to \infty$, first through the sequence of values for which $2u\xi(x)$ is an even multiple of π, and secondly through the sequence for which $2u\xi(x)$ is an odd multiple of π. Comparing the resulting expressions with (5.42) and (5.43) we see that

$$A = \csc\left(\frac{\pi}{2p}\right) e^{2(1-i)u\xi(c)} \{1 + O(\chi_{2p-2})\}, \tag{5.50}$$

$$B = i \csc\left(\frac{\pi}{2p}\right) e^{2u\xi(c)} \{1 + O(\chi_{2p-2})\}. \tag{5.51}$$

These are the required results. The necessary value of $\xi(c)$ may be found from (5.45) with the aid of the substitution $t^p = c^p \sin v$; thus

$$\xi(c) = \pi c^{2p}/(4p). \tag{5.52}$$

Let us turn now to the exact solution of (5.40) in terms of Whittaker functions. Following Heading [11] we find that when $x<0$ a solution is $\hat{w}(u,x)$, given by

$$\hat{w}(u,x)=|x|^{(1/2)-p}W_{\kappa,1/(4p)}(iu|x|^{2p}/p), \tag{5.53}$$

where

$$\kappa=iuc^{2p}/(4p), \tag{5.54}$$

and $W_{\kappa,\mu}(z)$ is Whittaker's function in the usual notation. For large $|z|$ it is known that[6]

$$W_{\kappa,\mu}(z)\sim z^{\kappa}e^{-z/2}, \qquad |\mathrm{ph}\, z|\leqq\tfrac{3}{2}\pi-\delta(<\tfrac{3}{2}\pi). \tag{5.55}$$

Hence

$$\hat{w}(u,x)\sim|x|^{(1/2)-p}\left(\frac{iu|x|^{2p}}{p}\right)^{\kappa}\exp\left(-\frac{iu|x|^{2p}}{2p}\right), \qquad x\to-\infty. \tag{5.56}$$

In order to compare the last result with (5.41), we first evaluate the integral (5.45) when $x\geqq c$. This may be done by means of the substitution $t^p=c^p\cosh v$; thus

$$\xi(x)=\xi(c)+\frac{c^{2p}}{2p}\left\{\frac{x^p}{c^p}\left(\frac{x^{2p}}{c^{2p}}-1\right)^{1/2}-\cosh^{-1}\left(\frac{x^p}{c^p}\right)\right\}, \qquad x\geqq c.$$

Hence

$$\xi(x)=\xi(c)+\frac{c^{2p}}{2p}\left\{\frac{x^{2p}}{c^{2p}}-\frac{1}{2}-\ln\left(\frac{2x^p}{c^p}\right)\right\}+o(1), \qquad x\to+\infty. \tag{5.57}$$

Using this result and the fact that $\xi(-x)=-\xi(x)$, we see from (5.41), (5.54), and (5.56) that the relation between $w_1(u,x)$ and $\hat{w}(u,x)$ is given by

$$w_1(u,x)=e^{\kappa}\kappa^{-\kappa}e^{-iu\xi(c)}\hat{w}(u,x). \tag{5.58}$$

Next, we need the continuation of $\hat{w}(u,x)$ to positive values of x. If we pass from the negative real x-axis to the positive real x-axis via an indentation that lies in the upper half of the x-plane, then from (5.53) we see that

$$\hat{w}(u,x)=(e^{-\pi i}x)^{(1/2)-p}W_{\kappa,1/(4p)}(e^{-2p\pi i}iux^{2p}/p), \qquad x>0.$$

Applying the connection formula for Whittaker functions with arguments $e^{-2p\pi i}z$, z, and $ze^{-\pi i}$ obtainable, for example, from [11] or [24, p. 262], we derive

$$\hat{w}(u,x)=\csc\left(\frac{\pi}{2p}\right)x^{(1/2)-p}\left\{\hat{B}W_{\kappa,1/(4p)}\left(\frac{iux^{2p}}{p}\right)+\hat{A}W_{-\kappa,1/(4p)}\left(-\frac{iux^{2p}}{p}\right)\right\}, \tag{5.59}$$

where

$$\hat{A}=2\pi e^{-\kappa\pi i}\Big/\left\{\Gamma\left(\frac{1}{2}+\frac{1}{4p}-\kappa\right)\Gamma\left(\frac{1}{2}-\frac{1}{4p}-\kappa\right)\right\}, \tag{5.60}$$

$$\hat{B}=i\left\{e^{-2\kappa\pi i}+\cos\left(\frac{\pi}{2p}\right)\right\}. \tag{5.61}$$

[6] See, for example, [24, p. 260].

The asymptotic form of $w_1(u, x)$ as $x \to +\infty$ may be found from (5.58) and (5.59) by replacing the Whittaker functions by their asymptotic forms for large argument, again given by (5.55). Then by using (5.44), (5.54), and (5.57) we are able to cast the result into the form

$$|f(x)|^{1/4} w_1(u, x) = \csc\left(\frac{\pi}{2p}\right)\{e^{-2iu\xi(c)}\, e^{\kappa\pi i}\, e^{2\kappa}\kappa^{-2\kappa}\hat{A}\, e^{iu\xi(x)} + \hat{B}\, e^{-iu\xi(x)} + o(1)\}.$$

Comparing this result with (5.46) and referring to (5.42) and (5.43), we see that

$$A = \csc\left(\frac{\pi}{2p}\right) e^{-2iu\xi(c)}\, e^{\kappa\pi i}\, e^{2\kappa}\kappa^{-2\kappa}\hat{A}, \tag{5.62}$$

$$B = \csc\left(\frac{\pi}{2p}\right)\hat{B}. \tag{5.63}$$

The last two equations give exact expressions for the coefficients in the wanted connection formula (5.46). To check the approximate formulas (5.50) and (5.51) found by our present theory, we assume u (and therefore, also, $|\kappa|$) is large and calculate the asymptotic form of the right-hand side of (5.60) by means of Stirling's formula. On substituting the result in (5.62) and using (5.52) and (5.54), we find that (5.50) agrees with (5.62) within the tolerance of the uniform error term $O(\chi_{2p-2})$; similarly for (5.51) and (5.63). This confirms the soundness of our asymptotic theory as applied to this example.

6. Previous results and conclusions. It will be assumed in this section that the reader is acquainted with the survey that was made in [27, § 6] of work on connection formulas for second-order differential equations having a single turning point. We shall continue to use the terms *central connection*, *lateral connection*, and *pseudo-lateral connection* in the same sense as in this reference, and remind the reader that central connection and pseudo-lateral connection may be used for turning points of any multiplicity, whereas at present true lateral connection may be employed only for simple turning points. The methods developed in the present paper are all of central connection type.

In the first part of this section we survey applications of the methods of central connection, lateral connection, and pseudo-lateral connection to equations having more than one turning point, proceeding roughly in increasing order of complexity. Because all three methods proceed step by step through one turning point at a time, they may be classified as having a *local* nature, even though the regions of validity of the corresponding asymptotic solutions of the differential equation may extend to infinity. Later in this section we discuss a significantly different type of method in which the approach is of *global* nature.

Two simple turning points. As in Examples 3 and 4 of § 5, two distinct cases arise depending whether the function $f(x)$ in (1.01) is positive or negative in the interval between its zeros. In the latter case there is also a subsidiary eigenvalue problem. Physical models corresponding to the two cases are the potential barrier and the potential well, respectively.

Solutions date back almost 50 years, and references can be found, for example, in [7, Chap. 1], [22], and [23]. From the mathematical standpoint, the

first fully satisfactory applications of the central connection method were those of Jeffreys [13], [14].

Rigorous solutions by means of lateral connection were published independently (and in the same year) by Fedoryuk [3], Fröman and Fröman [5], and the present writer [23]; the last two references also include error bounds.

Three simple turning points. Special examples have been solved by Nishimoto [21, § 6] and the present writer [24, p. 516, Exercises 15.4 and 15.5], in both cases by lateral connection.

Four simple turning points. This situation arises with double potential barriers or wells. As in the case of two simple turning points there are two cases, depending on the sign of $f(x)$. A formal solution of the barrier problem was given by Bohm [1, pp. 283–295], using central connection. Rigorous solutions by lateral connection, both for barriers and wells, have been supplied in a series of papers by N. Fröman and her collaborators [4], [6], [32]. Other work on these problems has been discussed by Heading [9], [10].

Arbitrary number of simple turning points. Problems of this type have been treated by Murphy and Good [20] (central connection), Fedoryuk [3] (lateral connection), Evgrafov and Fedoryuk [2] (lateral connection), and Heading [9] (lateral connection). Some of the analysis in these references is formal or given incompletely. Sibuya [29] and Weinberg [31] have given completely rigorous analyses, using lateral connection, for the case in which $g(x) = 0$ and $f(x)$ is a polynomial, or has polynomial growth rate as $|x| \to \infty$.

Two turning points, at least one of which is multiple. Heading [11, § 8] has solved the general problem in which one of the turning points is simple and the other is of any odd multiplicity by lateral connection, generalizing a device used by the present writer when both turning points are simple [23]. More recently, Heading [12] has made a further extension to solve the general problem in which the two turning points are of arbitrary odd multiplicities.

Leung [15], [16] has treated the example

$$d^2w/dx^2 = u^2x^2(x-1)^2w$$

by rigorous application of pseudo-lateral connection although the solution in these references is incomplete.[7]

Arbitrary number of multiple turning points. Again using pseudo-lateral connection, Leung [17] has solved the eigenvalue problem in the general case in which $g(x) \equiv 0$, $f(x)$ is positive as $x \to \pm\infty$, and there are any number of multiple turning points, provided that none of the turning points is of type II (defined in § 2). This problem was treated in Example 4 of § 5; compare the comments in the closing paragraph in this example. Leung's formula for the eigenvalues agrees with (5.24), except that the order of his estimate for the error term is weaker than our $O\{\chi_{2l+1}(k)\}$.

Global central connection. Methods of this type have been proposed by several writers, including Heading [8], Pike [28], and Lynn and Keller [18]. The idea is to try to approximate, in a uniform manner, the solutions of the given

[7] The complete solution has been communicated by Dr. Leung to the present writer. It agrees with the specialization of Example 1 of § 5 above.

differential equation by those of an equation of the form

$$d^2w/dx^2 = u^2P(x)w, \tag{6.01}$$

in which $P(x)$ is a polynomial whose degree is the sum of the order of the turning points in the interval (or complex domain) under consideration.

When we are dealing with problems in which the turning points are fixed this approach has no advantages to offer compared with the methods of the present paper. Indeed, the reverse is true; little is known about the solutions of equations of the form (6.01) even for quite moderate values of the degree of $P(x)$; compare [26, § 6.3], and [30]. Clearly unless connection formulas are independently available for (6.01) there is little to be gained by trying to use the solutions of this equation to approximate those of the original differential equation.

As noted by Lynn and Keller on page 393 of [18], however, if the locations of the turning points are variable owing to the presence of a second parameter in the differential equation, then there is an important potential advantage in the global approach: it may yield connection formulas that are uniform with respect to the second parameter, *even when two or more of the turning points coalesce.* As we noted in the Remarks at the close of Example 1 in § 5, the approach used in the present paper fails in these circumstances.

At present, the only rigorous development of the global approach that permits coalescence of turning points appears to be that of the present writer [25]. This reference treats the case in which $P(x)$ is a quadratic polynomial; the solutions of (6.01) are then expressible as parabolic cylinder functions. Approximate connection formulas are derived for large u, and they continue to be uniformly valid when the two simple turning points in the original differential equation coalesce into a double turning point. These results could be combined with those of the present paper to treat problems in which there are movable turning points, provided that the only kind of coalescence that occurs is of pairs of simple turning points.

Conclusions. The present paper provides a rigorous solution, and in fact the first general solution, of the problem of constructing connection formulas for second-order differential equations having an arbitrary number of turning points of arbitrary multiplicities.

The method used is a generalization of the central connection procedure developed by Jeffreys for two simple turning points, and employs only real-variable theory in the analysis. In consequence, the coefficients in the differential equation need not be analytic functions of the independent variable, nor is there any need to investigate the topology in the complex plane of the principal curves (or anti-Stokes lines).

Acknowledgment. The numerical integrations of the differential equation (5.35) used to check Example 5 were carried out by Mr. R. E. Kaylor. The author is pleased to acknowledge this assistance.

REFERENCES

[1] D. BOHM, *Quantum Theory*, Prentice-Hall, Englewood Cliffs, N.J., 1951.

[2] M. A. EVGRAFOV AND M. V. FEDORYUK, *Asymptotic behavior as $\lambda \to \infty$ of solutions of the equation $w''(z)-p(z,\lambda)w(z)=0$ in the complex z-plane*, Uspehi Mat. Nauk, 21 (1966), pp. 3–50 = Russian Math. Surveys, 21 (1966), pp. 1–48.

[3] M. V. FEDORYUK, *Asymptotics of the discrete spectrum of the operator $w''(x)-\lambda^2 p(x)w(x)$*, Mat. Sb., 68 (1965), pp. 81–110. (In Russian.)

[4] N. FRÖMAN AND Ö. DAMMERT, *Tunneling and super-barrier transmission through a system of two real potential barriers*, Nuclear Phys. A, 147 (1970), pp. 627–649.

[5] N. FRÖMAN AND P. O. FRÖMAN, *JWKB Approximation—Contributions to the Theory*, North-Holland, Amsterdam, 1965.

[6] N. FRÖMAN, P. O. FRÖMAN, U. MYHRMAN AND R. PAULSSON, *On the quantal treatment of the double-well potential problem by means of certain phase-integral approximations*, Ann. Physics, 74 (1972), pp. 314–323.

[7] J. HEADING, *An Introduction to Phase-integral Methods*, John Wiley, New York, 1962.

[8] ———, *Phase-integral methods I*, Quart, J. Mech. Appl. Math., 15 (1962), pp. 215–244.

[9] ———, *Exact and approximate methods for the investigation of the propagation of waves through a system of barriers*, Proc. Cambridge Philos. Soc., 74 (1973), pp. 161–178.

[10] ———, *The approximate use of the complex gamma function in some wave propagation problems*, J. Phys. A, 6 (1973), pp. 958–973.

[11] ———, *Barriers bounded by a transition point of order greater than unity*, Physica, 77 (1974), pp. 263–278.

[12] ———, *Barriers bounded by two transition points of arbitrary odd order*, Proc. Roy. Soc. Edinburgh, 73A (1974/5), pp. 51–64.

[13] H. JEFFREYS, *On approximate solutions of linear differential equations*, Proc. Cambridge Philos. Soc., 49 (1953), pp. 601–611.

[14] ———, *On the use of asymptotic approximations of Green's type when the coefficient has zeros*, Ibid., 52 (1956), pp. 61–66.

[15] A. W-K. LEUNG, *Connection formulas for asymptotic solutions of second order turning points in unbounded domains*, this Journal, 4 (1973), pp. 89–103.

[16] ———, *Errata: connection formulas for asymptotic solutions of second order turning points in unbounded domains*, this Journal, 6 (1975), p. 600.

[17] ———, *Distribution of eigenvalues in the presence of higher order turning points*, Trans. Amer. Math. Soc., to appear.

[18] R. Y. S. LYNN AND J. B. KELLER, *Uniform asymptotic solutions of second order linear ordinary differential equations with turning points*, Comm. Pure Appl. Math., 23 (1970), pp. 379–408.

[19] J. C. P. MILLER, *On the choice of standard solutions for a homogeneous linear differential equation of the second order*, Quart. J. Mech. Appl. Math., 3 (1950), pp. 225–235.

[20] E. L. MURPHY AND R. H. GOOD, JR., *WKB Connection formulas*, J. Math. and Phys., 43 (1964), pp. 251–254.

[21] T. NISHIMOTO, *On an extension theorem and its application for turning point problems of large order*, Kōdai Math. Sem. Rep., 25 (1973), pp. 458–489.

[22] F. W. J. OLVER, *Error analysis of phase-integral methods. I. General theory for simple turning points*, J. Res. Nat. Bur. Standards Sect. B, 69 (1965), pp. 271–290.

[23] ———, *Error analysis of phase-integral methods. II. Application to wave-penetration problems*, Ibid., 69 (1965), pp. 291–300.

[24] ———, *Asymptotics and Special Functions*, Academic Press, New York, 1974.

[25] ———, *Second-order linear differential equations with two turning points*, Philos. Trans. Roy. Soc. London Ser. A, 278 (1975), pp. 137–174.

[26] ———, *Unsolved problems in the asymptotic estimation of special functions*, Theory and Applications of Special Functions, R. A. Askey, ed., Academic Press, New York, 1975, pp. 99–142.

[27] ———, *Connection formulas for second-order differential equations with multiple turning points*, this Journal, 8 (1977), pp. 127–154.

[28] E. R. PIKE, *On the related-equation method of asymptotic approximation (W.K.B. or A-A method) I. A proposed new existence theorem*, Quart. J. Mech. Appl. Math., 17 (1964), pp. 105–124.

[29] Y. SIBUYA, *Subdominant solutions of the differential equation* $y''-\lambda^2(x-a_1)(x-a_2)\cdots(x-a_m)y=0$, Acta Math., 119 (1967), pp. 235–272.

[30] ———, *Global Theory of a Second Order Linear Ordinary Differential Equation with a Polynomial Coefficient,* Mathematics Studies 18, North-Holland, Amsterdam, 1975.

[31] L. WEINBERG, *The asymptotic distribution of eigenvalues for the boundary value problem* $y''(x)-\lambda^2 p(x)y(x)=0$, $y\in L_2(-\infty,+\infty)$, this Journal, 2 (1971), pp. 546–566.

[32] S. YNGVE, *Normalization of certain higher-order phase-integral approximations for wavefunctions of bound states in a potential well. II,* J. Mathematical Phys., 13 (1972), pp. 324–331.

SIAM J. MATH. ANAL.
Vol. 8, No. 4, August 1977

NOTE ON SOME CONVOLVED POWER SUMS*

L. CARLITZ†

Abstract. Neuman and Schonbach have obtained explicit formulas for the sum

$$S(i,j;N)=\sum_{k=0}^{N} k^i(N-k)^j \qquad (i,j\geqq 0)$$

by using known results involving Bernoulli numbers. In the present paper the functions

$$S(i,j;N;a)=\sum_{k=0}^{N}(k+a)^i(N-k-a)^j \qquad (i,j\geqq 0),$$

$$S'(i,j;N;a)=iS(i-1,j;N;a)-jS(i,j-1;N;a),$$

where a is arbitrary, are evaluated. The evaluation of $S'(i,j;N;a)$ makes use of an appropriate generating function. The final formula for $S(i,j;N;a)$ reduces to the Neuman–Schonbach result when $a=0$.

In addition it is shown that the sum $S(i,j;N)$ is closely related to the Eulerian numbers.

1. Neuman and Schonbach [3] have obtained explicit formulas for the sum

$$S(i,j;N)=\sum_{k=0}^{N} k^i(N-k)^j, \tag{1.1}$$

where i, j and arbitrary nonnegative integers. They show in particular that

$$S(i,j;N)=P_{i,j}(N), \tag{1.2}$$

where $P_{i,j}(x)$ is a polynomial of degree $i+j+1$ in x. The coefficients are simple multiples of Bernoulli numbers; also

$$P_{ij}(-x)=(-1)^{i+j+1}P_{i,j}(x) \qquad (i+j>0). \tag{1.3}$$

In this note we make some additional comments concerning the problem. It is well known that the sum [4, Chap. 2]

$$\sum_{k=0}^{n-1} i(k+a)^{i-1}=B_i(n+a)-B_i(a) \qquad (i>0), \tag{1.4}$$

where a is an arbitrary parameter and $B_i(a)$ is the Bernoulli polynomial of degree i defined by

$$\frac{xe^{ax}}{e^x-1}=\sum_{i=0}^{\infty} B_i(a)\frac{x^i}{i!}. \tag{1.5}$$

Thus

$$B_i(a)=\sum_{k=0}^{i}\binom{i}{k}B_k a^{i-k}, \qquad B_k=B_k(0), \tag{1.6}$$

where the B_k are the Bernoulli numbers.

* Received by the editors June 19, 1975, and in revised form February 28, 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.

In view of (1.4) it is natural to consider the sum

$$S(i,j;N;a)=\sum_{k=0}^{N}(k+a)^i(N-k-a)^j, \tag{1.7}$$

where again i, j are nonnegative integers and a is arbitrary. Then clearly

$$\sum_{i,j=0}^{\infty}S(i,j;N;a)\frac{x^iy^i}{i!j!}=\sum_{k=0}^{N}e^{(k+a)}e^{(N-k-a)y},$$

so that

$$\sum_{i,j,N=0}^{\infty}S(i,j;N;a)\frac{x^iy^j}{i!j!}z^N=\frac{e^{a(x-y)}}{(1-e^xz)(1-e^yz)}. \tag{1.8}$$

Since

$$\frac{1}{(1-az)(1-bz)}=\frac{1}{a-b}\left(\frac{a}{1-az}-\frac{b}{1-bz}\right),$$

we get

$$\begin{aligned}\frac{e^{a(x-y)}}{(1-e^xz)(1-e^yz)}&=\frac{e^{a(x-y)}}{e^x-e^y}\left(\frac{e^x}{1-e^xz}-\frac{e^y}{1-e^yz}\right)\\&=\frac{e^{a(x-y)}}{1-e^{y-x}}\frac{1}{1-e^xz}-\frac{e^{a(x-y)}}{e^{x-y}-1}\frac{1}{1-e^yz}.\end{aligned}$$

Thus

$$\frac{(x-y)e^{-a(x-y)}}{(1-e^xz)(1-e^yz)}=\frac{(y-x)e^{-a(y-x)}}{e^{y-x}-1}\frac{1}{1-e^xz}-\frac{(x-y)e^{a(x-y)}}{e^{x-y}-1}\frac{1}{1-e^yz}. \tag{1.9}$$

We now apply (1.5). Thus

$$\frac{(x-y)e^{a(x-y)}}{e^{x-y}-1}=\sum_{n=0}^{\infty}B_n(a)\frac{(x-y)^n}{n!}=\sum_{i,j=0}^{\infty}(-1)^jB_{i+j}(a)\frac{x^iy^j}{i!j!}$$

and

$$\frac{(y-x)e^{-a(y-x)}}{e^{y-x}-1}=\sum_{i,j=0}^{\infty}(-1)^iB_{i+j}(-a)\frac{x^iy^j}{i!j!}.$$

Since

$$\frac{1}{1-e^xz}=\sum_{N=0}^{\infty}e^{Nx}z^N=\sum_{t,N=0}^{\infty}\frac{N^rx^r}{r!}z^N,$$

it follows that the right-hand side of (1.9) is equal to

$$\sum_{i,j,N=0}^{\infty}\frac{x^iy^j}{i!j!}z^N\left\{\sum_{r=0}^{i}(-1)^{i-r}\binom{i}{r}B_{i+j-r}(-a)-\sum_{r=0}^{j}(-1)^{j-r}\binom{j}{r}B_{i+j-r}(a)\right\}. \tag{1.10}$$

By (1.8),

$$\frac{(x-y)e^{a(x-y)}}{(1-e^xz)(1-e^yz)}=\sum_{i,j,N=0}^{\infty}S'(i,j;N;a)\frac{x^iy^j}{i!j!}z^N,$$

where

$$S'(i,j;N;a)=iS(i-1,j;N;a)-jS(i,j-1;N;a). \tag{1.11}$$

Comparison with (1.10) now gives

$$S'(i,j;N;a)=\sum_{r=0}^{i}(-1)^{i-r}\binom{i}{r}B_{i+j-r}(-a)N^r-\sum_{r=0}^{j}(-1)^{j-r}\binom{j}{r}B_{i+j-r}(a)N^r. \tag{1.12}$$

For $a=0$, (1.12) reduces to

$$\begin{aligned} &iS(i-1,j;N)-jS(i,j-1;N)\\ &\quad=\sum_{r=0}^{i}(-1)^{i-r}\binom{i}{r}B_{i+j-r}N^r-\sum_{r=0}^{j}(-1)^{j-r}\binom{j}{r}B_{i+j-r}N^r. \end{aligned} \tag{1.13}$$

Note that by (1.1) and (1.11),

$$S'(i,j;N;a)=\sum_{k=0}^{n}\{i(N-k-a)-j(k+a)\}(k+a)^{i-1}(N-k-a)^{j-1}, \tag{1.14}$$

so that

$$S'(i,j;N;a)=\frac{\partial}{\partial a}S(i,j;N;a), \tag{1.15}$$

thus justifying the notation.

2. It should be remarked that one frequently encounters summation formulas involving an appropriate derivative. Indeed (1.4) can be thought of as an instance of this. A better example is the following.

Put

$$S_n(m;\xi)=\sum_{t=0}^{m-1}(t+\xi_1)^{n_1}\cdots(t+\xi_k)^{n_k},$$

$$S'_n(m;\xi)=\sum_{t=0}^{m-1}(d/dt)\{(t+\xi_1)^{n_1}\cdots(t+\xi_k)^{n_k}\}.$$

Then [2, p. 752]

$$S'_n(m;\xi)=B_{n_1,\cdots,n_k}(\xi_1+m,\cdots,\xi_k+m)-B_{n_1,\cdots,n_k}(\xi_1,\cdots,\xi_k), \tag{2.1}$$

where

$$\sum_{n_i=0}^{\infty}B_{n_1,\cdots,n_k}(\xi_1,\cdots,\xi_k)\frac{x_1^{n_1}\cdots x_k^{n_k}}{n_1!\cdots n_k!}=\frac{(x_1+\cdots+x_k)\,e^{\xi_1x_1+\cdots+\xi_kx_k}}{e^{x_1+\cdots+x_k}-1}.$$

On the other hand

$$S_n(m;\xi)=\sum_{s_i+t_i=n_i}\binom{n_1}{t_1}\cdots\binom{n_k}{t_k}B_{s_1,\cdots,s_k}(\xi_1,\cdots,\xi_k)\frac{m^{t_1+\cdots+t_k+1}}{t_1+\cdots+t_k+1}. \tag{2.2}$$

If we integrate (1.12) with respect to a, we get

$$S(i,j;N;a)=A(i,j;N)+\sum_{r=0}^{i}(-1)^{i-r+1}\binom{i}{r}\frac{B_{i+j-r+1}(-a)}{i+j-r+1}N^r + \sum_{j=0}^{j}(-1)^{j-r+1}\binom{j}{r}\frac{B_{i+j-r+1}(a)}{i+j-r+1}N^r, \tag{2.3}$$

where $A(i,j;N)$ is independent of a.

To find $A(i,j;N)$ we return to (1.11). Changing the notation slightly, we have

$$rS(r-1,i+j-r+1;N;a)-(i+j-r+1)S(r,i+j-r;N;a) = S'(r,i+j-r+1;N;a) \qquad (1\leqq r\leqq i). \tag{2.4}$$

Multiply both sides of (2.4) by

$$\frac{1}{i+j-r+1}\binom{i+j}{r}=\frac{1}{i+j+1}\binom{i+j+1}{r}$$

so that

$$\binom{i+j}{r-1}S(r-1,i+j-r+1;N;a)-\binom{i+j}{r}S(r,i+j-r;N;a) = \frac{1}{i+j+1}\binom{i+j+1}{r}S'(r,i+j-r+1;N;a) \qquad (1\leqq r\leqq i). \tag{2.5}$$

For $r=0$ we have

$$-S(0,i+j;N;a)=\frac{1}{i+j+1}S'(0,i+j+1;N;a), \tag{2.6}$$

since by (1.12)

$$\begin{aligned} S'(0,i+j+1;N;a) &= B_{i+j+1}(-a)-\sum_{r=0}^{i+j+1}(-1)^{i+j-r+1}\binom{i+j+1}{r}B_{i+j-r+1}(a)N^r \\ &= B_{i+j+1}(-a)-(-1)^{i+j-r+1}(a-N) \\ &= B_{i+j+1}(-a)-B_{i+j+1}(N-a+1) \end{aligned}$$

and

$$B_n(1-x)=(-1)^nB_n(x).$$

On the other hand

$$\begin{aligned} S(0,i+j;N;a) &= \sum_{k=0}^{N}(N-k-a)^{i+j}=\sum_{k=0}^{N}(k-a)^{i+j} \\ &= \frac{1}{i+j+1}\{B_{i+j+1}(N-a+1)-B_{i+j+1}(-a)\}. \end{aligned}$$

It now follows from (2.5) and (2.6) that

$$-\binom{i+j}{i}S(i,j;N;a)=\frac{1}{i+j+1}\sum_{r=0}^{i}\binom{i+j+1}{r}S'(r,i+j-r+1;N;a). \tag{2.7}$$

Substituting from (1.12) in (2.7) we get

$$\binom{i+j}{i}S(i,j;N;a)=\frac{1}{i+j+1}\sum_{r=0}^{i}\binom{i+j+1}{r}\Big\{-\sum_{t=0}^{r}(-1)^{r-t}\binom{r}{t}B_{i+j-t+1}(-a)N^t + \sum_{t=0}^{i+j-r+1}(-1)^{i+j-r-t+1}\binom{i+j-r+1}{t}B_{i+j-r+1}(a)N^t\Big\}. \tag{2.8}$$

We have

$$\frac{1}{i+j+1}\sum_{r=0}^{i}\binom{i+j+1}{r}\sum_{t=0}^{r}(-1)^{r-t}B_{i+j-t+1}(-a)N^t$$
$$=\frac{1}{i+j+1}\sum_{t=0}^{i}B_{i+j-t+1}(-a)N^t\sum_{r=t}^{i}(-1)^{r-t}\binom{i+j+1}{r}\binom{r}{t}$$
$$=\binom{i+j}{i}\sum_{t=0}^{i}(-1)^{i-t}\binom{i}{t}\frac{B_{i+j-t+1}(-a)}{i+j-t+1}N^t.$$

Similarly

$$\frac{1}{i+j+1}\sum_{r=0}^{i}\binom{i+j+1}{r}\sum_{t=0}^{i+j-r+1}(-1)^{i+j-r-t+1}\binom{i+j-r+1}{t}B_{i+j-t+1}(a)N^t$$
$$=\frac{1}{i+j+1}\sum_{t=0}^{i+j+1}B_{i+j-t+1}(a)N^t\sum_{r=0}^{\min(i,i+j-t+1)}(-1)^{i+j-r-t+1}\binom{i+j+1}{r}\binom{i+j-r+1}{t}.$$

We find that

$$\frac{1}{i+j+1}\sum_{r=0}^{\min(i,i+j-t+1)}(-1)^{i+j-r-t+1}\binom{i+j+1}{r}\binom{i+j-r+1}{t}$$
$$=\begin{cases}\dfrac{-(-1)^{j-t}}{i+j-t+1}\binom{i+j}{i}\binom{j}{t} & (0\leqq t\leqq j),\\ 0 & (j<t\leqq i+j),\\ \dfrac{1}{i+j+1} & (t=i+j+1).\end{cases}$$

Thus (2.8) becomes

$$S(i,j;N;a)=\frac{i!j!}{(i+j+1)!}N^{i+j+1}+\sum_{t=0}^{i}(-1)^{i-t+1}\binom{i}{t}\frac{B_{i+j-t+1}(-a)}{i+j-t+1}N^t + \sum_{t=0}^{j}(-1)^{j-t+1}\binom{j}{t}\frac{B_{i+j-t+1}(a)}{i+j-t+1}N^t. \tag{2.9}$$

This may be written in the form

$$S(i,j;N;a)=\frac{i!j!}{(i+j+1)!}N^{i+j+1}+\sum_{k=j+1}^{i+j}(-1)^{j-k}\binom{i}{i+j-k+1}\frac{B_k(-a)}{k}N^{i+j-k+1} + \sum_{k=i+1}^{i+j}(-1)^{i-k}\binom{j}{i+j-k+1}\frac{B_k(a)}{k}N^{i+j-k+1}. \tag{2.10}$$

For $a=0$, (2.10) reduces to the formula of Neuman and Schonbach.

Comparing (2.10) with (2.3), we get

$$A(i,j;N)=\frac{i!j!}{(i+j+1)!}N^{i+j+1}. \tag{2.11}$$

3. The Eulerian numbers [1], [5, Chap. 8] may be defined in the following way. Put

$$a_n(z)=(1-z)^{n+1}\sum_{k=0}^{\infty}k^n z^k. \tag{3.1}$$

Then $a_n(z)$ is a polynomial of degree n:

$$a_n(z)=\sum_{k=1}^{n}A_{nk}z^k \qquad (n\geqq 1); \tag{3.2}$$

the A_{nk} are the Eulerian numbers.

It follows from (3.1) that

$$\sum_{n=0}^{\infty}a_n(z)\frac{x^n}{n!}=\frac{1-z}{1-z\,e^{(1-z)x}}. \tag{3.3}$$

Thus by (1.8), with $a=0$,

$$\sum_{i,j,N=0}^{\infty}S(i,j;N)\frac{x^iy^j}{i!j!}z^N=\frac{1}{(1-e^xz)(1-e^yz)}$$

$$=(1-z)^{-2}\sum_{i,j=0}^{\infty}\frac{a_i(z)}{(1-z)^i}\frac{a_j(z)}{(1-z)^j}\frac{x^iy^j}{i!j!}.$$

Comparing coefficients of x^iy^j we get

$$\sum_{N=0}^{\infty}S(i,j;N)z^N=(1-z)^{-i-j-2}a_i(z)a_j(z). \tag{3.4}$$

The right-hand side of (3.4) is equal to

$$\sum_{r=0}^{\infty}\binom{i+j+r+1}{r}\sum_{s=0}^{i}A_{is}z^s\sum_{t=0}^{j}A_{jt}z^t.$$

For convenience we take

$$A_{n0}=\begin{cases}1 & (n=0),\\ 0 & (n=1,2,3,\cdots).\end{cases} \tag{3.5}$$

Thus (3.4) gives

$$S(i,j;N)=\sum_{\substack{r+s+t=N\\ s\leqq i,t\leqq j}}\binom{i+j+r+1}{r}A_{is}A_{jt}=\sum_{\substack{s+t\leqq N\\ s\leqq i,t\leqq j}}\binom{i+j+N-s-t+1}{N-s-t}A_{is}A_{jt}. \tag{3.6}$$

Alternatively if we write (3.4) in the form

$$(1-z)^{i+j+2}\sum_{N=0}^{\infty}S(i,j;N)z^N=a_i(z)a_j(z),$$

we get

$$\sum_{r+s=n} A_{ir}A_{js} = \sum_{r+s=n} (-1)^s \binom{i+j+2}{s} S(i,j;r). \tag{3.7}$$

It is well known that

$$A_{n,k} = A_{n,n-k+1} \qquad (1 \leqq k \leqq n). \tag{3.8}$$

Thus for $i, j \geqq 1$,

$$\sum_{r+s=n} A_{ir}A_{js} = \sum_{r+s=n} A_{i,i-r+1}A_{j,j-s+1} = \sum_{r+s=i+j-n+2} A_{ir}A_{js}.$$

Let $T(i,j;n)$ denote the right-hand side of (3.7), so that we have proved that

$$T(i,j;n) = T(i,j;i+j-n+2). \tag{3.9}$$

It is not difficult to prove (3.9) directly without (3.8). Indeed

$$\begin{aligned} T(i,j;i+j-n+2) &= \sum_{r=0}^{i+j-n+2} (-1)^{i+j-n-r} \binom{i+j+2}{i+j-n-r+2} S(i,j;r) \\ &= (-1)^{i+j} \sum_{r=0}^{i+j-n+2} (-1)^{n+r} \binom{i+j+2}{n+r} S(i,j;r) \\ &= (-1)^{i+j} \sum_{r=n}^{i+j+2} (-1)^r \binom{i+j+2}{r} S(i,j;r-n). \end{aligned}$$

Since $S(i,j;N)$ is a polynomial in N of degree $i+j+1$, we have

$$\sum_{r=0}^{i+j+2} (-1)^r \binom{i+j+2}{r} S(i,j;r-n) = 0,$$

so that

$$\begin{aligned} T(i,j;i+j-n+2) &= (-1)^{i+j+1} \sum_{r=0}^{n-1} (-1)^r \binom{i+j+2}{r} S(i,j;r-n) \\ &= (-1)^{i+j+1} \sum_{r=0}^{n} (-1)^r \binom{i+j+2}{r} S(i,j;r-n). \end{aligned}$$

Finally by (1.3)

$$S(i,j;r-n) = (-1)^{i+j+1} S(i,j;n-r).$$

This completes the proof of (3.9).

We remark that (1.8) is easily generalized. Thus if we put

$$S(i,j,k;N;a,b) = \sum_{r+s \leqq N} (r+a)^i (s+b)^j (N-r-s-a-b)^k \tag{3.10}$$

then

$$\sum_{i,j,k=0}^{\infty} S(i,j,k;N;a,b) \frac{x^i y^j z^k}{i!j!k!} = \sum_{r+s \leqq N} e^{(r+a)x} e^{(s+b)y} e^{(N-r-s-a-b)z}.$$

Hence

$$(3.11)\qquad \sum_{N=0}^{\infty}\sum_{i,j,k=0}^{\infty} S(i,j,k;N;a,b)\frac{x^i y^j z^k}{i!j!k!}u^N=\frac{e^{a(x-z)+b(y-z)}}{(1-e^x u)(1-e^y u)(1-e^z u)}.$$

For $a=b=0$, it is clear how (3.6) and (3.7) can be generalized. On the other hand, generalization of (1.12) and (2.10) seems much more complicated. The extension of (1.2) makes use of the partial fraction decomposition

$$\frac{1}{(1-au)(1-bu)(1-cu)}=\sum\frac{a^2}{(a-b)(a-c)}\frac{1}{1-au}$$

where $\sum$ indicates cyclic permutation of a, b, c. Then

$$\frac{1}{(1-e^x u)(1-e^y u)(1-e^z u)}=\sum\frac{e^{2x}}{(e^x-e^y)(e^x-e^z)}\frac{1}{1-e^x u},$$

$$\frac{(y-z)(z-x)(x-y)}{(1-e^x u)(1-e^y u)(1-e^z u)}=-\sum\frac{(y-x)(z-x)}{(e^{y-x}-1)(e^{z-x}-1)}\frac{y-z}{1-e^x u}.$$

We have

$$\frac{(y-x)(z-x)}{(e^{y-x}-1)(e^{z-x}-1)}=\sum_{r,s,j,k=0}^{\infty}(-1)^{r+s}B_{r+j}B_{s+k}\frac{x^{r+s}y^j z^k}{r!s!j!k!}$$

$$=\sum_{i,j,k=0}^{\infty}(-1)^i b(i,j,k)\frac{x^i y^j z^k}{i!j!k!},$$

where, for brevity, we put

$$(3.12)\qquad b(i,j,k)=\sum_{r+s=i}\binom{i}{r}B_{r+j}B_{s+k}.$$

Put

$$(3.13)\qquad S'(i,j,k;N)=\sum ij(j-1)S(i-1,j-2,k;N)-\sum i(i-1)jS(i-2,j-1,k;N)$$

where

$$S(i,j,k;N)=S(i,j,k;N;0,0)$$

and the $\sum$ refers to cyclic permutation of i, j, k. Then we have

$$(3.14)\qquad S'(i,j;k;N)=\sum_{i,j,k}\sum_{r=0}^{i}\binom{i}{r}b'(i-r,j,k)N^r,$$

where

$$b'(i,j,k)=jb(i,j-1,k)-kb(i,j,k-1)$$

and $\sum_{i,j,k}$ indicates cyclic permutation of i, j, k.

REFERENCES

[1] L. CARLITZ, *Eulerian numbers and polynomials*, Math. Mag., 33 (1959), pp. 247–260.
[2] ———, *On arrays of numbers*, Amer. J. Math., 54 (1932), pp. 739–752.
[3] C. P. NEUMAN AND D. I. SCHONBACH, *Evaluation of sums of convolved powers using Bernoulli numbers*, SIAM Rev., to appear.
[4] N. E. NÖRLUND, *Vorlesungen über Differenzenrechnung*, Springer-Verlag, Berlin, 1924.
[5] JOHN RIORDAN, *An Introduction to Combinatorial Analysis*, John Wiley, New York, 1958.

SIAM J. MATH. ANAL.
Vol. 8, No. 4, August 1977

UNIFORM ASYMPTOTIC SOLUTIONS OF NONHOMOGENEOUS DIFFERENTIAL EQUATIONS WITH TURNING POINTS*

DONATUS U. ANYANWU†

Abstract. Langer's methods of the comparison equation and its generalization by R. Y. S. Lynn and J. B. Keller is extended for the treatment of a nonhomogeneous second order ordinary differential equation with an arbitrary number of turning points. The development is purely formal but the method is verified with some examples.

1. Introduction. We wish to obtain a formal uniform asymptotic solution, for $|\lambda|$ large, of the differential equation

$$\frac{d^2u}{dz^2}+\lambda^2R(z,\lambda)u=\lambda^2G(z,\lambda), \qquad z\in\mathscr{D}, \quad \lambda\in\mathscr{S}. \tag{1.1}$$

Here $\mathscr{D}$ is either a complex domain or a real interval and $\mathscr{S}$ is either a sector of the complex λ-plane or a semi-infinite interval of the real axis. To do so we shall utilize a generalization of R. E. Langer's method [1] by R. Y. S. Lynn and J. B. Keller [2]. We will assume that $R(z,\lambda)$ and $G(z,\lambda)$ are analytic or infinitely differentiable in $\mathscr{D}$.

Let z_j, $j=0,\cdots,N$, be zeros of orders m_j of $R(z,\infty)$, i.e. the turning points, and define $\mu=\sum_{j=0}^{N}m_j$. We propose a solution to (1.1) of the form

$$\begin{aligned}u(z,\lambda)=&B(z,\lambda)V[\xi(z,\lambda)]\\&+\lambda^{-\mu/(\mu+2)}G(z,\lambda)V'[\xi(z,\lambda)]+H(z,\lambda)\end{aligned} \tag{1.2}$$

where $V(\xi)$ is a solution of the comparison equation

$$V''(\xi)+\left(\sum_{k=0}^{\mu}\gamma_k\xi^k\right)V(\xi)=\sum_{k=0}^{\mu}\Gamma_k\xi^k. \tag{1.3}$$

The constants γ_k, $k=0,1,\cdots,\mu$, and the functions $B(z,\lambda)$, $C(z,\lambda)$ were introduced by Lynn and Keller [2] in the treatment of the linear homogeneous problem. The constants Γ_k and the function $H(z,\lambda)$ are now introduced for the nonhomogeneous problem. If $V_1(\xi)$ and $V_2(\xi)$ are two linearly independent solutions of the homogeneous form of (1.3), then

$$V(\xi)=Q_1(\xi)V_1(\xi)+Q_2(\xi)V_2(\xi) \tag{1.4}$$

* Received by the editors May 8, 1975, and in final revised form March 20, 1976.

† Department of Mathematics, University of Nigeria, Nsukka, Nigeria.

is a solution to (1.3), with

$$
\begin{aligned}
Q_1(\xi) &= -\frac{1}{W}\int^{\xi}\Big(\sum_{k=0}^{\mu}\Gamma\eta^k\Big)V_2(\eta)\,d\eta + c_1,\\
Q_2(\xi) &= \frac{1}{W}\int^{\xi}\Big(\sum_{k=0}^{\mu}\Gamma_k\eta^k\Big)V_1(\eta)\,d\eta + c_2.
\end{aligned}
\tag{1.5}
$$

In (1.5) the constant W is the Wronskian of $V_1(\xi)$ and $V_2(\xi)$, and c_1, c_2 are arbitrary constants. It follows that once the asymptotic properties of $V_1(\xi)$ and $V_2(\xi)$ are known, those of $V(\xi)$ can be derived.

A uniform representation of $u(z,\lambda)$ in (1.1) for the case $\mu = 0$ has been obtained by F. W. J. Olver [3, pp. 386–390] for some special $R(z,\lambda)$. The case $\mu = 1$ was treated by R. A. Clark [4], C. R. Steele [5], A. H. Nayfeh [6, pp. 352–356] and again by Olver [3, pp. 429–433]; while the case $\mu = 2$ and $N = 0$ was treated by Nayfeh [6, pp. 356–358] and by D. J. McGuiness [7]. The above treatments were for special forms of $R(z,\lambda)$. Our purpose here is to obtain the formal solution for the general problem.

2. The equation for $H(z,\lambda)$. We substitute (1.2) into (1.1) and we also use (1.3) to obtain the following equation:

$$
\begin{aligned}
&\Big\{B''-(\xi')^2B\sum_{k=0}^{\mu}\gamma_k\xi^k-\lambda^{-\mu/(\mu+2)}\Big[2\xi'C'\sum_{k=0}^{\mu}\gamma_k\xi^k+\xi''C\sum_{k=0}^{\mu}\gamma_k\xi^k\\
&\qquad+(\xi')^2C\Big(\sum_{k=0}^{\mu}\gamma_k\xi^k\Big)'\Big]+\lambda^2RB\Big\}V\\
&+\Big\{2\xi'B'+\xi''B+\lambda^{-\mu/(\mu+2)}\Big[C''-(\xi')^2C\sum_{k=0}^{\mu}\gamma_k\xi^k+\lambda^2RC\Big]\Big\}V'\\
&+\Big\{(\xi')^2B\sum_{k=0}^{\mu}\Gamma_k\xi^k+\lambda^{-\mu/(\mu+2)}\xi'C'\sum_{k=0}^{\mu}\Gamma_k\xi^k+\lambda^{-\mu/(\mu+2)}\Big(\xi'C\sum_{k=0}^{\mu}\Gamma_k\xi^k\Big)'\\
&\qquad+\lambda^2(RH-G)+H''\Big\}=0.
\end{aligned}
\tag{2.1}
$$

Here primes denote differentiation with respect to z, and we have assumed that $\xi = \xi(z)$. Next we equate the coefficients of V and V' to zero and thus obtain (2.2) and (2.3) of Lynn and Keller [2], for the determination of $B(z,\lambda)$and $C(z,\lambda)$. There is in addition the following equation for $H(z,\lambda)$:

$$
\begin{aligned}
&(\xi')^2B\sum_{k=0}^{\mu}\Gamma_k\xi^k+\lambda^{-\mu/(\mu+2)}\xi'C'\sum_{k=0}^{\mu}\Gamma_k\xi^k\\
&\qquad+\lambda^{-\mu/(\mu+2)}\Big(\xi'C\sum_{k=0}^{\mu}\Gamma_k\xi^k\Big)'+\lambda^2(RH-G)+H''=0.
\end{aligned}
\tag{2.2}
$$

Lynn and Keller in [2] set $\xi(z) = \lambda^{2/(\mu+2)}\phi(z)$. By assuming a representation for the constants $\gamma_k(\lambda)$ of the form

$$
\gamma_k \sim \lambda^{2(\mu-k)/(\mu+2)}\sum_{p=0}^{\infty}\lambda^{-p}\gamma_{kp}, \qquad k = 0, 1, \cdots, \mu,
\tag{2.3}
$$

and by also assuming

$$R(z,\lambda)\sim\sum_{p=0}^{\infty}\lambda^{-p}R_p(z), \tag{2.4}$$

they determined $B(z,\lambda)$ and $C(z,\lambda)$ in the form

$$B(z,\lambda)\sim\sum_{p=0}^{\infty}\lambda^{-p}B_p(z);\qquad C(z,\lambda)\sim\sum_{p=0}^{\infty}\lambda^{-p}C_p(z). \tag{2.5}$$

The function $\phi(z)$ was found to satisfy the differential equation

$$\left(\frac{d\phi}{dz}\right)^2=\frac{R_0(z)}{\sum_{k=0}^{\mu}\gamma_{k_0}\phi^k}=\frac{R_0(z)}{\gamma_{\mu 0}\prod_{j=0}^{N}(\phi-\phi_j)^{m_j}}, \tag{2.6}$$

where $\phi_j=\phi(z_j), j=0,1,\cdots,N$, are constants. Furthermore, the functions $B_p(z)$ and $C_p(z)$ in (2.5) were explicitly determined and have the form:

$$B_p(z)=\left(\frac{1}{\phi'}\right)^{1/2}\cos(\theta(z)-\psi_p(z)), \tag{2.7}$$

$$C_p(z)=\left(\frac{\phi'}{R_0}\right)^{1/2}\sin(\theta(z)-\psi_p(z)) \tag{2.8}$$

where

$$\theta=\theta(z,z_0,\gamma_{k1});\quad \psi_p=\psi_p(z,z_0,\gamma_{k,p+1}),\qquad p=1,2,\cdots, \tag{2.9}$$

with $\psi_0=0$. Regularity conditions on $B_p(z)$ and $C_p(z)$ required that $\mu-1$ of the $\mu+1$ constants γ_{kp}, $p=1,\cdots$, be specially chosen. $\theta(z)$ and $\psi_p(z)$ contain arbitrary integration constants.

We will in the nonhomogeneous case assume

$$G(z,\lambda)\sim\sum_{p=0}^{\infty}\lambda^{-p}G_p(z) \tag{2.10}$$

and seek to determine $H(z,\lambda)$ in (2.2) in the form

$$H(z,\lambda)\sim\sum_{p=0}^{\infty}\lambda^{-p}H_p(z). \tag{2.11}$$

Regularity of the coefficients, $H_p(z)$, at the turning points will be achieved by utilizing the coefficients Γ_{kp} in the expansion

$$\Gamma_k\sim\lambda^{2(\mu-k)/(\mu+2)}\sum_{p=0}^{\infty}\lambda^{-p}\Gamma_{kp},\qquad k=0,1,\cdots,\mu. \tag{2.12}$$

It will then follow that the asymptotic representation of $u(z,\lambda)$ in (1.2) will be valid everywhere in, at least, a subdomain of $\mathscr{D}$ containing the turning points, once the appropriate linear combination $V(\xi)$ of (1.3) is chosen.

3. The equation for $H_p(z)$. On substituting (2.4), (2.5), (2.10)–(2.12) into (2.2) and equating coefficients of like powers of λ, we obtain the following

equation for $H_p(z)$:

$$R_0(z)H_p(z) = G_p(z) - (\phi')^2 B_0(z) \sum_{k=0}^{\mu} \Gamma_{kp}\phi^k - J_p(z), \qquad p = 0, 1, \cdots. \tag{3.1}$$

Here

$$\begin{aligned} J_p(z) = {} & \sum_{q=1}^{p} R_q H_{p-q} + (\phi')^2 \sum_{q=0}^{p-1} B_{p-q} \sum_{k=0}^{\mu} \Gamma_{kq}\phi^k + H''_{p-2} \\ & + \phi' \sum_{q=0}^{p-1} C'_{p-q-1} \sum_{k=0}^{\mu} \Gamma_{kq}\phi^k + \left(\phi' \sum_{q=0}^{p-1} C_{p-q-1} \sum_{k=0}^{\mu} \Gamma_{kq}\phi^k\right)', \end{aligned} \tag{3.2}$$

where terms with negative sub-index are zero. We observe that $J_0(z) = 0$. Since $R_0(z)$ has a zero or order m_j at z_j the R.H.S. (right-hand side) of (3.1) together with its first $m_j - 1$ derivatives must also vanish at z_j in order that $H_p(z)$ be regular at that point. Thus we have the following regularity conditions:

$$\frac{d^s}{dz^s}\left\{G_p(z) - (\phi')^2 B_0(z) \sum_{k=0}^{\mu} \Gamma_{kp}\phi^k - J_p(z)\right\} = 0 \quad \text{at } z = z_j, \tag{3.3}$$

$$s = 0, \cdots, m_j - 1, \quad j = 0, \cdots, N, \quad p = 0, 1, \cdots.$$

Equation (3.3) can be viewed as μ linear equations on the $\mu + 1$ constants Γ_{kp}, $k = 0, \cdots, \mu$. Thus since the rank of the coefficient matrix cannot exceed μ, special choices of the constants Γ_{kp} can be made to ensure the regularity of $H_p(z)$ at every point of $\mathscr{D}$, provided $G_p(z)$ is regular there also. We note from (2.6) and (2.7) that $(\phi')^2$ and $B_0(z)$ do not vanish at the turning points.

It is clear that at least one of these constants can be chosen arbitrarily. Such a choice will be aimed at simplifying (1.3) and making it easier to solve. The most obvious way is to aim at reducing the degree of the polynomial in the R.H.S. of (1.3).

4. Examples. We now consider some specific examples. First we look at the case for which $R_0(z)$ has no zero in D. Thus $\mu = 0$. We choose $\gamma_0 = \gamma_{00} = 1$. Since no regularity conditions on $B_p(z)$, $C_p(z)$ and $H_p(z)$ have to be satisfied we choose $\gamma_{0p} = 0, p = 1, 2, \cdots$, and $\Gamma_0 = 0$. Consequently the comparison equation (1.3) for $\mu = 0$ becomes

$$V''(\xi) + V(\xi) = 0 \tag{4.1}$$

with the solution

$$V(\xi) = \sin(\xi - \xi_0). \tag{4.2}$$

From [2] we obtain the result

$$\theta(z) = \frac{1}{2}\int_{z_0}^{z} R_0^{-1/2}(s) R_1(s)\, ds. \tag{4.3}$$

Furthermore (2.6) yields $\phi'(z)$ and $\phi(z)$:

$$(\phi')^2 = R_0(z), \tag{4.4}$$

$$\phi(z) = \int_{z_0}^{z} R_0^{1/2}(s)\, ds, \tag{4.5}$$

where z_0 is an arbitrary point in D, and where we have chosen the constant $\phi_0 = \phi(z_0) = 0$. Hence in (4.2) we also set $\xi_0 = 0$. On using (4.4), (2.7) and (2.8) give

$$B_0(z) = (\cos\theta(z))/R_0^{1/4}(z), \qquad C_0(z) = (\sin\theta(z))/R_0^{1/4}(z). \tag{4.6}$$

Finally we obtain $H_0(z)$ from (3.1) with $p = 0$ as follows:

$$H_0(z) = G_0(z)/R_0(z). \tag{4.7}$$

Thus we have the following first approximation for (1.2) in this case:

$$u(z, \lambda) = R_0^{-1/4} \sin[\lambda\phi(z) + \theta(z)] + G_0(z)/R_0(z) + O(\lambda^{-1}). \tag{4.8}$$

To obtain more terms in the expansion we only need to compute $B_p(z)$, $C_p(z)$ from (2.7) and (2.8) (or the more explicit forms given in [2]), and $H_p(z)$ from (3.1). This result corresponds to Olver's [3, pp. 386–390] when $R_1(z) = R_3(z) = \cdots = 0$.

Next we look at the case in which $R_0(z_0) = 0$, but $R'(z_0) \neq 0$ for some z_0 in D. Then $\mu = 1$ and $N = 0$. Equation (2.6) reduces to

$$(\phi')^2 = \frac{R_0(z)}{-\phi(z)} \tag{4.9}$$

where we have chosen the free constants $\phi_0 = \phi(z_0) = 0$ and $\gamma_1 = \gamma_{10} = -1$. The choice $\gamma_0 = 0$ implies $\gamma_{00} = 0$. Integrating (4.9) we obtain

$$\phi(z) = \left\{\frac{3}{2}\int_{z_0}^{z} [-R_0(s)]^{1/2}\, ds\right\}^{2/3}. \tag{4.10}$$

Since no regularity conditions on $B_p(z)$, $C_p(z)$ need to be satisfied we set $\gamma_{1p} = 0$, $p = 1, 2, \cdots$, and hence $\theta(z)$ is also given by (4.3) (see [2]). Using (4.9) in (2.7) and (2.8) for this case we obtain (for $p = 0$)

$$B_0(z) = [-\phi(z)/R_0(z)]^{1/4} \cos\theta(z), \qquad C_0(z) = [\phi(z)R_0(z)]^{-1/4} \sin\theta(z). \tag{4.11}$$

We now seek to determine $H_0(z)$ and choose $\Gamma_1 = 0$. The comparison equation (1.3) for $\mu = 1$ becomes

$$V''(\xi) - \xi V(\xi) = \Gamma_0. \tag{4.12}$$

On using (1.4) and (1.5) for $\mu = 1$, $V_1(\xi) = \mathrm{Ai}(\xi)$, $V_2(\xi) = \mathrm{Bi}(\xi)$ the solution to (4.12) can be seen to have the form

$$V(\xi) = c_1 \mathrm{Ai}(\xi) + c_2 \mathrm{Bi}(\xi) + \Gamma_0(-\xi)^{1/2} S_{0,1/3}[\tfrac{2}{3}(-\xi)^{3/2}] \tag{4.13}$$

where $\mathrm{Ai}(\xi)$, $\mathrm{Bi}(\xi)$ are the Airy functions and $S_{0,1/3}(\eta)$ is the Lommel function (see Watson [8]). Now (3.1) for $p=0$ yields

$$H_0(z)=[G_0(z)-(\phi')^2B_0(z)\Gamma_{00}]/R_0(z). \tag{4.14}$$

Regularity of $H_0(z)$ at $z=z_0$ is achieved if we choose

$$\Gamma_{00}=G_0(z_0)/([\phi'(z_0)]^2B_0(z_0)), \tag{4.15}$$

having used (3.3) for $s=0, j=0$ and $p=0$. Thus a first approximation to (1.2) for this case is given by

$$\begin{aligned} u(z,\lambda)&=B_0(z)V[\lambda^{2/3}\phi(z)]+\lambda^{-1/3}C_0(z)V'[\lambda^{2/3}\phi(z)]\\ &\quad+G_0(z)/R_0(z)-[\phi'(z)/\phi'(z_0)]^2[B_0(z)/B_0(z_0)]\\ &\quad\cdot[G_0(z_0)/R_0(z)]\\ &\quad+O(\lambda^{-1}), \end{aligned} \tag{4.16}$$

where $V(\xi)$ is given by (4.13), $\phi(z)$ by (4.10) and $B_0(z)$, $C_0(z)$ are given by (4.11). To obtain more terms we again compute $B_p(z)$ and $C_p(z)$ from [2]. We also use (3.3) with $s=0$ to obtain Γ_{0p} and finally use (3.1) to obtain $H_p(z)$. This case was considered by Clark [4] and Nayfeh [6, p. 356]. If $R_1(z)=0$, $\theta(z)=0$ and hence $C_0(z)=0$. Also if $R_p(z)=0$, $p=1,3,\cdots$, then our result can be shown to be equivalent to Olver's [3].

When $R_0(z_0)=R'(z_0)=0$, but $R''(z_0)\neq 0$, then we have a second order turning point at $z=z_0$. In this case $m_0=\mu=2$ and $N=0$. We choose $\gamma_2=\gamma_{20}=-\frac{1}{4}$ and $\gamma_1=0$. The other constant γ_0 now has to be employed to ensure regularity of $B_p(z)$, $C_p(z)$ at $z=z_0$. Consequently, on choosing $\phi_0=0$, (2.6) becomes

$$(\phi')^2=\frac{-4R_0(z)}{\phi^2(z)}, \tag{4.17}$$

and integrating this equation we have

$$\phi(z)=2\left\{\int_{z_0}^{z}[-R_0(s)]^{1/2}\,ds\right\}^{1/2}, \tag{4.18}$$

since $\phi_0=0$, $\gamma_{00}=0$. In [2], $\theta(z)$ for this case was found to be

$$\theta(z)=-\frac{1}{2}\int_{z_0}^{z}\{R_0^{-1/2}(s)R_1(s)+4\gamma_{01}R_0^{1/2}[\phi(s)]^{-2}\}\,ds, \tag{4.19}$$

where

$$\gamma_{01}=[-2R_0''(z_0)]^{-1/2}R_1(z_0). \tag{4.20}$$

We have, from (2.7) and (2.8),

$$\begin{aligned} B_0(z)&=\{[\phi(z)]^2/[-4R_0(z)]\}^{1/4}\cos\theta(z),\\ C_0(z)&=\{-4/([\phi(z)]^2R_0(z))\}^{1/4}\sin\theta(z). \end{aligned} \tag{4.21}$$

Next we set $\Gamma_2=0$. Equation (1.3) becomes

$$V''(\xi)+(\gamma_0-\tfrac{1}{4}\xi^2)V=\Gamma_0+\Gamma_1\xi \tag{4.22}$$

with the homogeneous solutions

(4.23)
$$V_1(\xi) = D_{\gamma_0 - 1/2}(\xi), \quad V_2(\xi) = \frac{1}{\pi}\Gamma(\tfrac{1}{2} + \gamma_0)[\sin \pi\gamma_0 D_{\gamma_0 - 1/2}(\xi) + D_{\gamma_0 - 1/2}(-\xi)],$$

where $D_{\gamma_0 - 1/2}(\xi)$ is the parabolic cylinder function. Thus $V(\xi)$ is given by (1.4) with

(4.24)
$$Q_1(\xi) = -\sqrt{\pi/4}\int^{\xi} (\Gamma_0 + \Gamma_1\eta) V_2(\eta) d\eta + C_1,$$
$$Q_2(\xi) = \sqrt{\pi/4}\int^{\xi} (\Gamma_0 + \Gamma_1\eta) V_1(\eta)\, d\eta + C_2.$$

Defining

(4.25)
$$\Omega(z) = [-4R_0(z)/\phi^2(z)]^{3/4},$$

(3.1) for $p = 0$ yields

(4.26)
$$H_0(z) = \{G_0(z) - \Omega(z)\cos\theta(z)[\Gamma_{00} + \Gamma_{10}\phi(z)]\}/R_0(z),$$

where we have used (4.17) and (4.21) for $\phi'(z)$ and $B_0(z)$. Condition (3.3) for $p = 0$, $s = 0, 1$, yields the following:

(4.27)
$$\Gamma_{00} = G_0(z_0)/\Omega(z_0),$$

(4.28)
$$\Gamma_{10} = G_0'(z_0)/\Omega^{5/3}(z_0) - G_0(z_0)\Omega'(z_0)/\Omega^{8/3}(z_0),$$

having used the fact that $\theta(z_0) = \theta'(z_0) = 0$ for this case. Thus we have

(4.29)
$$u(z, \lambda) = B_0(z) V[\lambda^{1/2}\phi(z)] + \lambda^{-1/2} C_0(z) V'[\lambda^{1/2}\phi(z)] + \{G_0(z) - \Omega(z)\cos\theta(z)[\Gamma_{00} + \Gamma_{10}\phi(z)]\}/R_0(z) + O(\lambda^{-1}),$$

where Γ_{00} and Γ_{01} are given by (4.27) and (4.28), $B_0(z)$, $C_0(z)$ are defined in (4,21). Previous treatment of this problem, instead of (1.3)–(1.5), employed $V(\xi)$ in (1.2) given by

$$V(\xi) = \Gamma_0(\lambda) T_0(\xi) + \Gamma_1(\lambda)\xi T_1(\xi),$$

where $T_0(\xi)$ and $T_1(\xi)$ satisfy

$$T_0'' + (\gamma_0 - \xi^2) T_0 = 1 \quad \text{and} \quad T_1'' + (\gamma_0 - \xi^2) T_1 = \xi.$$

Nayfeh in [6, pp. 356–358] and McGuinness in [7] used this representation. From [7] we see that such a representation required four coefficient functions in the expansion. However, that representation is easily seen to be completely equivalent to ours by identifying the four coefficient functions as: $\Gamma_0 B(z, \lambda)$, $\Gamma_1[\xi B(z, \lambda) + \lambda^{-1/2} C(z, \lambda)]$, $\lambda^{-1/2}\Gamma_0 C(z, \lambda)$ and $\lambda^{-1/2}\Gamma_1 \xi C(z, \lambda)$. We note that for $\mu = 2$, Γ_0 and $\xi\Gamma_1$ are both $O(\lambda)$. Setting $R_0 = -z^2$, $R_1(z) = -p(z)$ and $R_j(z) = -q_{j-2}(z)$, $j = 2, 3, \cdots$, McGuinness's results are recovered.

Finally we look at the case $R_0(z_0) = R_0(z_1) = 0$ but $R_0'(z_0) \neq 0$, $R_0'(z_1) \neq 0$. Here again $\mu = 2$, but $N = 1$. With the same choices of γ_1, γ_2 and Γ_2 as in the preceding case for $N = 0$, our comparison equation has the same form as (4.22)

except that now γ_0, Γ_0 and Γ_1 are different. Choosing $\phi_1 = -\phi_0$ as in [2], (2.6) yields $\gamma_{00} = \frac{1}{4}\phi_0^2$ and

$$[\phi'(z)]^2 = 4R_0(z)/[\phi_0^2 - \phi^2(z)] \tag{4.30}$$

By integrating (4.3) Lynn and Keller determined $\phi(z)$ as follows:

$$-\frac{1}{2}\phi_0^2\left[\frac{1}{2}\pi - \sin^{-1}\frac{\phi}{\phi_0} + 1 - \left(\frac{\phi}{\phi_0}\right)^2\right] = 2\int_{z_0}^{z} [-R_0(s)]^{1/2}\, ds, \tag{4.31}$$

and

$$\phi_0 = 2\left\{\frac{1}{\pi}\int_{z_0}^{z_1} [-R_0(s)]^{1/2}\, ds\right\}^{1/2}. \tag{4.32}$$

Also

$$\theta(z) = -\frac{1}{2}\int_{z_0}^{z}\left[\frac{R_1(s)}{R_0^{1/2}(s)} - \frac{4\gamma_{01}R_0^{1/2}(s)}{\phi_0^2 - \phi^2(s)}\right] ds \tag{4.33}$$

where the regularity condition at $z = z_1$ determines γ_{01} as

$$\gamma_{01} = \frac{\int_{z_0}^{z_1} R_0^{-1/2}(s)R_1(s)\, ds}{\int_{z_0}^{z_1}[4R_0^{-1/2}(s)/(\phi_0^2 - \phi^2(s))]\, ds}. \tag{4.34}$$

Using (4.30), $B_0(z)$, $C_0(z)$ are obtained from (2.7), (2.8):

$$\begin{aligned} B_0(z) &= \left(\frac{\phi_0^2 - \phi^2(z)}{4R_0(z)}\right)^{1/4} \cos\theta(z); \\ C_0(z) &= \left(\frac{4}{R_0(z)[\phi_0^2 - \phi^2(z)]}\right)^{1/4} \sin\theta(z). \end{aligned} \tag{4.35}$$

In this case (3.1) also yields (4.26) for $H_0(z)$ except that we replace $\Omega(z)$ in that equation by

$$\hat{\Omega}(z) = \left(\frac{4R_0(z)}{\phi_0^2 - \phi^2(z)}\right)^{3/4}, \tag{4.36}$$

with (3.3) for $p = 0$, $s = 0$, $j = 0$, 1, giving

$$\Gamma_{00} = (G_0(z_0)\hat{\Omega}_1 + G_0(z_1)\hat{\Omega}_0)/(2\hat{\Omega}_0\hat{\Omega}_1), \tag{4.37}$$

and

$$\Gamma_{01} = (G_0(z_0)\hat{\Omega}_1 - G_0(z_1)\hat{\Omega}_0)/(2\phi_0\hat{\Omega}_0\hat{\Omega}_1). \tag{4.38}$$

Here $\hat{\Omega}_0 = \hat{\Omega}(z_0)$, $\hat{\Omega}_1 = \hat{\Omega}(z_1)$ and also $\theta(z_0) = \theta(z_1) = 0$. Thus $u(z, \lambda)$ is given by (4.29) with $B_0(z)$, $C_0(z)$ given by (4.35), $\phi(z)$ by (4.31), Γ_{00} and Γ_{01} by (4.37) and (4.38) with $\Omega(z)$ replaced by $\hat{\Omega}(z)$. The author is not aware of any previous treatment of this case.

It is interesting to see that when z_0 converges to z_1, the result for two first order turning points goes over into that for one second order turning point. This fact was demonstrated in [2] for the linear homogeneous problem. To show this for the nonhomogeneous case we only need to show that Γ_{00} and Γ_{01} in the two

cases become identical when $z_0 \to z_1$. Since $\phi_0 = 0$ when $z_0 = z_1$ we observe that $\hat{\Omega}(z) \to \Omega(z)$. Hence (4.37) goes over into (4.27). To show that (4.38) becomes (4.28) we only need to see that

$$\lim_{z_1 \to z_0} \left[\frac{G_0(z_0)\hat{\Omega}(z_1) - G_0(z_1)\hat{\Omega}(z_0)}{-\phi(z_1)} \right] = \frac{G'(z_0)\Omega(z_0) - G_0(z_0)\Omega'(z_0)}{\Omega^{2/3}(z_0)},$$

having used the fact that $\phi'(z_0) = \Omega^{2/3}(z_0)$ from (4.17) and (4.25).

Conclusion. The development here is a generalization of the method of Clark [4] and McGuinness [7] much in the same way that Lynn and Keller [2] represent a generalization of the method of Langer [1].

Equation (1.1) occurs in the analysis of the stability of toroidal shells (Clark [9], Tumarkin [10]), in the analysis of thin elastic shells (Clark [11]) and in heat conduction in hollow cylinder (Holstein [12]).

Acknowledgment. The author wishes to acknowledge with gratitude some helpful criticisms by Professors Joseph B. Keller, F. W. J. Olver and W. Wasow. The referee's comments are also acknowledged with thanks.

REFERENCES

[1] R. E. LANGER, *The asymptotic solutions of ordinary linear differential equations of the second order, with special reference to a turning point*, Trans. Amer. Math. Soc., 67 (1949), pp. 461–490.

[2] J. B. KELLER AND R. Y. S. LYNN, *Uniform asymptotic solutions of second order linear ordinary differential equations with turning points*, Comm. Pure Appl. Math., 23 (1970), pp. 379–408.

[3] F. W. J. OLVER, *Asymptotics and Special Functions*, Academic Press, New York, 1974.

[4] R. A. CLARK, *Asymptotic solutions of a nonhomogeneous differential equation with a turning point*, Arch. Rational Mech. Anal., 12 (1963), pp. 34–51.

[5] C. R. STEELE, *On the asymptotic solution of nonhomogeneous ordinary differential equations with a large parameter*, Quart. Appl. Math. 23 (1965), pp. 193–201.

[6] A. H. NAYFEH, *Perturbation Methods*, John Wiley, New York, 1973.

[7] D. J. MCGUINNESS, *Nonhomogeneous differential equation with a second-order turning point.* J. Mathematical Phys., 7 (1966), pp. 1030–1037.

[8] G. N. WATSON, *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, London, 1958.

[9] R. A. CLARK, *Asymptotic solutions of toroidal shell problems*, Quart. Appl. Math., 16 (1958), pp. 47–60.

[10] S. A. TUMARKIN, *Asymptotic solution of a linear homogeneous second order differential equation with a transition point and its application to the computations of toroidal shells and propellor blades*, Prikl. Mat. Meh. = J. Appl. Math. Mech., 23 (1959), pp. 1549–1565.

[11] R. A. CLARK, *Asymptotic solutions of elastic shell problems*, Asymptotic Solutions of Differential Equations and Their Applications, C. H. Wilcox, ed., John Wiley, New York, 1964, pp. 185–209.

[12] V. H. HOLSTEIN, *Über die ausere and innere Reibungsschicht bei Storungen laminarer Stromungen*, Z. Angew. Math. Mech., 30 (1950), pp. 25–49.

SIAM J. MATH. ANAL.
Vol. 8, No. 4, August 1977

CAUCHY'S PROBLEM FOR SYSTEMS OF FIRST ORDER ANALYTIC ELLIPTIC EQUATIONS IN THE PLANE*

CHUNG-LING YU†

Abstract. Let a, b, c, d, f, g be analytic functions of two real variables x, y in the $z = x + iy$ plane. Consider the elliptic system:

$$\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} = au + bv + f, \qquad \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} = cu + dy + g.$$

The following topics connected with the Cauchy problem of the abovementioned system will be investigated: (i) The necessary and sufficient conditions for the existence of the global solutions. (ii) The estimate of the "stability" of the solutions. (iii) The method for constructing the global solution. (iv) The construction of the approximate solution with the imprecise data. (v) The generalizations to the quasi-linear analytic elliptic system:

$$\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} = f_1(u, v, x, y) \quad \text{and} \quad \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} = f_2(u, v, x, y).$$

1. Introduction. The aim of this paper is to study the Cauchy problem for the first order linear elliptic equations (in normal form)

$$\begin{aligned} &\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} = au + bv + f, \\ &\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} = cu + dv + g \end{aligned} \tag{1.1}$$

and first order quasi-linear elliptic equations (in normal form)

$$\begin{aligned} &\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} = f_1(u, v, x, y), \\ &\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} = f_2(u, v, x, y) \end{aligned} \tag{1.2}$$

where $a, b, c, d, f, g, f_1, f_2$ are analytic functions of their arguments.

The following areas will be investigated: (i) The necessary and sufficient conditions for the existence of global solutions for the Cauchy problem of (1.1). (ii) The methods for constructing the global solutions of the problem of (1.1) and the solution of (1.2). (iii) The "stability" for the problem of (1.1) and (1.2). (iv) The method for constructing the approximate solutions of (1.1) with imprecise data.

The first part of this paper will deal with the problem with analytic data; the second part will deal with the problem with nonanalytic data. For general reference, the reader is referred to the two survey articles by Payne [16], [17], and the books of Gilbert [7], [8], Wendland and Haack [23], and of Vekua [21], [22].

Some results in this paper, especially the necessary and sufficient conditions for existence and the methods for constructing the solution, may be extended to

* Received by the editors November 2, 1973, and in final revised form April 26, 1976.

† Department of Mathematics, Florida State University, Tallahassee, Florida. Now at Faculty of Engineering, Benghazi University, Benghazi, Libya.

equation (1.1) with nonanalytic coefficients [26] and to higher order analytic elliptic equations [27].

2. Definitions and notations. Let $f(z_1, \cdots, z_n)$ be a holomorphic function of the complex variables $z_1, \cdots, z_n$ in a domain $G \subset C^n$. We can associate with $f(z_1, \cdots, z_n)$ another function, defined in the conjugate domain $\bar{G} = \{(\zeta_1, \cdots, \zeta_n); (\zeta_1, \cdots, \zeta_n) \in C^n; (\bar{\zeta}_1, \cdots, \bar{\zeta}_n) \in G\}$ by the formula

$$f^*(\zeta_1, \cdots, \zeta_n) = \overline{f(\bar{\zeta}_1, \cdots, \bar{\zeta}_n)}, (\zeta_1, \cdots, \zeta_n) \in \bar{G}, \tag{2.1}$$

We call $f^*(\zeta_1, \cdots, \zeta_n)$ the * *conjugate* function to $f(z_1, \cdots, z_n)$. It is easily seen that $f^*(\zeta_1, \cdots, \zeta_n)$ is a holomorphic function of $\zeta_1, \cdots, \zeta_n$ in $\bar{G}$.

By introducing the complex notation

$$\frac{\partial}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right), \tag{2.2}$$

we can write (1.1) in the complex form

$$\frac{\partial}{\partial \bar{z}} w = Aw + B\bar{w} + F \tag{2.3}$$

and (1.2) in the form

$$\frac{\partial}{\partial \bar{z}} w = H(w, \bar{w}, x+iy, x-iy) \tag{2.4}$$

where

$$w = u + iv, \qquad A = \tfrac{1}{4}(a + d + ic - ib),$$
$$B = \tfrac{1}{4}(a - d + ic + ib), \qquad F = \tfrac{1}{2}(f + g),$$
$$H = \tfrac{1}{2}(f_1 + if_2).$$

If we continue $a, b, c, d, f, g, f_1, f_2$ into the space of two complex variables, we obtain A, B, F, H as holomorphic functions of the two complex variables

$$z = x + iy, \qquad \zeta = x - iy.$$

From now on we will assume D to be a simply connected domain in the $z = x + iy$ plane whose boundary ∂D is supposed to contain a segment σ, $\sigma = \{x: a < x < b\}$. We assume σ to contain the origin as an interior point. We also assume $A(z, \zeta)$, $B(z, \zeta)$, $F(z, \zeta)$, $H(\xi_1, \xi_2, z, \zeta)$ are holomorphic for $z, \zeta \in D \cup \sigma \cup \bar{D}$, $|\xi_1|, |\xi_2| < M$.

The Cauchy problem for the equation (1.1) (or (1.2)) can be stated as follows: Find a solution $w(x, y)$ of the equation (1.1) (or (1.2)) satisfying the following condition:

$$w(x, 0) = \rho(x), \qquad x \in \sigma, \tag{2.5}$$

where $\rho(x)$ is defined for $x \in \sigma$. Note there is no loss in generality in assuming the Cauchy data to be prescribed on the x axis.

PART I. CAUCHY PROBLEM WITH ANALYTIC DATA

It is well known that there have been only a few constructive methods for the numerical integration of the analytic elliptic Cauchy problem.

Garabedian [6] has introduced a method which is suitable for numerical integration by converting the analytic elliptic problem to a hyperbolic Cauchy problem. Henrici [10] has used conjugate coordinates and the Riemann function to obtain an explicit representation for the solution of the second order linear equation in terms of its Cauchy data. He thus obtained clear information about the domain of existence and the continuous dependence of the solutions on the Cauchy data in a certain topology. In [3], [4], [5] Colton (see also Gilbert [8]) reduces the analytic Cauchy problem for certain higher order almost linear analytic elliptic equations to finding a fixed point of a contraction mapping. The problem then can be solved by classical iterative procedures.

In this part, we shall make use of * conjugate functions (see § 2) and Vekua's integral representation [21] to reduce the Cauchy problem for (1.1) to the problem of finding a solution to a linear complex Volterra integral equation, and thus obtain the integral representation, continuous dependence and the domain of existence for the solution. For equations (1.2), we again use the * conjugate function to reduce the problem to a nonlinear integral equation, which then can be solved by the method of successive approximations.

Our methods and results may be considered as continuations and generalizations of those of Henrici [10], Colton [3] and Aziz, Gilbert and Howard [1], [8].

3. Analytic Cauchy problem for (1.1). The integral representation in Lemma 3.1 below for the solution (2.3) in a simply connected $G \subset D \cup \sigma \cup \bar{D}$ has been established by Vekua [21], and later extended to the boundary ∂G of G by Yu [25].

LEMMA 3.1. *Every C^1 solution $w(z)$ of differential equation* (2.3) *in G has the integral representation*

$$w(z)=\left\{\phi(z)+\int_{z_0}^{z}\Gamma_1(z,\bar{z},t,\zeta_0)\phi(t)\,dt + \int_{\zeta_0}^{\bar{z}}\Gamma_2(z,\bar{z},z_0,\tau)\phi^*(\tau)\,d\tau+U_0(z,\bar{z})\right\}\exp\int_{\zeta_1}^{\bar{z}}A(z,t)\,dt, \tag{3.1}$$

where (z_0,ζ_0) is a fixed point in $(G,\bar{G})$, $\phi(z)$ is a holomorphic function in G, $U_0(z,\zeta)$ is a holomorphic function for $z,\ \zeta\in D\cup\sigma\cup\bar{D}$, and $\Gamma_1(z,\zeta,t,\tau)$, $\Gamma_2(z,\zeta,t,\tau)$ are two holomorphic functions of the four variables $z,\ t,\ \zeta,\ \tau\in D\cup\sigma\cup\bar{D}$. In (3.1),

$$U_0(z,\zeta)=\int_{\zeta_0}^{\zeta}F_0(z,\tau)\,d\tau+\int_{\zeta_0}^{\zeta}d\tau\int_{z_0}^{z}\Gamma_1(z,\zeta,t,\tau)F_0(t,\tau)\,dt + \int_{\zeta_0}^{\zeta}d\tau\int_{z_0}^{z}\Gamma_2(z,\zeta,t,\tau)F_0^*(\tau,t)\,dt, \tag{3.2}$$

$$\Gamma_1(z,\zeta,t,\tau)=\int_{\tau}^{\zeta}\Gamma(z,\zeta,t,\eta)\,d\eta, \tag{3.3}$$

$$\Gamma_2(z, \zeta, t, \tau) = C(z, \tau) + \int_t^z C(\xi, \tau)\Gamma_1(z, \zeta, \xi, \tau)\, d\xi \tag{3.4}$$

where Γ *is the unique holomorphic solution of the Volterra integral equation*

$$\Gamma(z, \zeta, t, \tau) = C(z, \tau)C^*(\tau, t) + \int_\tau^\zeta d\eta \int_t^z C(\xi, \tau)C^*(\tau, t)\Gamma(z, \zeta, \xi, \eta)\, d\xi, \tag{3.5}$$

for $z, \zeta, t, \tau \in D \cup \sigma \cup \bar{D}$, *and*

$$F_0(z, \zeta) = F(z, \zeta) \exp\left[-\int_{\zeta_1}^{\zeta} A(z, t)\, dt\right].$$

Conversely, the function $w(z)$ *which is given by formula* (3.1) *is a solution of differential equation* (2.3) *in* G.

Moreover, $w(z)$ *is continuous in* $G \cup \partial G$ *if and only if* $\phi(z)$ *is continuous there.*

Theorem 3.1 below provides a constructive method for obtaining a unique solution of the Cauchy problem for (2.3) and its domain of existence.

THEOREM 3.1. *Let the Cauchy data* $\rho(z)$ *be holomorphic throughout* $D \cup \sigma \cup \bar{D}$. *Then the solution of the Cauchy problem* (2.3), (2.5) *is analytic for* $(x, y) \in D \cup \sigma \cup \bar{D}$, *has the integral representation* (3.13) *below and can be constructed by the method of successive approximation through* (3.11), (3.12), *and* (3.13).

Proof. Let us consider the integral equation

$$\begin{aligned}\rho(z) = \Big\{\phi(z) &+ \int_0^z \Gamma_1(z, z, t, 0)\phi(t)\, dt \\ &+ \int_0^z \Gamma_2(z, z, 0, \tau)\phi^*(\tau)\, d\tau + U_0(z, z)\Big\} \exp \int_{\zeta_1}^z A(z, t)\, dt\end{aligned} \tag{3.6}$$

for the unknown function $\phi(z)$, where $\Gamma_1(z, \zeta, t, 0)$, $\Gamma_2(z, \zeta, 0, \tau)$, and $U_0(z, \zeta)$ are functions given in Lemma 3.1 (for convenience, we set $z_0 = \zeta_0 = 0$), and recall that they are holomorphic for $z, \zeta, t, \tau \in D \cup \sigma \cup \bar{D}$.

Set

$$h(z) = \rho(z) \exp\left[-\int_\zeta^z A(z, t)\, dt\right] - U_0(z, \zeta), \tag{3.7}$$

$$G_1(z, t) = \Gamma_1(z, z, t, 0), \tag{3.8}$$

$$G_2(z, t) = \Gamma_2(z, z, 0, \tau). \tag{3.9}$$

It then follows immediately from the definition of the * conjugate function that we have

$$\begin{aligned}h^*(z) = \phi^*(z) &+ \int_0^z G_1^*(z, t)\phi^*(t)\, dt \\ &+ \int_0^z G_2^*(z, t)\phi(t)\, dt.\end{aligned} \tag{3.10}$$

Using matrix notation, (3.6) and (3.10) are equivalent to the following Volterra integral equation:

$$H(z) = \Phi(z) + \int_0^z G(z, t)\Phi(t)\, dt \tag{3.11}$$

where

$$H(z)=\begin{pmatrix} h(z) \\ h^*(z) \end{pmatrix}, \qquad \Phi(z)=\begin{pmatrix} \phi(z) \\ \phi^*(z) \end{pmatrix},$$

$$G(z,t)=\begin{pmatrix} G_1(z,t), & G_2(z,t) \\ G_2^*(z,t), & G_1^*(z,t) \end{pmatrix}.$$

Therefore, the existence and the uniqueness of the solution (2.6) follow from the existence and uniqueness of solutions to the Volterra integral equation (3.11), and the solution $\phi(z)$ has the following integral representation:

$$\phi(z)=h(z)+\int_0^z I_1(z,t)h(t)\,dt+\int_0^z I_2(z,t)h^*(t)\,dt \tag{3.12}$$

where the kernel function $I_1(z,t)$, $I_2(z,t)$ are holomorphic for z, $t\in D\cup\sigma\cup\bar{D}$, depend only on G_1 and G_2, and can be constructed explicitly by standard procedures (see Vekua [21]).

Define the function

$$\begin{aligned} w(z)=&\left\{\phi(z)+\int_0^z \Gamma_1(z,\bar{z},t,0)\phi(t)\,dt\right. \\ &\left.+\int_0^z \Gamma_2(z,z,0,\tau)\phi(\tau)\,d\tau+U_0(z,z)\right\} \\ &\cdot\exp\int_{\zeta_1}^{\bar{z}} A(z,t)\,dt \end{aligned} \tag{3.13}$$

where $\phi(z)$ is given by (3.12).

By Lemma 3.1, the function $w(z)$ is a solution of (2.3) in $D\cup\sigma\cup\bar{D}$, and from (3.13) and (3.6), $w(x)=\rho(x)$. Hence $w(z)$ is the desired solution of the Cauchy problem.

Theorem 3.1 admits of the following converse:

THEOREM 3.2. *Let $w(z)$ be a solution of* (1.1) *in $D\cup\sigma\cup\bar{D}$. Then the function*

$$w(x)=\rho(x), \qquad x\in\sigma, \tag{3.14}$$

is analytic on σ and can be continued analytically into the whole $D\cup\sigma\cup\bar{D}$; i.e., $\rho(z)$ is holomorphic for $z\in D\cup\sigma\cup\bar{D}$.

Proof. It follows immediately from (3.1).

In view of Theorem 3.1 and Theorem 3.2, we may associate with every holomorphic function in $D\cup\sigma\cup\bar{D}$ a solution $w(z)$ of (1.1) in $D\cup\sigma\cup\bar{D}$. We write accordingly

$$w=T\rho.$$

Let $W(D\cup\sigma\cup\bar{D})$ be the space of C^1 solutions of the homogeneous equation

$$\frac{\partial w}{\partial\bar{z}}=Aw+B\bar{w}$$

over the real field in $D \cup \sigma \cup \bar{D}$, with the topology of uniform convergence on compact subsets. Also let $H(D \cup \sigma \cup \bar{D})$ be the vector space of holomorphic functions in $D \cup \sigma \cup \bar{D}$ over the real field, with the topology of uniform convergence on compact subsets. Then we have the following result connected with the continuous dependence on the Cauchy data.

THEOREM 3.3. *The transformation T is an isomorphism from $H(D \cup \sigma \cup \bar{D})$ onto $W(D \cup \sigma \cup \bar{D})$.*

Proof. By (3.12) and (3.13), it is clear that T is 1–1, onto, linear and continuous from $H(D \cup \sigma \cup \bar{D})$ to $W(D \cup \sigma \cup \bar{D})$. Therefore, by the interior mapping theorem T is an isomorphism between $H(D \cup \sigma \cup \bar{D})$ and $W(D \cup \sigma \cup \bar{D})$.

4. Analytic Cauchy problem for (1.2). We now seek a solution $w(z)$ for (1.2) which satisfies the Cauchy data (2.5).

Since the solution of the analytic elliptic equation is again analytic, we can transform the Cauchy problem (2.4), (2.5) into complex form

$$W_\zeta = H(W, W^*, z, \zeta), \tag{4.1}$$

$$W(z, \zeta)|_{\zeta = z} = \rho(z) \tag{4.2}$$

where $W(z, \zeta) = w((z+\zeta)/2, (z-\zeta)/2i)$, $z = x + iy$, $\zeta = x - iy$, and W^* is the * conjugate function of W.

We assume that $H(\xi_1, \xi_2, z, t)$ is holomorphic for $|\xi_1|\ |\xi_2| < M$, $z,\ t \in D \cup \sigma \cup \bar{D}$, and that $|\rho(z)| < M$ for $z \in D \cup \sigma \cup \bar{D}$. We then have

$$W(z, \zeta) = f(z) + \int_0^\zeta H(W(z, t), W^*(t, z), z, t)\, dt \tag{4.3}$$

where $f(z)$ is a holomorphic function for z and W^* is the * conjugate function of W. Conversely, if $W(z, \zeta)$ satisfies (4.3) then $W(z, \bar{z})$ is a solution of (1.2).

We also find that the initial condition (4.2) is equivalent to

$$\rho(z) = f(z) + \int_0^z H(W(z, t), W^*(t, z), z, t)\, dt. \tag{4.4}$$

We conclude that the Cauchy problem (2.4), (2.5) is equivalent to the following nonlinear complex integral equation

$$\begin{aligned} W(z, \zeta) = \rho(z) &- \int_0^z H(W(z, t), W^*(t, z), z, t)\, dt \\ &+ \int_0^\zeta H(W(z, t), W^*(t, z), z, t)\, dt. \end{aligned} \tag{4.5}$$

Define

$$\begin{aligned} (TW)(z, \zeta) = \rho(z) &- \int_0^z H(W(z, t), W^*(t, z), z, t)\, dt \\ &+ \int_0^\zeta H(W(z, t), W^*(t, z), z, t)\, dt. \end{aligned} \tag{4.6}$$

Let $|\rho(z)|, |W(z,\zeta)|<M$, for $|z|, |\zeta|<r$; then

$$
\begin{aligned}
|TW| &\leqq |\rho(z)| + \left|\int_0^{\zeta} H(W(z,t), W^*(t,z), z, t)\, dt\right| \\
&\quad + \left|\int_0^{\zeta} H(W(z,t), W^*(t,z), z, t)\, dt\right| \\
&\leqq \max|\rho(z)| + (\max|H|)|z| + (\max|H|)|\zeta| \\
&< M
\end{aligned} \tag{4.7}
$$

for r sufficiently small.

From Schwarz's lemma H satisfies a Lipschitz condition with respect to the first two arguments, i.e.,

$$
|H(\xi_1, \eta_1, z, t) - H(\xi_1, \eta_2, z, t)| \leqq C_0\{|\xi_1 - \xi_2| + |\eta_1 - \eta_2|\} \tag{4.8}
$$

where C_0 is a positive constant. Hence for $|W_1|, |W_2|<M$, $|z|, |\zeta|<r$,

$$
\begin{aligned}
|TW_1 - TW_2| &\leqq \left|\int_0^{\zeta} \{H(W_1, W_1^*, z, t) - H(W_2, W_2^*, z, t)\}\, dt\right| \\
&\quad + \left|\int_0^{z} \{H(W_1, W_1^*, z, t) - H(W_2, W_2^*, z, t)\}\, dt\right| \\
&\leqq 4C_0\Big(\sup_{|z|,|\zeta|<r} |W_1 - W_2|\Big) r.
\end{aligned} \tag{4.9}
$$

Therefore

$$
\sup_{|z|,|\zeta|<r} |TW_1 - TW_2| \leqq 4rC_0 \sup_{|z|,|\zeta|<r} |W_1 - W_2|. \tag{4.10}
$$

Let r be small enough such that $4C_0 r < 1$. Then in view of (4.10), the equation (4.5) can be solved by the method of successive approximation. We then have:

THEOREM 4.1. *If*

(i) $H(\xi_1, \xi_2, z, \zeta)$ *is holomorphic for* $z, \zeta \in D \cup \sigma \cup \bar{D}$, *and for* $|\xi_1|, |\xi_2|<M$, *and*

(ii) $\rho(z)$ *is holomorphic in* $D \cup \sigma \cup \bar{D}$ *and* $|\rho(z)|<M$,

then the Cauchy problem (2.4), (2.5) *has an unique analytic solution in* $|z| \leqq r$, *for* r *sufficiently small, and can be constructed by the method of successive approximation through* (4.5).

PART II. CAUCHY PROBLEM WITH NONANALYTIC DATA IN A GENERAL DOMAIN

The methods mentioned in Part I only work for analytic Cauchy data defined in a conformally symmetric domain (see Gilbert [7, p. 196], [8]). For nonanalytic Cauchy data, Hadmard [9] established a necessary and sufficient condition for the existence of a solution of the Laplace equation in a neighborhood of the initial line. Payne and Sather [18], and D. Sather and J. Sather [20] extended this result to a class of elliptic-parabolic equations. John [13] also studied the existence

theory of holomorphic functions, and suggested a method for constructing a global solution.

Carleman [2] introduced the famous two constants theorem which gives information on the continuous dependence of holomorphic functions on their boundary data. In particular, a class of uniformly bounded holomorphic functions is "stable". Work along this line was continued by John [11]–[14], Pucci [19], Lavrentiev [15], Payne [16] and many others (see references cited in Payne [17]).

In real physical problems, the Cauchy data are obtained by measurement and hence are not known precisely. Carleman [2] introduced an auxiliary function (the so-called Carleman function) to construct an approximate solution for holomorphic functions with given imprecise data and Lavrentiev [15] extended this idea to the Laplace equation in 3 variables.

We shall study the abovementioned topics for the equations (1.1) and (1.2). From now on in this part G will denote a bounded simply connected subdomain in $D \cup \sigma \cup \bar{D}$, Γ' is a part of boundary ∂G of G and $\Gamma'' = \partial G \backslash \Gamma'$.

5. "Stability" of the Cauchy problem for (1.1) in a bounded domain G. In this section an estimate of "stability" for the solution of (1.1) in a (not necessarily conformal symmetric) simply connected domain G is given.

Let $w(z)$ be a bounded solution of (1.1) in G such that

$$|w(z)| \leqq M, \qquad z \in G. \tag{5.1}$$

Let the values of $w(z)$ on Γ' be known, and suppose it is required to determined $w(z)$ in some part of G. Theorem 5.3 in this section will characterize the stability of the solution of the problem.

Carleman [2] proved the following so-called two-constant theorem characterizing the stability of a holomorphic function for the above mentioned problem.

THEOREM 5.1. *Let $f(z)$ be a bounded holomorphic function in G, and suppose that*

$$|f(z)| < M, \qquad z \in G,$$

and

$$|f(z)| < \varepsilon, \qquad z \in \Gamma'.$$

Then the inequality

$$|f(z)| < M^{(1-\omega(z))} \varepsilon^{\omega(z)}$$

will also hold where $\omega(z)$ is the harmonic measure of the curve Γ' with respect to the point z and the domain G.

Proof. See Lavrentiev [15].

Let

$$M_1 = \sup_{z,t \in G \cup \partial G} \{\Gamma_1(z, \bar{z}, t, \bar{z}_0), \Gamma_2(z, \bar{z}, z_0, t)\}, \tag{5.2}$$

$$M_2 = \sup_{z,t \in G \cup \partial G} 4M_1 \exp\{4M_1|z-t|\}, \tag{5.3}$$

$$M_3 = \sup_{z \in G \cup \partial G} \exp\left[-\int_{\zeta_1}^{\bar{z}} A(z, t)\, dt\right], \tag{5.4}$$

$$M_4 = \sup_{z \in G \cup \partial G} \left[\int_{\zeta_1}^{\bar{z}} A(z, t)\, dt \right]. \tag{5.5}$$

THEOREM 5.2. *If the solution $w(z)$ of* (2.3) (*with $F = 0$*) *in G is continuous in $G \cup \partial G$, bounded by a constant M in G and ε in Γ', i.e.*

$$|w(z)| < M, \qquad z \in G, \tag{5.6}$$

$$|w(z)| < \varepsilon, \qquad z \in \Gamma', \tag{5.7}$$

then the holomorphic function $\phi(z)$ in (3.1) *satisfies the following inequalities:*

$$|\phi(z)| < KM, \qquad z \in G, \tag{5.8}$$

$$|\phi(z)| < K\varepsilon, \qquad z \in \Gamma' \tag{5.9}$$

where K depends on the coefficients of (2.3) *and G only, and can be determined through* (5.13) *and* (5.14) *below. Here we assume that the length of the path γ connecting z and a fixed point z_0 is less than a fixed constant.*

Proof. Let

$$k(z) = w(z) \exp\left[-\int_{\zeta_1}^{\bar{z}} A(z, t)\, dt \right]; \tag{5.10}$$

therefore, (3.1) becomes (we may take $\zeta_0 = \bar{z}_0$)

$$\begin{aligned} k(z) = \phi(z) &+ \int_{z_0}^{\bar{z}} \Gamma_1(z, \bar{z}, t, \bar{z}_0)\phi(t)\, dt \\ &+ \int_{\bar{z}_0}^{\bar{z}} \Gamma_2(z, \bar{z}, z_0, t)\phi^*(t)\, dt \end{aligned} \tag{5.11}$$

for $z \in G \cup \partial G$.

Let γ denote a path from z_0 to z. In view of the method of successive approximation, we have

$$\phi(z) = k(z) + \int_{z_0}^{z} H_1(z, t)k(t)\, dt + \int_{\bar{z}_0}^{z} H_2(z, t)k^*(t)\, dt \tag{5.12}$$

for $z \in \gamma$. Furthermore,

$$|H_1(z, t)| \leqq 4M_1 \exp(4M_1|z - t|) < M_2,$$

$$|H_2(z, t)| \leqq 4M_1 \exp(4M_1|z - t|) < M_2$$

for $z, t \in G \cup \partial G$.

Hence, for $z \in \Gamma'$

$$|\phi(z)| \leqq \varepsilon M_3 \left(1 + 2M_2 \int_{\gamma} |dt| \right) \tag{5.13}$$

and for $z \in G$

$$|\phi(z)| \leqq M M_3 \left(1 + 2M_2 \int_{\gamma} |dt| \right). \tag{5.14}$$

The desired inequality therefore follows from (5.13) and (5.14).

Now we are in the position to give a solution for the "stability problem".

THEOREM 5.3. *If $w(z)$ is a bounded solution of* (1.1) *in G, continuous in $G \cup \Gamma'$, such that* (5.6) *and* (5.7) *hold, then*

$$|w(z)| \leqq KM_4\left\{M^{1-\omega(z)}\varepsilon^{\omega(z)} + 2\int_{z_0}^{z} M^{1-\omega(t)}\varepsilon^{\omega(t)}|dt|\right\} \tag{5.15}$$

for $z \in G$, where $\omega(z)$ is the harmonic measure of the curve Γ' with respect to the point z and the domain G, K is the constant given in Theorem 5.2, *and the line integral is along the arc in G which takes the shortest length from z_0 to z. Here, we again assume that* (1.1) *is homogeneous.*

Proof. According to Lemma 3.1 and Theorem 5.2, $w(z)$ has the representation (3.1) and the function $\phi(z)$ in (3.1) satisfies the inequalities (5.8) and (5.9).

In view of the Theorem 5.1,

$$|\phi(z)| \leqq (\varepsilon K)^{\omega(z)}(KM)^{1-\omega(z)} = K\varepsilon^{\omega(z)}M^{1-\omega(z)}. \tag{5.16}$$

Therefore, from (3.1), we have

$$|w(z)| \leqq M_4\left\{KM^{1-\omega(z)}\varepsilon^{\omega(z)} + 2KM_1\int_{z_0}^{z} M^{1-\omega(t)}\varepsilon^{\omega(t)}|dt|\right\}. \tag{5.17}$$

Then the desired result follows from (5.7).

Remark 5.1. The above theorem has demonstrated that if one prescribes in addition to the Cauchy data a supplementary condition, namely that the solution of (1.1) be uniformly bounded by some constant M in its region of definition G, then the Cauchy problem for (1.1) is stable.

Remark 5.2. The two-constant theorem also can be generalized to (1.1) with bounded measurable coefficients, and to higher order elliptic equations. They will be given in the forthcoming papers.

6. The construction of the elementary solution for (1.1) in G. The next few sections will study the existence and the construction of solutions for (1.1) in a general domain. Since our methods depend on the use of a generalized Cauchy formula, we shall first construct the elementary solution for (2.3). Our method simplifies the one given in Vekua [21]; in particular, we can omit the formulas on pp. 79–80 in Vekua [21].

A convenient technical simplification results from using the transformation

$$w(z, \bar{z}) = w_0(z, \bar{z}) \exp\int_{\zeta_1}^{\bar{z}} A(z, t)\, dt \tag{6.1}$$

where ζ_1 is a fixed point in G. Substituting (6.1) in (2.3), we obtain the differential equation (see Vekua [21, (15.10) and (15.35b)])

$$\frac{\partial w_0}{\partial \bar{z}} = C(z, \bar{z})\overline{w_0(z, \bar{z})} + F_0(z, \bar{z}), \tag{6.2}$$

where

$$C(z,\bar{z}) = B(z,\bar{z}) \exp\left[\int_{\bar{\zeta}_1}^{z} A^*(\bar{z},t)\,dt - \int_{\zeta_1}^{\bar{z}} A(z,t)\,dt\right], \tag{6.3}$$

$$F_0(z,\bar{z}) = F(z,\bar{z}) \exp\left[-\int_{\zeta_1}^{\bar{z}} A(z,t)\,dt\right]. \tag{6.4}$$

Now we are in the position to construct the elementary solution of the homogeneous differential equation

$$\frac{\partial w}{\partial \bar{z}} = C(z,\bar{z})\overline{w(z)}. \tag{6.5}$$

The following lemma can be found in Vekua [22]; for the sake of completeness, we give a proof here.

LEMMA 6.1. *Let u, v, w be the solutions of*

$$\frac{\partial u}{\partial \bar{z}} + \bar{C}\bar{v} = 0, \tag{6.6}$$

$$\frac{\partial v}{\partial \bar{z}} + \bar{C}\bar{u} = 0, \tag{6.7}$$

$$\frac{\partial w}{\partial \bar{z}} - C\bar{w} = 0 \tag{6.8}$$

in G, continuous in $G \cup \partial G$. Then

$$\int_{\partial G} (uw\,dz - \bar{v}\bar{w}\,d\bar{z}) = 0. \tag{6.9}$$

Proof. By Green's formula, we have

$$\frac{1}{2i}\int_{\partial G} uw\,dz = \iint_G \frac{\partial(uw)}{\partial \bar{z}}\,dx\,dy, \tag{6.10}$$

$$-\frac{1}{2i}\int_{\partial G} \bar{v}\bar{w}\,d\bar{z} = \iint_G \frac{\partial(\bar{v}\bar{w})}{\partial z}\,dx\,dy. \tag{6.11}$$

Hence,

$$\begin{aligned}&\frac{1}{2i}\int_{\partial G} (uw\,dz - \bar{v}\bar{w}\,d\bar{z})\\ &= \iint_G \left\{u\left(\frac{\partial w}{\partial \bar{z}} - C\bar{w}\right) + \bar{v}\left(\frac{\partial \bar{w}}{\partial z} - \bar{C}w\right) + w\left(\frac{\partial u}{\partial \bar{z}} + \bar{C}\bar{v}\right) + \bar{w}\left(\frac{\partial \bar{v}}{\partial z} + Cu\right)\right\} dx\,dy = 0.\end{aligned} \tag{6.12}$$

LEMMA 6.2. *Let*

$$\phi(z) = \phi_0(z) + \left[\frac{1}{2\pi i}\int_{\bar{L}} \Gamma_2(z,\bar{z}_0,z_0,\tau)\phi_0^*(\tau)\,d\tau\right]\log(z - z_0), \tag{6.13}$$

where $\phi_0(z)$ is analytic in a double connected domain $G_1 \subset D \cup \sigma \cup \bar{D}$, L is a closed curve in G_1, and z_0 is a fixed point. Then the function $w(z, \bar{z})$ defined by

$$
\begin{aligned}
w(z, \bar{z}) &= \phi(z) + \int_{z_1}^{z} \Gamma_1(z, \bar{z}, t, \bar{z}_0)\phi(t)\, dt \\
&\quad + \int_{\bar{z}_1}^{\bar{z}} \Gamma_2(z, \bar{z}, z_0, t)\phi^*(t)\, dt \\
&= T\phi
\end{aligned} \tag{6.14}
$$

is a single valued solution of (6.5) *in G.*

Proof. See Vekua [21, pp. 72–75].

LEMMA 6.3. *Let*

$$\phi_0(z) = \frac{1}{z - z_0}. \tag{6.15}$$

Then the function $\phi(z)$ in (6.13) *is given by the formula*

$$\phi(z) = \frac{1}{z - z_0} + C(z, \bar{z}_0) \log (z - z_0). \tag{6.16}$$

Proof. See Vekua [21, p. 75].

LEMMA 6.4. *Let*

$$\phi_0(z) = \frac{1}{i(z - z_0)}. \tag{6.17}$$

Then the function $\phi(z)$ in (6.13) *is given by the formula*

$$\phi(z) = \frac{1}{i(z - z_0)} - \frac{1}{i} C(z, \bar{z}_0) \log (z - z_0). \tag{6.18}$$

Proof. We see that

$$\phi_0^*(\zeta) = \frac{1}{i(\zeta - z_0)}, \tag{6.19}$$

but, by formula (3.4)

$$\Gamma_2(z, \bar{z}_0, z_0, \bar{z}_0) = C(z, \bar{z}_0). \tag{6.20}$$

Hence (6.13) takes the form (6.18).

Let

$$
\begin{aligned}
X_1(z, \bar{z}) &= T\left\{ \frac{1}{t - z_0} + C(t, \bar{z}_0) \log (t - z_0) \right\} \\
X_2(z, \bar{z}) &= T\left\{ \frac{1}{i(t - z_0)} - \frac{1}{i} C(t, \bar{z}_0) \log (t - z_0) \right\}
\end{aligned} \tag{6.21}
$$

and

$$\Omega_1(z, \bar{z}) = \tfrac{1}{2}(X_1 + iX_2), \qquad \Omega_2(z, \bar{z}) = \tfrac{1}{2}(X_1 - iX_2). \tag{6.22}$$

Then we have the following lemma.

LEMMA 6.5. *The functions Ω_1 and Ω_2 are the solutions of the equations*

$$\frac{\partial \eta_1}{\partial \bar{z}} + \bar{C}\bar{\eta}_2 = 0, \qquad \frac{\partial \eta_2}{\partial \bar{z}} + \bar{C}\bar{\eta}_1 = 0. \tag{6.23}$$

Furthermore,

$$\begin{aligned}\Omega_1(z, \bar{z}, z_0, \bar{z}_0) = {} & \frac{1}{z - z_0} + \int_{z_1}^{z} \Gamma_1(z, \bar{z}, t, \bar{z}_0) \frac{1}{t - z_0}\, dt \\ & - \int_{\bar{z}_1}^{\bar{z}} \Gamma_2(z, \bar{z}, z_0, \tau) C^*(\tau, z_0) \log(\tau - \bar{z}_0)\, d\tau,\end{aligned} \tag{6.24}$$

and

$$\begin{aligned}\Omega_2(z, \bar{z}, z_0, \bar{z}_0) = {} & \int_{\bar{z}_1}^{\bar{z}} \Gamma_2(z, \bar{z}, z_0, \tau) \frac{1}{\tau - \bar{z}_0} d\tau + C(z, \bar{z}_0) \log(z - z_0) \\ & + \int_{z_1}^{z} \Gamma_1(z, \bar{z}, t, \bar{z}_0) C(t, \bar{z}_0) \log(t - z_0)\, dt.\end{aligned} \tag{6.25}$$

Proof. By direct substitution, the desired results follow from (6.21).

DEFINITION 6.1. We call the functions $\{\Omega_1, \Omega_2\}$ a *pair of elementary solutions for* (6.5)

THEOREM 6.1. *Let $w(z)$ be a solution of* (6.5) *on G, continuous in $G \cup \partial G$. Then*

$$\begin{aligned}w(z_0) = {} & \frac{1}{2\pi i} \int_{\partial G} w(z) \Omega_1(z, \bar{z}, z_0, \bar{z}_0)\, dz \\ & - \overline{w(z)} \Omega_2(z, \bar{z}, z_0, \bar{z}_0)\, d\bar{z}\end{aligned} \tag{6.26}$$

for $z_0 \in G$.

Proof. If we remove the circle $|z - z_0| \leqq \rho$ from G, the functions Ω_1, Ω_2 will be analytic in the rest of G. We now have from Lemma 6.1 that

$$\int_{\partial G \cup \gamma_\rho} [\Omega_1(z, \bar{z}, z_0, \bar{z}_0) w(z)\, dz - \Omega_2(z, \bar{z}, z_0, \bar{z}_0) \overline{w(z)}\, d\bar{z} = 0, \tag{6.27}$$

where $w(z)$ is a solution of (1.1) in G, continuous in $G \cup \partial G$, and $\gamma_\rho = \{z \mid |z - z_0| = \rho\}$. If we pass to the limit as $\rho \to 0$ in (6.27), using Lemma 6.5 we have

$$w(z_0) = \frac{1}{2\pi i} \int_{\partial G} [w(z) \Omega_1(z, \bar{z}, z_0, \bar{z}_0)\, dz - w(z) \Omega_2(z, \bar{z}, z_0, \bar{z}_0)\, d\bar{z}].$$

7. The existence and construction of solutions to (1.1) in G. In this section the Cauchy data $\rho(x)$ in (2.5) is assumed to be bounded and continuous on σ. We shall give necessary and sufficient conditions for the existence of the solution to

the Cauchy problem for (2.3) in D and methods for constructing it. Since the principal part of (2.3) is invariant under conformal mappings, the results in $D \cup \sigma$ can be extended to a more general domain G with nonanalytic initial curve. This problem has been studied by Hadamard [9] for harmonic functions and John [13] for holomorphic functions.

THEOREM 7.1. *Let*

$$G_0(z) = \frac{1}{2\pi i} \int_\sigma \rho_0(t)\Omega_1 \, dt - \overline{\rho_0(t)}\Omega_2 \, d\bar{t} \tag{7.1}$$

where $\rho_0(t)$ is a bounded continuous function on σ. Then $G_0(z)$ is a solution of (6.5) *in $D \cup \bar{D}$ (off σ). Furthermore, if $\rho_0(t)$ is Hölder continuous in σ, then*

$$G_0(z) \to G_0^+(x) = \tfrac{1}{2}\rho(x) + G_0(x) \quad \text{as } z \to x \in \sigma,$$

for $z \in D$, the first integral in (7.1) *being understood as the Cauchy principal value, and the second integral converging in the ordinary sense (we assume that D is on the left of σ).*

Proof. The proof is accomplished by direct calculation and properties of singular integrals.

THEOREM 7.2. *Let*

$$F(x) = \rho_0(x) - G_0^+(x, x). \tag{7.2}$$

Then $\rho_0(x) = w_0(x, x)$ for a solution $w_0(z, \bar{z})$ of (6.5) *in D if and only if $F(x)$ is analytic on σ and its analytic continuation $F(z)$ is a holomorphic function in $D \cup \sigma \cup \bar{D}$.*

Proof. If $w_0(z, \bar{z})$ is a solution of the problem in D, then the function

$$\begin{aligned} H(z, \bar{z}) &= w_0(z, \bar{z}) - G_0(z, \bar{z}) \\ &= \frac{1}{2\pi i} \int_{\partial D} w_0\Omega_1 \, dt - \bar{w}_0\Omega_2 \, d\bar{t} - G_0 \\ &= \frac{1}{2\pi i} \int_{\partial D - \sigma} w_0\Omega_1 \, dt - \bar{w}_0\Omega_2 \, d\bar{t} \end{aligned} \tag{7.3}$$

is a solution in $D \cup \sigma \cup \bar{D}$; hence $F(z) = H(z, z)$ is holomorphic in $D \cup \sigma \cup \bar{D}$.

Conversely, if $F(z)$ is holomorphic in $D \cup \sigma \cup \bar{D}$, by Theorem 3.1 there exists a solution $H(z, \bar{z})$ of (6.5) in $D \cup \sigma \cup \bar{D}$ such that $H(x, x) = F(x)$. Thus

$$w_0(z, \bar{z}) = H(z, \bar{z}) + G_0(z, \bar{z}) \tag{7.4}$$

is a solution of the problem in D.

The proof of the above theorem also gives a constructive method for the solution of the problem which we shall restate in the following theorem:

THEOREM 7.3. *Let $w_0(z, \bar{z})$ be a solution of* (6.5) *in D, continuous in $D \cup \sigma$. Let the values of $w_0(x, x)$ be known on σ. Also let*

$$F(x) = w_0(x, x) - G_0^+(x, x). \tag{7.5}$$

Then the analytic continuation $F(z)$ is holomorphic in $D \cup \sigma \cup \bar{D}$, and we can construct a solution $H(z, \bar{z})$ of (6.5) *in $D \cup \sigma \cup \bar{D}$ such that $H(x, x) = F(x)$ by using*

Theorem 3.1. *Then* $w_0(z, \bar{z})$ *is given by the formula*

$$w_0(z, \bar{z}) = H(z, \bar{z}) + G_0(z, \bar{z}). \tag{7.6}$$

By using the transformation (6.1) and formula (3.1) we can immediately obtain the necessary and sufficient condition for the existence of a solution to (2.3) and a procedure for its construction.

THEOREM 7.4. *Let* $\rho_0(x) = \rho(x) \exp[-\int_{\zeta_1}^{x} A(x, t)\, dt] - U_0(x, x)$, *where* $U_0(z, \bar{z})$ *is given in* (3.2). *Then* $\rho(x)$ *is the Cauchy data for a solution* $w(z, \bar{z})$ *of* (2.3) *in* D *if and only if* $\rho_0(x)$ *satisfies the condition for existence in Theorem* 7.2. *Furthermore, if the solution* $w(z, \bar{z})$ *exists in* $D \cup \sigma$, $w(z, \bar{z})$ *is given by the formula*

$$w(z, \bar{z}) = \{H(z, \bar{z}) + G_0(z, \bar{z}) + U_0(z, \bar{z})\} \exp \int_{\zeta_1}^{z} A(z, t)\, dt \tag{7.7}$$

where H *and* G *are given in Theorem* 7.2.

Another version for the necessary and sufficient condition can be stated as follows.

THEOREM 7.5. *The function* $w(x, x)$ *is the Cauchy data for a solution* $w(z, \bar{z})$ *of* (2.3) *in* D *if and only if the function* $\phi(x)$ *in the equation*

$$\phi(x) = k(x) + \int_{x_0}^{x} H_1(x, t)k(t)\, dt + \int_{x_0}^{x} H_2(x, t)k^*(t)\, dt \tag{7.8}$$

is the Cauchy data for a holomorphic function $\phi(z)$ *in* $D \cup \sigma$. *Furthermore, if the solution* $w(z, \bar{z})$ *exists,* $w(z, \bar{z})$ *is given by the* (3.1), *where* H_1, H_2 *and* k *are defined in* (5.10) *and* (5.12).

The proof of Theorem 7.5 follows from formula (3.1).

Remark 7.1. The construction of $\phi(z)$ in (7.8) can be obtained as follows: Let

$$G(z) = \frac{1}{2\pi i} \int_{\sigma} \frac{\phi(t)}{t - z}\, dt$$

and let

$$F(x) = \phi(x) - G^{+}(x).$$

Then $F(z)$ is holomorphic in $D \cup \sigma \cup \bar{D}$, and thus $\phi(z)$ is given by the formula

$$\phi(z) = F(z) + G(z).$$

Remark 7.2. Theorem 7.2 depends only on Green's formula and an elementary solution; therefore its method should be applicable to more general elliptic equations.

8. The construction of an approximate solution for (1.1) in *G*. We now come to the case where $w(z)$ is given on Γ' only within an error ε. We shall give two methods for constructing the approximate solution of (6.5) in G. The methods are based on the use of the Carleman function.

DEFINITION 8.1. We call $G(z, \zeta, \delta)$ a *Carleman function for the domain G and the curve* Γ' if $G(z, \zeta, \delta)$ has the following properties:

(a) $$G(z, \zeta, \delta) = \frac{1}{\zeta - z} + \tilde{G}(z, \zeta, \delta)$$

where $\tilde{G}$ is a holomorphic function of the variable ζ, holomorphic and bounded in G.

(b) The function $G(z, \zeta, \delta)$ satisfies the inequality

$$\int_{\Gamma''} |G(z, \zeta, \delta)|\,|d\zeta| \leqq \delta.$$

(c) The function

$$\mu(z, \tau) \cdot \varepsilon\gamma' = M\tau$$

has a solution $\tau(z, \varepsilon)$ for each ε and M, and $\tau(z, \varepsilon) \to 0$ as $\varepsilon \to 0$, where γ' is the length of Γ' and $\mu(z, \delta) = \max_{\zeta \in \Gamma'} |G|$.

The existence and the construction of the Carleman function in a simply connected domain G can be found in Lavrentiev [15].

Let the Carleman function $G(z, \zeta, \delta)$ for the domain G and the curve Γ' be known where $\partial G = \Gamma' + \Gamma''$. Let the values of $f(z)$ on the curve Γ' be known with accuracy ε, i.e. a function $f_\varepsilon(z)$ is known on the curve Γ' such that

$$|f_\varepsilon(z) - f(z)| \leqq \varepsilon, \qquad z \in \Gamma'. \tag{8.1}$$

Under the assumption that

$$|f(z)| \leqq M, \qquad z \in \Gamma'', \tag{8.2}$$

we now construct the approximate solution for $f(z)$.

THEOREM 8.1. *Let the function* $f_{\delta\varepsilon}(z)$ *be defined by*

$$f_{\delta\varepsilon}(z) = \frac{1}{2\pi i} \int_{\Gamma'} G(z, \zeta, \delta) f_\varepsilon(\zeta)\, d\zeta. \tag{8.3}$$

Then

$$|f(z) - f_{\delta\varepsilon}(z)| \leqq \frac{1}{\pi} M \cdot \tau(z, \varepsilon), \tag{8.4}$$

for $z \in G$.

(This theorem can be found in [15]; however, for the sake of completeness, we give a proof here.)

Proof. According to Cauchy's formula

$$f(z) = \frac{1}{2\pi i} \int_{\partial G} G(z, \zeta, \delta) f(\zeta)\, d\zeta. \tag{8.5}$$

Therefore,

$$f(z)-f_{\zeta\varepsilon}(z)=\frac{1}{2\pi i}\int_{\Gamma''} G(z,\zeta,\delta)f(\zeta)\,d\zeta + \frac{1}{2\pi i}\int_{\Gamma'} G(z,\zeta,\delta)[f(\zeta)-f_\varepsilon(\zeta)]\,d\zeta. \tag{8.6}$$

By virtue of the properties of the Carleman function we see

$$\left|\int_{\Gamma''} Gf\,d\zeta\right| \leqq \delta\cdot M, \tag{8.7}$$

and

$$\left|\int_{\Gamma'} G[f(\zeta)-f_\varepsilon(\zeta)]\,d\zeta\right| \leqq \mu(z,\delta)\cdot\varepsilon\cdot\gamma'. \tag{8.8}$$

Substituting (8.7), (8.8) in (8.6) we obtain

$$|f(z)-f_{\delta\varepsilon}(z)| \leqq \frac{1}{2\pi}[\delta M+\mu(z,\delta)\cdot\varepsilon\cdot\gamma'] = \frac{1}{\pi}M\cdot\tau(z,\varepsilon). \tag{8.9}$$

Let

$$G_1(z,\zeta,\delta)=\int_{\zeta_1}^{\zeta} G(z,t,\delta)\,dt = \int_{\zeta_1}^{\zeta}\frac{1}{t-z}\,dt+\int_{\zeta_1}^{\zeta}\tilde{G}(z,t,\delta)\,dt = \ln(\zeta-z)+\tilde{\tilde{G}}(z,\zeta,\delta), \tag{8.10}$$

where ζ_1 is an endpoint of Γ''. Then

$$\int_{\Gamma''} |G_1(z,\zeta,\delta)|\,|d\zeta| \leqq \delta\gamma'' \tag{8.11}$$

where γ'' is the length of Γ''. And

$$\max_{\zeta\in\Gamma'} |G_1(z,\zeta,\delta)| = \gamma'\mu(z,\delta). \tag{8.12}$$

Define (see (6.24) and (6.25))

$$\tilde{\Omega}_1(z,\bar{z},z_0,\bar{z}_0,\delta)=G(z,z_0,\delta)+\int_{z_1}^{z}\Gamma_1(z,\bar{z},t,\bar{z}_0)G(z_0,t,\delta)\,dt - \int_{\bar{z}_1}^{\bar{z}}\Gamma_2(z,\bar{z},z_0,\tau)C^*(\tau,z_0)G_1^*(z_0,t,\delta)\,dt, \tag{8.13}$$

$$\tilde{\Omega}_2(z, \bar{z}, z_0, \bar{z}_0, \delta) = C(z, \bar{z}_0)G_1(z_0, z, \delta)$$

$$+\int_{\bar{z}_1}^{\bar{z}} \Gamma_2(z, \bar{z}, z_0, \tau)G^*(z_0, \tau, \delta)\, d\tau \tag{8.14}$$

$$+\int_{z_1}^{z} \Gamma_1(z, \bar{z}, t, \bar{z}_0)C(t, \bar{z}_0)G_1(z_0, t, \delta)\, dt,$$

and let

$$M_5 = \sup_{\tau, z \in D \cup \sigma \cup \bar{D}} C^*(\tau, z). \tag{8.15}$$

Then, by (8.13) and (8.14), we see

$$\int_{\Gamma''} |\tilde{\Omega}_1(z, \bar{z}, t, \bar{t}, \delta)|\, |dt| \leqq \delta + M_1\delta + M_5M_1\delta\gamma'', \tag{8.16}$$

$$\int_{\overline{\Gamma''}} |\bar{\Omega}_2(z, \bar{z}, t, \bar{t}, \delta)|dt| \leqq M_5\delta\gamma'' + M_1\delta + M_5M_1\delta\gamma'', \tag{8.17}$$

and

$$\max_{\zeta \in \Gamma'} |\tilde{\Omega}_1(z, z, \zeta, \zeta, \delta_0)| \leqq \mu(z, \delta)\{1 + M_1\gamma' + M_1 + M_5\gamma'\gamma'\}, \tag{8.18}$$

$$\max_{\zeta \in \Gamma'} |\tilde{\Omega}_2(z, z, \zeta, \zeta, \delta_0)| \leqq \mu(z, \delta)\{M_5\gamma' + \gamma'M_1 + M_5M_1\gamma'\gamma'\}. \tag{8.19}$$

Let $\delta\delta_1$, $\delta\delta_2$, $\mu(z, \delta)\delta_3$ and $\mu(z, \delta)\delta_4$ denote the values on the right side of inequalities (8.16), (8.17), (8.18) and (8.19) respectively. We now construct the approximate solutions for a solution $w(z)$ *of* (6.5) in G.

THEOREM 8.2. *Suppose*

$$|w(z)| \leqq M \quad \textit{for } z \in G, \tag{8.20}$$

$$|w(z) - w_\varepsilon(z)| < \varepsilon \quad \textit{for } z \in \Gamma'. \tag{8.21}$$

Let

$$w_{\varepsilon\delta}(z) = \frac{1}{2\pi i}\int_{\Gamma'} w_\varepsilon(t)\tilde{\Omega}_1(t, \bar{t}, z, \bar{z}, \delta)\, dt$$

$$-\overline{w_\varepsilon(t)}\tilde{\Omega}_2(t, \bar{t}, z, \bar{z}, \delta)\, d\bar{t}. \tag{8.22}$$

Then

$$|w(z) - w_{\varepsilon\delta}(z)| \leqq \frac{1}{\pi}M_6\tau(z, \varepsilon) \quad \textit{for } z \in G \tag{8.23}$$

where

$$M_6 = M(\delta_1 + \delta_2),$$

$\tau(z, \varepsilon)$ is a solution of

$$\varepsilon\mu(z, \tau)\gamma'(\delta_3+\delta_4) = M\tau(\delta_1+\delta_2), \tag{8.24}$$

and $\tau(z, \varepsilon)\to 0$ *as* $\varepsilon\to 0$.

Proof. By (8.13), (8.14), Lemma 6.5, and the argument using for deriving the generalized Cauchy formula (Theorem 6.1), we have

$$\begin{aligned} w(z) = \frac{1}{2\pi i}\int_{\partial G} & w(t)\tilde{\Omega}_1(t, \bar{t}, z, \bar{z}, \delta)\, dt \\ & -\overline{w(t)}\tilde{\Omega}_2(t, \bar{t}, z, \bar{z}, \delta)\, d\bar{t}. \end{aligned} \tag{8.25}$$

Therefore,

$$\begin{aligned} w(z) - w_{\varepsilon\delta}(z) = \frac{1}{2\pi i}\int_{\Gamma''} & w(t)\tilde{\Omega}_1(t, \bar{t}, z, \bar{z}, \delta)\, dt \\ & -\overline{w(t)}\tilde{\Omega}_2(t, \bar{t}, z, \bar{z}, \delta)\, d\bar{t} \\ & +\frac{1}{2\pi i}\int_{\Gamma'} [w(t) - w_\varepsilon(t)]\tilde{\Omega}_1(t, \bar{t}, z, \bar{z}, \delta)\, dt \\ & -[\overline{w(t)} - \overline{w_\varepsilon(t)}]\tilde{\Omega}_2(t, \bar{t}, z, \bar{z}, \delta)\, d\bar{t}. \end{aligned} \tag{8.26}$$

By virtue of (8.16), (8.17), (8.18) and (8.19), we have

$$\left|\int_{\Gamma''} w(t)\tilde{\Omega}_1\, dt - \overline{w(t)}\tilde{\Omega}_2\, d\bar{t}\right| \leqq M\delta\delta_1 + M\delta\delta_2 = M\delta(\delta_1+\delta_2), \tag{8.27}$$

and

$$\left|\int_{\Gamma'} [w - w_\varepsilon]\tilde{\Omega}_1\, dt - [\bar{w}_0 - \bar{w}_\varepsilon]\tilde{\Omega}_2\, d\bar{t}\right| \leqq \varepsilon\mu(z, \delta)(\delta_3+\delta_4)\gamma'. \tag{8.28}$$

Substituting (8.27) and (8.28) into (8.26), we obtain

$$|w(z) - w_{\varepsilon\delta}(z)| \leqq \frac{1}{2\pi}[M\delta(\delta_1+\delta_2) + \varepsilon\mu\gamma'(\delta_3+\delta_4)]. \tag{8.29}$$

Hence the desired inequality follows from (8.29) and (8.24).

We now give the second method for constructing an approximate solution. Let

$$k_\varepsilon(z) = w_\varepsilon(z)\exp\left[-\int_{\zeta_1}^{\bar{z}} A(z, t)\, dt\right], \qquad z\in\Gamma'. \tag{8.30}$$

We define the function $\phi_\varepsilon(z)$ by

$$\phi_\varepsilon(z) = k_\varepsilon(z) + \int_{z_0}^{z} H_1(z, t)k_\varepsilon(t)\, dt + \int_{\bar{z}_0}^{z} H_2(z, t)k_\varepsilon^*(t)\, dt, \tag{8.31}$$

for $z \in \Gamma'$, where H_1 and H_2 are given by the formula (5.12), and z_0 can be assumed on Γ'. According to (8.20), (8.21),

$$\begin{aligned} |\phi(z)-\phi_\varepsilon(z)| &\leqq M_3\varepsilon + 2M_3\varepsilon M_2\gamma' \\ &= \varepsilon M_3(1+2M_2\gamma') \\ &= \varepsilon' \quad \text{on } \Gamma'. \end{aligned} \tag{8.32}$$

Again, by formula (5.8), we see

$$|\phi(z)| \leqq KM = M' \tag{8.33}$$

for $z \in G$. Then by Theorem 8.1, we can construct a function $\phi_{\delta\varepsilon'}(z)$ to approximate $\phi(z)$, and have the estimate

$$|\phi_{\delta\varepsilon'}(z)-\phi(z)| \leqq \frac{1}{\pi} M'\tau(z, \varepsilon') \quad \text{for } z \in G, \tag{8.34}$$

where $\tau(z, \varepsilon')$ is a solution of

$$\frac{\tau}{\mu(z, \tau)} = \frac{\varepsilon'\gamma'}{M'}. \tag{8.35}$$

Then the function

$$\begin{aligned} w_{\delta\varepsilon}(z) = \Big\{ \phi_{\delta\varepsilon'}(z) &+ \int_{z_0}^{\bar{z}} \Gamma_1(z, \bar{z}, t, z_0)\phi_{\delta\varepsilon'}(t)\, dt \\ &+ \int_{\bar{z}_0}^{\bar{z}} \Gamma_2(z, \bar{z}, \bar{z}_0, t)\phi^*_{\delta\varepsilon'}(t)\, dt \Big\} \exp \int_{\zeta_1}^{\bar{z}} A(z, t)\, dt \end{aligned} \tag{8.36}$$

is an approximate solution for the problem. Note that the line integrals in (8.36) are along the shortest path from z_0 to z in G. Furthermore

$$|w_{\delta\varepsilon}(z)-w(z)| \leqq M_4 \cdot \frac{1}{\pi} M'\tau(z, \varepsilon')\{1+2M_1\gamma\}, \tag{8.37}$$

where γ is the length of the integral path from z_0 to z. We thus have

THEOREM 8.3. *If $w(z)$ is a solution of the homogeneous equation* (2.3) *($F = 0$) in G, satisfies the condition* (8.20), *and $w_\varepsilon(z)$ is a continuous function such that* (8.21) *holds, then the function $w_{\delta\varepsilon}(z)$ in* (8.36) *is an approximate solution of $w(z)$ in G. Furthermore, the estimate of "stability" is given by* (8.37).

9. "Stability" of the Cauchy problem for (2.4) in a simply connected domain *G*. In order to establish "stability" for the Cauchy problem, we make use of the fact that the difference, $w = w_1 - w_2$, of two solutions satisfies a system of linear elliptic equations. In fact, we have that

$$w_{1\bar{z}} = H(w_1, \bar{w}_1, z, \bar{z}), \tag{9.1}$$

$$w_{2\bar{z}} = H(w_2, \bar{w}_2, z, \bar{z}) \tag{9.2}$$

and subtracting these two equations gives

$$w_{\bar{z}} = H(w_1, \bar{w}_1, z, \bar{z}) - H(w_2, \bar{w}_2, z, \bar{z}) \tag{9.3}$$

Let $H_i = H(w_i, \bar{w}_i, z, \bar{z})$. Then the difference $H_1 - H_2$ can be computed as follows. Set

$$F(t) = H[tw_1 + (1-t)w_2, t\bar{w}_1 + (1-t)\bar{w}_2, z, \bar{z}].$$

Then

$$H_1 - H_2 = \int_0^1 F'(t)\,dt = w\int_0^1 H_w\,dt + \bar{w}\int_0^1 H_{\bar{w}}\,dt. \tag{9.4}$$

Thus (9.3) may be written in the form

$$w_{\bar{z}} = Aw + B\bar{w}, \tag{9.5}$$

where

$$A = \int_0^1 H_w\,dt, \qquad B = \int_0^1 H_{\bar{w}}\,dt. \tag{9.6}$$

Since we assume w_1 and w_2 have prior bound M, i.e.

$$|w_1|, |w_2| < M \quad \text{for } z \in G, \tag{9.7}$$

the functions A and B are bounded by a fixed constants. Hence from (3.3), (3.4), (3.5) the functions Γ_1, Γ_2 (which depend only on w_1 and w_2) in (3.1) will have a fixed bound M_1 if w_1 and w_2 satisfy condition (9.7). Applying Theorem 5.3 to (9.5) we have (using the notation given in Theorem 5.3) the following theorem.

THEOREM 9.1. *If w_1 and w_2 satisfy* (9.7) *in G and*

$$|w_1(z) - w_2(z)| < \varepsilon \quad \textit{for } z \in \Gamma',$$

then

$$|w(z)| \leqq KM_4\left\{M^{1-\omega(z)}\varepsilon^{\omega(z)} + 2\int_{z_0}^{z} M^{1-\omega(t)}\varepsilon^{\omega(t)}\,dt\right\}$$

for $z \in G$.

Acknowledgment. The author is very grateful for the valuable suggestions and comments made by Professors Robert Gilbert and David Colton.

REFERENCES

[1] A. K. AZIZ, R. P. GILBERT AND H. C. HOWARD, *A second order non-linear elliptic boundary value problem with generalized Goursat data*, Ann. Mat. Pura Appl., 72 (1966), pp. 325–341.

[2] T. CARLEMAN, *Les Fonctions Quasi Analytiques*, Paris, 1928.

[3] D. COLTON, *Cauchy's problem for almost linear elliptic equations in two independent variables*, J. Approximation Theory, 3 (1970), pp. 66–71.

[4] ———, *Cauchy's problem for almost linear elliptic equations in two independent variables*, II, Ibid., 4 (1971), pp. 288–294.

[5] ———, *Cauchy's problem for a class of fourth order elliptic equations in two independent variables*, Applicable Anal., (1971), pp. 13–22.

[6] P. GARABEDIAN, *Partial Differential Equations*, Interscience, New York, 1964.

[7] R. P. GILBERT, *Function Theoretic Methods in Partial Differential Equations*, Academic Press, New York, 1969.

[8] ———, *Constructive Methods for Elliptic Equations*, Lecture Notes in Mathematics, vol. 365, Springer-Verlag, New York, 1974.
[9] J. HADAMARD, *Lectures on the Cauchy Problem in Linear Partial Differential Equations*, Yale University Press, New Haven, CT, 1923.
[10] P. HENRICI, *A survey of I. N. Vekua's theory of elliptic partial differential equations with analytic coefficients*, Z. Angew. Math. Phys., 8 (1957), pp. 189–203.
[11] F. JOHN, *Numerical solution of the head equation for preceding time*, Ann. Mat. Pure Appl., 40 (1955), pp. 129–142.
[12] ———, *A note on "improper" problems in partial differential equations*, Comm. Pure Appl. Math., 8 (1955), pp. 494–495.
[13] ———, *Differential Equations with Approximate and Improper Data*, New York University, New York, 1955.
[14] ———, *Continuous dependence on data for solutions of partial differential equations with a prescribed bound*, Comm. Pure Appl. Math., 13 (1960), pp. 551–585.
[15] M. M. LAVRENTIEV, *Some Improperly Posed Problems in Mathematical Physics*, Springer-Verlag, New York, 1967.
[16] L. E. PAYNE, *On some non-well-posed problems for partial differential equations*, Numer. Sol. of Nonlinear Diff. Eqtns., M.R.C. Conference, Univ. of Wisconsin, John Wiley, 1966, pp. 239–263.
[17] ———, *Some remarks on improperly posed problems for partial differential equations*, Lecture Notes in Mathematics, vol. 316, Springer-Verlag, New York, 1973, pp. 1–30.
[18] L. E. PAYNE AND D. SATHER, *On an initial-boundary value problem for a class of degenerate elliptic operators*, Ann. Mat. Pura Appl., 78 (1968), pp. 323–338.
[19] C. PUCCI, *Discussions del problema di Cauchy per le equazioni ditipo ellitico*, Ibid., 30 (1958), pp. 391–412.
[20] D. SATHER AND J. SATHER, *The Cauchy problem for an elliptic parabolic operator*, Ibid., 30 (1958), pp. 197–214.
[21] I. N. VEKUA, *New Methods for Solving Elliptic Equations*, North-Holland, Amsterdam, 1967.
[22] ———, *Systeme von Differentialgleichungen erster Ordaung vom elliptischen Typus und Randwertaufgaben*, Berlin, 1956.
[23] W. WENDLAND AND W. HAACK, *Lectures on Partial and Pfaffian Differential Equations*, Pergamon Press, Oxford, 1972.
[24] C. L. YU, *Reflection principle for systems of first order elliptic equations with analytic coefficients*, Trans. Amer. Math. Soc., 164 (1972), pp. 489–501.
[25] ———, *Cauchy problem and analytic continuation for systems of first order elliptic equations with analytic coefficients*, Ibid., 185 (1973), pp. 429–443.
[26] ———, *On the global solvability of linear elliptic Cauchy problem*, to appear.
[27] ———, *Solvability of the Cauchy problem for linear elliptic equations with analytic coefficients*, to appear.

SIAM J. MATH. ANAL.
Vol. 8, No. 4, August 1977

SECOND ORDER SINGULAR BOUNDARY VALUE PROBLEMS*

T. C. LEE AND D. WILLETT†

Abstract. A theorem for the existence of a solution to singular differential equations of the form $Ly = f(t, y, y')$ satisfying generalized boundary conditions at two points a and b is proven. The operator L is assumed to be a generalized disconjugate second order linear differential operator on $[a, b]$ and may be singular in the sense that the coefficients in L may be discontinuous at a or b, or a may be $-\infty$ and b may be ∞. The solution satisfies $\alpha \leqq y \leqq \beta$ where it is assumed α and β are functions satisfying $\alpha \leqq \beta$, $L\beta \leqq f(t, \beta, \beta')$ and $L\alpha \geqq f(t, \alpha, \alpha')$.

1. The result. The use of differential inequalities in the existence theory of second order boundary value problems for ordinary differential equations is a well-established proposition (cf., e.g., [1]–[6], [9]–[12], [14]–[20]). Recently [8], [10] applications of this theory have been made in the area of singular perturbation theory. The point of this paper is to present a comprehensive theorem in this area extending a variety of known results and yet remaining relatively easy to apply.

Let $-\infty \leqq a < b \leqq \infty$ and define

$$L = D^2 + p(t)D + q(t) \qquad (D = d/dt), \tag{1.1}$$

where p and q are generally assumed to be continuous functions on the open interval (a, b), i.e., $p, q \in C(a, b)$. We assume L is *disconjugate in the generalized sense on the closed interval* $[a, b]$. This and other less widely known concepts will be made precise in the succeeding sections. Then $Ly = 0$ has a *fundamental principal system* (u_1, u_2) of solutions on $[a, b]$. Let $u = u_1 + u_2$ and define

$$\mathscr{D}^0 y(t) = \lim_{s \to t} \frac{y(s)}{u(s)}, \qquad \mathscr{D}^1 y(t) = \lim_{s \to t} \frac{w(u, y)(s)}{w(u_1, u_2)(s)} \qquad (a \leqq t \leqq b), \tag{1.2}$$

where $w(g, h)$ always denotes the 2×2 Wronskian determinant of the functions g and h and the above limits are the appropriate one-sided limits at a and b.

We will consider in this paper a boundary value problem for the differential equation

$$Ly = f(t, y, y'), \qquad a < t < b, \quad (y' = Dy) \tag{1.3}$$

where $f \in C[(a, b) \times R^2]$ $[R = (-\infty, \infty)]$ and f is further restricted in the following manner. Let α, β be functions of class $C'(a, b)$ such that $\alpha \leqq \beta$ and $\mathscr{D}^0 a$, $\mathscr{D}^0 \beta$, $\mathscr{D}^1 \alpha$ and $\mathscr{D}^1 \beta$ are defined and bounded on $[a, b]$. Let

$$F(t, y, z) = f\left(t, y, \frac{u'(t)y + w(u_1, u_2)(t)z}{u(t)}\right), \tag{1.4}$$

$$\lambda = \max\left[\mathscr{D}^0\beta(b) - \mathscr{D}^0\alpha(a), \mathscr{D}^0\beta(a) - \mathscr{D}^0\alpha(b)\right]. \tag{1.5}$$

* Received by the editors October 6, 1975, and in revised form April 26, 1976.

† Department of Mathematics and Statistics, University of Calgary, Calgary, Alberta, Canada T2N 1N4. The research of the second author was supported in part by the National Research Council of Canada under Grants A5593 and A4069.

We assume f satisfies a *generalized Nagumo condition on* (a, b) with respect to such a pair (α, β), i.e., there exists a positive function ϕ on $[\lambda, \infty)$ such that

$$\frac{u^3(t)}{w^2(u_1, u_2)(t)}|F(t, y, z)| \leqq \phi(|z|) \tag{1.6}$$

for $\lambda \leqq |z| < \infty$, $\alpha(t) \leqq y \leqq \beta(t)$, $a < t < b$ and

$$\int_\lambda^\infty \frac{s}{\phi(s)}\,ds > \sup_{a<t<b} \mathscr{D}^0\beta(t) - \inf_{a<t<b} \mathscr{D}^0\alpha(t) \equiv K. \tag{1.7}$$

Let N be any real number such that for all $a < t < b$,

$$\max(\mathscr{D}^1\alpha(t), \mathscr{D}^1\beta(t)) < N \quad \text{and} \quad K < \int_\lambda^N \frac{s}{\phi(s)}\,ds, \tag{1.8}$$

and let

$$\psi(t) = \sup\{|F(t, y, z)| : |z| \leqq N, \alpha(t) \leqq y \leqq \beta(t)\}. \tag{1.9}$$

Define

$$\bar{\mathscr{D}}^1 y(b) = \lim_{t\to b^-} \frac{w(y, u_2)(t)}{w(u_1, u_2)(t)}, \qquad \bar{\mathscr{D}}^1 y(a) = \lim_{t\to a^+} \frac{w(u_1, y)(t)}{w(u_1, u_2)(t)} \tag{1.10}$$

and

$$J = \int_a^b \frac{\psi(s)u(s)}{w(u_1, u_2)(s)}\,ds. \tag{1.11}$$

Finally, let

$$\begin{aligned}
\Phi_0 = {} & \{(x, \mathscr{D}^0\beta(b) + J) : \mathscr{D}^0\alpha(a) \leqq x \leqq \mathscr{D}^0\beta(a)\} \\
& \cup \{(x, \mathscr{D}^0\alpha(b) - J) : \mathscr{D}^0\alpha(a) \leqq x \leqq \mathscr{D}^0\beta(a)\} \\
& \cup \{(\mathscr{D}^0\beta(a), y) : \mathscr{D}^0\beta(b) + J \geqq y > \bar{\mathscr{D}}^1\beta(a)\} \\
& \cup \{(\mathscr{D}^0\alpha(a), y) : \mathscr{D}^0\alpha(b) - J \leqq y < \bar{\mathscr{D}}^1\alpha(a)\}, \\
\Phi_1 = {} & \{(x, \mathscr{D}^0\beta(a) + J) : \mathscr{D}^0\alpha(b) \leqq x \leqq \mathscr{D}^0\beta(b)\} \\
& \cup \{(x, \mathscr{D}^0\alpha(a) - J) : \mathscr{D}^0\alpha(b) \leqq x \leqq \mathscr{D}^0\beta(b)\} \\
& \cup \{(\mathscr{D}^0\beta(b), y) : \mathscr{D}^0\beta(a) + J \geqq y > \bar{\mathscr{D}}^1\beta(b)\} \\
& \cup \{(\mathscr{D}^0\alpha(b), y) : \mathscr{D}^0\alpha(a) - J \leqq y < \bar{\mathscr{D}}^1\alpha(b)\},
\end{aligned}$$

$$S_0 = \{(x, y) : \mathscr{D}^0\alpha(a) \leqq x \leqq \mathscr{D}^0\beta(a), \mathscr{D}^0\alpha(b) - J \leqq y \leqq \mathscr{D}^0\beta(b) + J\}$$

and

$$S_1 = \{(x, y) : \mathscr{D}^0\alpha(b) \leqq x \leqq \mathscr{D}^0\beta(b), \mathscr{D}^0\alpha(a) - J \leqq y \leqq \mathscr{D}^0\beta(a) + J\}.$$

THEOREM. *Assume* $p, q \in C(a, b)$, $f \in C((a, b) \times R^2)$, *and there exist functions* $\alpha, \beta \in C^2(a, b)$ *such that* $\beta \geqq \alpha$, $L\beta \leqq f(t, \beta, \beta')$, $L\alpha \geqq f(t, \alpha, \alpha')$, $\mathscr{D}^0\alpha$, $\mathscr{D}^0\beta$, $\mathscr{D}^1\alpha$, $\mathscr{D}^1\beta$ *exist and are bounded on* $[a, b]$, f *satisfies a generalized Nagumo*

condition on (a, b) *with respect to* (α, β), $(u/(w(u_1, u_2)))L\alpha$ *and* $(u/(w(u_1, u_2)))L\beta$ *are integrable on* $[a, b]$, *and* J, $\bar{\mathscr{D}}^1\alpha(a)$, $\bar{\mathscr{D}}^1\alpha(b)$, $\bar{\mathscr{D}}^1\beta(a)$, $\bar{\mathscr{D}}^1\beta(b)$ *exist. If* Ω_0 *and* Ω_1 *are continua in* S_0 *and* S_1, *respectively, such that*

$$\Omega_0 \cup \Phi_0 \quad \textit{is connected} \tag{1.12}$$

and

$$\Omega_1 \cup \Phi_1 \textit{ is connected,} \tag{1.13}$$

then there exists a solution $y \in C^2(a, b)$ *of* (1.3) *such that*

$$(\mathscr{D}^0 y(a), \bar{\mathscr{D}}^1 y(a)) \in \Omega_0 \quad \textit{and} \quad (\mathscr{D}^0 y(b), \bar{\mathscr{D}}^1 y(b)) \in \Omega_1 \quad \textit{and} \quad \alpha \leqq y \leqq \beta. \tag{1.14}$$

We will discuss in detail the main ideas of the above theorem in the next three sections, prove the theorem in § 5 and give some examples in § 6. For generalizations of this theorem to functional differential equations and equations with just one boundary condition, see Lee [8].

2. Generalized disconjugacy. A second order linear operator $L = D^2 + p(t)D + q(t)$ is said to be *disconjugate on an interval* I in which p and q are continuous functions if no nontrivial solution of $Ly = 0$ has more than one zero, counting multiplicity, in I. For results identifying conditions on p and q which imply disconjugacy of L see Willett [20]–[22]. If L is disconjugate on an open interval $I = (a, b)$, then the solutions of $Ly = 0$ form a hierarchy of functions at each endpoint of I. In other words, there always exists at the endpoint b, and similarly at a, a pair (y_1, y_2) of solutions such that

$$\lim_{t \to b^-} \frac{y_1(t)}{y_2(t)} = 0. \tag{2.1}$$

A pair (y_1, y_2) of nontrivial solutions of $Ly = 0$ satisfying (2.1) is called a *principal system at* b. Thus, (e^{-t}, e^t) is a principal system at ∞ for $D^2 - 1$, and $(1 - t, 1)$ is a principal system at 1 for D^2. For any nontrivial solution y of $Ly = 0$ and principal system (y_1, y_2) at b, there exist constants c_1 and c_2 not both zero such that $y = c_1 y_1 + c_2 y_2$. If $c_2 = 0$, we define y to have a *zero* (*generalized zero*) at b; if $c_2 \neq 0$, we define y to *not* have a zero at b. This definition is clearly independent of the principal system (y_1, y_2); hence, it defines a way to attach zeros to solutions at points (in this case, a and b) where the equation, and hence the solutions, are not defined. Thus, the function 1 has a zero at ∞ and $-\infty$ with respect to the operator D^2, but does not have a zero at ∞ with respect to the operator $D^2 + D$. Zeros defined in this way depend upon the operator. Of course, if an operator L is nonsingular at b (i.e., $b < \infty$ and p and q can be extended to be continuous at b), then a solution y has a generalized zero at b if and only if

$$\lim_{t \to b^-} y(t) = 0,$$

i.e., the continuous extension of y to b has an ordinary zero at b. Thus, if $p, q \in C(a, b)$ and $a < c < b$, generalized zero and ordinary zero mean the same thing at c.

A second order operator L with coefficient functions p and q assumed continuous on just the open interval (a, b) is defined to be *disconjugate in the generalized sense on the closed interval* $[a, b]$ if no nontrivial solution of $Ly = 0$ has more than one zero in $[a, b]$. Thus, D^2 is generalized disconjugate on $[a, \infty]$ if a is finite, but is not generalized disconjugate on $[-\infty, \infty]$ because the constant solutions of $y'' = 0$ have two zeros, one at ∞ and one at $-\infty$. A constant coefficient operator $D^2 + pD + q$ (p and q constants), is generalized disconjugate on $[-\infty, \infty]$ if and only if $p^2 - 4q > 0$, i.e., the characteristic numbers are real and distinct.

If L is disconjugate in the generalized sense on $[a, b]$, then there exist positive solutions u_1 and u_2 of $Ly = 0$ such that (u_1, u_2) is a principal system at b and (u_2, u_1) is a principal system at a. Any such pair (u_1, u_2) is called a *fundamental principal system* (f.p.s.) *on* $[a, b]$. In an f.p.s. (u_1, u_2), the functions u_1 and u_2 are each unique up to multiplication by positive constants. Thus, $(1, t)$ is an f.p.s. on $[0, \infty]$ of D^2, $(b - t, t - a)$ on $[a, b] \subset (-\infty, \infty)$ of D^2 and $(e^{-t}, e^t - e^{-t})$ on $[0, \infty]$ of $D^2 - 1$. For further discussion and extension of these ideas to higher order equations see Muldowney [13] or Willett [19].

The following two lemmas are consequences of the generalized disconjugacy assumption and will be needed to prove the theorem. In these lemmas we assume L is a generalized disconjugate second order linear operator on $[a, b]$ and $\mathscr{D}^0$ and $\mathscr{D}^1$ are defined in terms of an f.p.s. on $[a, b]$ of L by (1.2).

LEMMA 1. *If* $y \in C^1(a, b)$ *and* $\mathscr{D}^0 y(a)$, $\mathscr{D}^0 y(b)$ *exist, then there exists an* $\mathscr{E} \in [a, b]$ *such that*

$$\mathscr{D}^0 y(b) - \mathscr{D}^0 y(a) = \mathscr{D}^1 y(\mathscr{E}). \tag{2.2}$$

Proof. A generalized mean value theorem of Willett [19; Thm. 3.1, p. 1034] implies, for the differential equation

$$My = \frac{w(u, y)}{w(u_1, u_2)} = 0,$$

that there exists a point $\mathscr{E}_t \in (t, b)$ such that

$$y(t) = u(t)\mathscr{D}^0 y(b) - u_1(t) My(\mathscr{E}_t).$$

Since $\mathscr{D}^1 y(\mathscr{E}_t) = My(\mathscr{E}_t)$ and $\lim_{t \to a^+} u_2(t)/u_1(t) = 0$,

$$\lim_{t \to a^+} \mathscr{D}^1 y(\mathscr{E}_t) = \lim_{t \to a^+} \left[\frac{u(t)}{u_1(t)} \mathscr{D}^0 y(b) - \frac{y(t)}{u_1(t)} \right] = \mathscr{D}^0 y(b) - \mathscr{D}^0 y(a).$$

Hence, there must exist an $\mathscr{E}$ in the closed interval $[a, b]$ such that (2.2) holds.

LEMMA 2. *If* $y \in C^2(a, b)$, $y(t) > 0$ *for* $t \in (t_1, t_2) \subset (a, b)$ *and* $\mathscr{D}^0 y(t_1) = 0 = \mathscr{D}^0 y(t_2)$, *there exists an* $\mathscr{E} \in (t_1, t_2)$ *such that*

$$\mathscr{D}^1 y(E) = 0 \quad \textit{and} \quad Ly(\mathscr{E}) \leqq 0. \tag{2.3}$$

Proof. Since $y/u > 0$ on (t_1, t_2) and

$$\lim_{t \to t_2^-} \frac{y(t)}{u(t)} = 0 = \lim_{t \to t_1^+} \frac{y(t)}{u(t)},$$

there exists an $\mathscr{E} \in (t_1, t_2)$ at which y/u attains its maximum. Hence,

$$\left(\frac{y}{u}\right)'(\mathscr{E}) = 0,$$

which implies

$$\mathscr{D}^1 y(\mathscr{E}) = \frac{u^2(\mathscr{E})}{w(u_1, u_2)(\mathscr{E})}\left(\frac{y}{u}\right)'(\mathscr{E}) = 0;$$

and

$$\left(\frac{y}{u}\right)''(\mathscr{E}) \leqq 0,$$

which implies

$$Ly(\mathscr{E}) = \frac{w(u, u_2)}{u}\mathscr{D}\frac{u^2}{w(u, u_2)}\mathscr{D}\frac{y}{u}\Bigg|^{\mathscr{E}} = u(\mathscr{E})\left(\frac{y}{u}\right)''(\mathscr{E}) \leqq 0.$$

3. Generalized Nagumo condition. The generalized Nagumo condition defined in § 1 (cf. (1.4)–(1.7)) includes the classical Nagumo condition [16], for suppose $L = D^2$ and $[a, b] \subset (-\infty, \infty)$. Then, we can take $u_1(t) = b - t$, $u_2(t) = t - a$ and $u(t) = b - a = w(u_1, u_2)$. Hence, $F(t, y, z) = f(t, y, z)$ and (1.6) simply becomes

$$(b-a)|f(t, y, z)| \leqq \phi(|z|),$$

which is equivalent to the classical Nagumo condition.

To illustrate the nature of the generalized Nagumo condition and to indicate the direction where improvement is needed, consider the equation

$$y'' - y = r(t)yy', \qquad 0 < t < \infty. \tag{3.1}$$

Letting $L = D^2 - 1$ and $f(t, y, z) = r(t)yz$, we choose

$$u_1(t) = e^{-t} \quad \text{and} \quad u_2(t) = e^t - e^{-t},$$

so that

$$u(t) = e^t \quad \text{and} \quad w(u_1, u_2) = 2.$$

Hence, (1.6) becomes

$$|r(t)y(y + 2e^{-t}z)|e^{3t}/4 \leqq \phi(|z|), \qquad 0 < t < \infty. \tag{3.2}$$

So if $\gamma(t) = \max(|\beta(t)|, |\alpha(t)|)$, we are essentially required to assume $r(t)\gamma^2(t)\exp(3t)$ and $r(t)\gamma(t)\exp(2t)$ are bounded on $(0, \infty)$, which because of the exponential factors is a rather stringent, although quite natural, assumption on $r(t)$. The exponential factor comes from the factor $u^3(t)$ in (1.6) and could be improved if one had more freedom in choosing $u(t)$, e.g., if $u(t) = u_2(t) = e^{-t}$ were possible, then the factor $u^3(t)$ in (1.6) would be actually an asset. In general, $u(t) = c_1u_1(t) + c_2u_2(t)$ with $c_1, c_2 \geqq 0$ and not both zero is possible up to the final stages of the proof of the theorem in § 5 where technical difficulties arise unless $c_1 > 0$ and $c_2 > 0$; hence, our choice of $c_1 = 1 = c_2$ right from the beginning.

Assume now that f satisfies a generalized Nagumo condition on (a, b) with respect to α, β ((1.4)–(1.7)), and let the number N satisfy (1.8). Define

$$F_N(t, y, z) = \begin{cases} F(t, y, z) & \text{when } |z| \leqq N, \\ F\left(t, y, \frac{z}{|z|}N\right) & \text{when } |z| > N. \end{cases} \tag{3.3}$$

LEMMA 3. *Assume that f satisfies a generalized Nagumo condition on (a, b) with respect to α, β and*

$$Ly = F_N(t, y, \mathscr{D}^1 y), \qquad \alpha < t < b.$$

If

$$\alpha(t) \leqq y(t) \leqq \beta(t), \qquad \alpha < t < b,$$

then

$$|\mathscr{D}^1 y(t)| < N, \qquad \alpha < t < b,$$

i.e.,

$$Ly = f(t, y, y').$$

Proof. Suppose $|\mathscr{D}^1 y(t)| \geqq N$ at some point in (a, b). By Lemma 1, there exists an $\mathscr{E} \in [a, b]$ such that $|\mathscr{D}^1 y(\mathscr{E})| \leqq \lambda < N$. So by continuity there exist points $\tau \in (a, b)$ and $\sigma \in [a, b]$ such that $|\mathscr{D}^1 y(\tau)| = N$, $|\mathscr{D}^1 y(\sigma)| = \lambda$ and $\lambda \leqq |\mathscr{D}^1 y(t)| \leqq N$ for t between τ and σ. The proof now divides into four cases depending upon the sign of $\mathscr{D}^1 y(\tau)$ and whether τ is to the left or right of σ. Since the details of the cases are similar, we will do only one case explicitly.

Suppose that $\mathscr{D}^1 y(\tau) = N$, choose σ so that $\mathscr{D}^1 y(\sigma) = \lambda$ and suppose $\sigma > \tau$, i.e., $\lambda \leqq \mathscr{D}^1 y(t) \leqq N$ for $\tau < t < \sigma$. Consider $\tau < t < \sigma$. Since

$$F_N(t, y(t), \mathscr{D}^1 y(t)) = F(t, y(t), \mathscr{D}^1 y(t))$$

and

$$D\mathscr{D}^0 y(t) = \frac{w(u_1, u_2)(t)}{u^2(t)} \mathscr{D}^1 y(t) \geqq 0,$$

$$\mathscr{D}^1 y D\mathscr{D}^1 y = \frac{u^3(t)}{w^2(u_1, u_2)(t)} F(t, y, \mathscr{D}^1 y) D\mathscr{D}^0 y \geqq -\phi(|\mathscr{D}^1 y|) D\mathscr{D}^0 y.$$

Therefore, if $z = \mathscr{D}^1 y$, then

$$K < \int_\lambda^N \frac{z}{\phi(z)}\, dz = -\int_\tau^\sigma \frac{\mathscr{D}^1 y(t) D\mathscr{D}^1 y(t)}{\phi(\mathscr{D}^1 y(t))}\, dt \leqq \int_\tau^\sigma D\mathscr{D}^0 y(t)\, dt$$
$$= \mathscr{D}^0 y(\sigma) - \mathscr{D}^0 y(\tau) \leqq K,$$

which is a contradiction.

4. The boundary conditions. Figure 1 illustrates (1.12) in the general case; namely, Ω_0 must be a continuum which intersects or approaches the line segments composing Φ_0. A similar figure illustrates (1.13). These conditions are the same as those considered by Muldowney and Willett [14], which can be consulted to see

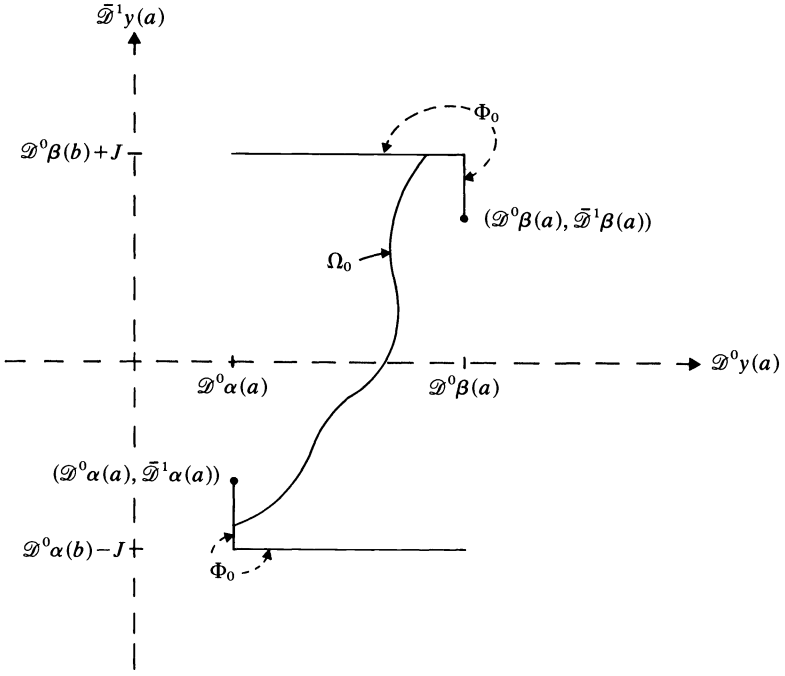


FIG. 1.

the various nonlinear boundary conditions in analytic form that are included in (1.12) and (1.13). Here we will just illustrate how (1.12) (and similarly (1.13)) includes the classical linear combination of solution and first derivative type of boundary condition.

Suppose the boundary condition at a $(a \geqq -\infty)$ is of the form

$$a_1\mathscr{D}^0 y(a) + a_2\bar{\mathscr{D}}^1 y(a) = A, \tag{4.1}$$

where a_1, a_2, A are real numbers. Thus, if $a_2 \neq 0$, Ω_0 is defined to be a straight line with slope $-a_1/a_2$ (see Fig. 2), and if $a_2 = 0$, Ω_0 is taken to be the vertical straight line $\mathscr{D}^0 y(a) = A/a_1$. In the cases $a_2 \neq 0$, Ω_0 will intersect or approach the upper part of Φ_0 if

$$\frac{A}{a_2} - \frac{a_1}{a_2}\mathscr{D}^0\beta(a) \geqq \bar{\mathscr{D}}^1\beta(a),$$

that is, if β satisfies

$$\frac{1}{a_2}[a_2\bar{\mathscr{D}}^1\beta(a) + a_1\mathscr{D}^0\beta(a) - A] \leqq 0. \tag{4.2}$$

Similarly, when $a_2 \neq 0$, Ω_0 intersects or approaches the lower part of Φ_0 if α satisfies

$$\frac{1}{a_2}[a_2\bar{\mathscr{D}}^1\alpha(a) + a_1\mathscr{D}^0\alpha(a) - A] \geqq 0. \tag{4.3}$$

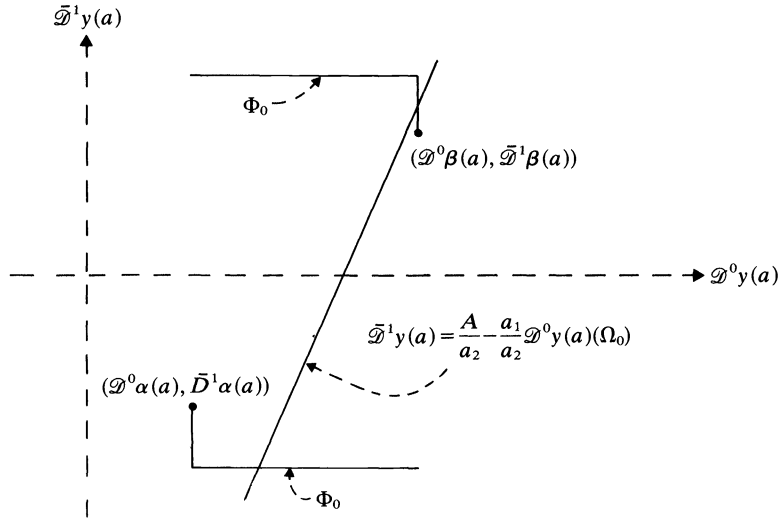


FIG. 2.

In case $a_2 = 0$, the corresponding conditions are

$$\mathscr{D}^0\alpha(a) \leqq A/a_1 \leqq \mathscr{D}^0\beta(a). \tag{4.4}$$

Conditions (4.2)–(4.4) are similar to the classical conditions for the boundary condition

$$a_1 y(a) + a_2 y'(a) = A$$

when a is a nonsingular point. Of course, when a is a nonsingular point, then

$$\mathscr{D}^0 y(a) = \frac{y(a)}{u_1(a)} \quad \text{and} \quad \bar{\mathscr{D}}^1 y(a) = \frac{u_1(a)y'(a) - u_1'(a)y(a)}{w(u_1, u_2)(a)},$$

and the classical conditions follow from (4.2)–(4.4) by direct substitution.

5. Proof of the theorem. Two further preliminary lemmas will be proven.

LEMMA 4. *If $y \in C^1(a, b)$, $Ly(t)$ is measurable and satisfies*

$$\lambda(t) \leqq Ly(t) \leqq \gamma(t), \qquad a < t < b,$$

with $\lambda u/(w(u_1, u_2))$, $\gamma u/(w(u_1, u_2)) \in L_1[a, b]$, and $\mathscr{D}^0 y(a)$, $\mathscr{D}^0 y(b)$, $\bar{\mathscr{D}}^1 y(a)$ and $\bar{\mathscr{D}}^1 y(b)$ exist, then

$$\bar{\mathscr{D}}^1 y(a) \leqq \mathscr{D}^0\alpha(b) - \int_a^b \frac{u_1(s)\gamma(s)}{w(u_1, u_2)(s)}\,ds \quad \textit{implies} \quad \mathscr{D}^0 y(b) \leqq \mathscr{D}^0\alpha(b); \tag{5.1}$$

$$\bar{\mathscr{D}}^1 y(a) \geqq \mathscr{D}^0\beta(b) - \int_a^b \frac{u_1(s)\lambda(s)}{w(u_1, u_2)(s)}\,ds \quad \textit{implies} \quad \mathscr{D}^0 y(b) \geqq \mathscr{D}^0\beta(b); \tag{5.2}$$

$$\mathscr{D}^0 y(a) - \int_a^b \frac{u_2(s)\gamma(s)}{w(u_1, u_2)(s)}\,ds \leqq \bar{\mathscr{D}}^1 y(b) \leqq \mathscr{D}^0 y(a) - \int_a^b \frac{u_2(s)\lambda(s)}{w(u_1, u_2)(s)}\,ds. \tag{5.3}$$

Proof. From

$$y(t)=\mathscr{D}^0y(a)u_1(t)+\bar{\mathscr{D}}^1y(a)u_2(t)+\int_a^t g(t,s)Ly(s)\,ds,$$

where

$$g(t,s)=[u_2(t)u_1(s)-u_1(t)u_2(s)]/w(u_1,u_2)(s) \tag{5.4}$$

is the Cauchy function for L, we obtain

$$\begin{aligned}\mathscr{D}^0y(b)=\lim_{t\to b^-}\frac{y(t)}{u_2(t)}&=\bar{\mathscr{D}}^1y(a)+\int_a^b\frac{u_1(s)Ly(s)}{w(u_1,u_2)(s)}\,ds\\&\leqq\bar{\mathscr{D}}^1y(a)+\int_a^b\frac{u_1(s)\gamma(s)}{w(u_1,u_2)(s)}\,ds,\end{aligned}$$

which proves (5.1). One proves (5.2) similarly.

To obtain (5.3), we consider

$$y(t)=\bar{\mathscr{D}}^1y(b)u_1(t)+\mathscr{D}^0y(b)u_2(t)-\int_t^b g(t,s)Ly(s)\,ds. \tag{5.5}$$

Then

$$\begin{aligned}\mathscr{D}^0y(a)=\lim_{t\to a^+}\frac{y(t)}{u_1(t)}&=\bar{\mathscr{D}}^1y(b)+\int_a^b\frac{u_2(s)}{w(u_1,u_2)(s)}Ly(s)\,ds\\&\leqq\bar{\mathscr{D}}^1y(b)+\int_a^b\frac{u_2(s)\gamma(s)}{w(u_1,u_2)(s)}\,ds,\end{aligned}$$

from which the first inequality in (5.3) follows. One can obtain the second inequality in (5.3) similarly from (5.5).

Let $k(t)$ be any positive continuous function on (a,b) such that

$$\begin{aligned}\frac{\pi}{2}\int_a^b\frac{u_2(s)k(s)}{w(u_1,u_2)(s)}\,ds&\leqq\int_a^b\frac{u_1(s)\psi(s)}{w(u_1,u_2)(s)}\,ds,\\ \frac{\pi}{2}\int_a^b\frac{u_1(s)k(s)}{w(u_1,u_2)(s)}\,ds&\leqq\int_a^b\frac{u_2(s)\psi(s)}{w(u_1,u_2)(s)}\,ds.\end{aligned} \tag{5.6}$$

Define

$$\bar{F}_N(t,y,z)=\begin{cases}F_N(t,\beta(t),z)+k(t)\,\mathrm{Tan}^{-1}(y-\beta(t)) & \text{for } y>\beta(t),\\ F_N(y,y,z) & \text{for } \alpha(t)\leqq y\leqq\beta(t),\\ F_N(t,\alpha(t),z)-k(t)\,\mathrm{Tan}^{-1}(\alpha(t)-y) & \text{for } y<\alpha(t),\end{cases}$$

for $\alpha\leqq\beta$.

LEMMA 5. *Assume that* $\alpha,\ \beta \in C^2(a, b)$, $\alpha \leqq \beta$, $L\alpha \geqq f(t, \alpha, \alpha')$, $L\beta \leqq f(t, \beta, \beta')$, $\mathscr{D}^0\alpha(a)$, $\mathscr{D}^0\alpha(b)$, $\mathscr{D}^0\beta(\alpha)$, $\mathscr{D}^0\beta(b)$ *exist and* $N > \max(|\mathscr{D}^1\alpha(t)|, |\mathscr{D}^1\beta(t)|)$ *for* $a < t < b$. *If*

$$
\begin{aligned}
&Ly = \bar{F}_N(t, y, \mathscr{D}^1 y), \qquad && \mathscr{D}^0\alpha(a) \leqq \mathscr{D}^0 y(a) \leqq \mathscr{D}^0\beta(a),\\
&&& \mathscr{D}^0\alpha(b) \leqq \mathscr{D}^0 y(b) \leqq \mathscr{D}^0\beta(b),
\end{aligned} \tag{5.7}
$$

then

$$\alpha \leqq y \leqq \beta; \quad \textit{i.e.,} \quad Ly = F_N(t, y, \mathscr{D}^1 y).$$

Proof. Assume $y(t) > \beta(t)$ for some $t \in (a, b)$. Then there exists an interval $[t_1, t_2] \subset [a, b]$ such that $z(t) = y(t) - \beta(t) > 0$ for $t \in (t_1, t_2)$ and $\mathscr{D}^0 z(t_1) = 0 = \mathscr{D}^0 z(t_2)$. Hence, Lemma 2 implies there exists an $\mathscr{E} \in (t_1, t_2)$ such that

$$\mathscr{D}^1 z(\mathscr{E}) = 0 \quad \text{and} \quad Lz(\mathscr{E}) \leqq 0.$$

But

$$
\begin{aligned}
Lz(\mathscr{E}) &= Ly(\mathscr{E}) - L\beta(\mathscr{E}) \geqq \bar{F}_N(\mathscr{E}, y(\mathscr{E}), \mathscr{D}^1 y(\mathscr{E})) - f(\mathscr{E}, \beta(\mathscr{E}), \beta'(\mathscr{E}))\\
&= F_N(\mathscr{E}, \beta(\mathscr{E}), \mathscr{D}^1\beta(\mathscr{E})) + k(\mathscr{E})\,\mathrm{Tan}^{-1}(z(\mathscr{E})) - f(\mathscr{E}, \beta(\mathscr{E}), \beta'(\mathscr{E}))\\
&= k(\mathscr{E})\,\mathrm{Tan}^{-1} z(\mathscr{E}) > 0,
\end{aligned}
$$

which is a contradiction. Therefore, $\beta \geqq y$; and similarly, $\alpha \leqq y$.

Proof of theorem. Let

$$
\begin{aligned}
\psi_{00} &= \{(x, \mathscr{D}^0\beta(b) + J) : \mathscr{D}^0\alpha(a) \leqq x \leqq \mathscr{D}^0\beta(a)\},\\
\psi_{01} &= \{(x, \mathscr{D}^0\alpha(b) - J) : \mathscr{D}^0\alpha(a) \leqq x \leqq \mathscr{D}^0\beta(a)\},\\
\psi_{10} &= \{(x, \mathscr{D}^0\beta(a) + J) : \mathscr{D}^0\alpha(b) \leqq x \leqq \mathscr{D}^0\beta(b)\},\\
\psi_{11} &= \{(x, \mathscr{D}^0\alpha(a) - J) : \mathscr{D}^0\alpha(b) \leqq x \leqq \mathscr{D}^0\beta(b)\}.
\end{aligned}
$$

We consider first the case where $\Omega_0 \cup \psi_{00} \cup \psi_{01}$ and $\Omega_1 \cup \psi_{10} \cup \psi_{11}$ are both connected.

Let N be chosen as in (1.8) and consider the equation

$$Lz = \bar{F}_N(t, z, \mathscr{D}^1 z), \qquad \alpha < t < b.$$

For ε_0 sufficiently small, let $0 < \varepsilon < \varepsilon_0$ and a_ε be a function such that $b > a_\varepsilon > a$ and $a_\varepsilon \downarrow a$, as $\varepsilon \downarrow 0$. Let $\tau_\varepsilon = \min(a_\varepsilon - a, \varepsilon)$, and define

$$
z(t) = \begin{cases} pu_1(t) + qu_2(t) & \text{for } a < t < a_\varepsilon,\\ pu_1(t) + qu_2(t) + \displaystyle\int_{a_\varepsilon}^{t} g(t, s)\bar{F}_N(s, z(s - \tau_\varepsilon), \mathscr{D}^1 z(s - \tau_\varepsilon))\, ds & \text{for } a_\varepsilon \leqq t < b, \end{cases} \tag{5.8}
$$

where $g(t, s)$ is the Cauchy function for L (see (5.4)) and $(p, q) \in R^2$. For each $(p, q) \in R^2$, (5.8) defines a unique function $z(t, p, q) \equiv z(t) \in C(a, b) \cap C^1(a_\varepsilon, b)$. Define a mapping $T \equiv T_\varepsilon : R^2 \to R^2$ by

$$T(p, q) = (\mathscr{D}^0 z(b; p, q), \bar{\mathscr{D}}^1 z(b; p, q)),$$

and consider $T\Omega_0$. Since

$$Lz(t)=\begin{cases}\bar{F}_N(t, z(t-\tau_\varepsilon), \mathscr{D}^1 z(t-\tau_\varepsilon)), & a_\varepsilon < t < b,\\ 0, & a < t < a_\varepsilon,\end{cases}$$

we have

$$|Lz(t)| \leqq \psi(t)+(\pi/2)k(t), \qquad a<t<b, \quad t \neq a_\varepsilon.$$

Since $p=\mathscr{D}^0 z(a)$ and $q=\bar{\mathscr{D}}^1 z(a)$, Lemma 4 and (5.6) imply

(5.9) $\quad \mathscr{D}^0\alpha(a)-J \leqq \bar{\mathscr{D}}^1 z(b) \leqq \mathscr{D}^0\beta(a)+J$ if $\mathscr{D}^0\alpha(a) \leqq p \leqq \mathscr{D}^0\beta(a)$,

(5.10) $\quad \mathscr{D}^0 z(b) \geqq \mathscr{D}^0\beta(b)$ if $q \geqq \mathscr{D}^0\beta(b)+J$,

(5.11) $\quad \mathscr{D}^0 z(b) \leqq \mathscr{D}^0\alpha(b)$ if $q \leqq \mathscr{D}^0\alpha(b)-J$.

Since $(p, q) \in \Omega_0$ implies $\mathscr{D}^0\alpha(a) \leqq p \leqq \mathscr{D}^0\beta(a)$ and $\mathscr{D}^0\alpha(b)-J \leqq q \leqq \mathscr{D}^0\beta(b)+J$, we conclude that $T\Omega_0$ is a continuum in the strip $\mathscr{D}^0\alpha(a)-J \leqq \bar{\mathscr{D}}^1 z(b) \leqq \mathscr{D}^0\beta(a)+J$ which intersects the lines $\mathscr{D}^0 z(b)=\mathscr{D}^0\alpha(b)$ and $\mathscr{D}^0 z(b)=\mathscr{D}^0\beta(b)$ (see Fig. 3). By a result of Muldowney and Willett [14, Lemma 3, p. 702], $T\Omega_0 \cap \Omega_1 \neq \phi$, i.e. there exists $(p_\varepsilon, q_\varepsilon) \in \Omega_0$ such that $(\mathscr{D}^0 z(b; p_\varepsilon, q_\varepsilon), \bar{\mathscr{D}}^1 z(b; p_\varepsilon, q_\varepsilon)) \in \Omega_1$. Let

$$z_\varepsilon(t) \equiv z(t; p_\varepsilon, q_\varepsilon).$$

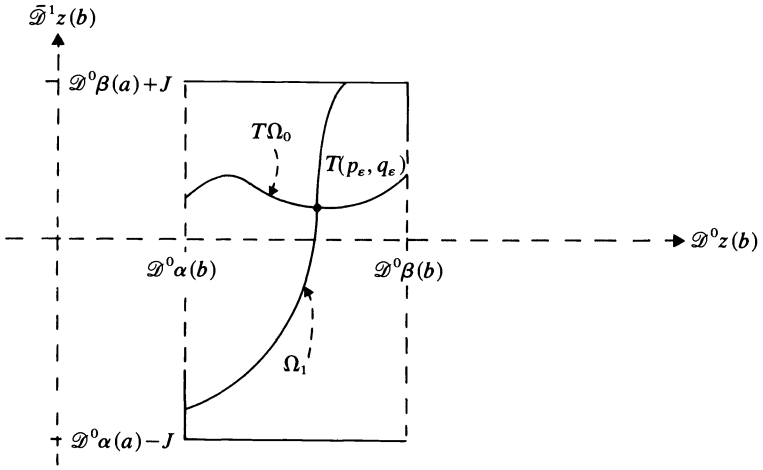


FIG. 3.

Since

$$\mathscr{D}^1 z_\varepsilon(t)=\begin{cases}q_\varepsilon-p_\varepsilon & \text{for } a<t<a_\varepsilon,\\ q_\varepsilon-p_\varepsilon+\displaystyle\int_{a_\varepsilon}^{t}\frac{u(s)}{w(u_1,u_2)(s)}\bar{F}_N(s, z_\varepsilon(s-\tau_\varepsilon), \mathscr{D}^1 z_\varepsilon(s-\tau_\varepsilon))\,ds & \\ & \text{for } a_\varepsilon \leqq t<b,\end{cases}$$

$(p_\varepsilon, q_\varepsilon) \in \Omega_0$, which is bounded, and $|\bar{F}_N| \leqq \psi$ with $\psi u/w(u_1, u_2)$ integrable on $[a, b]$, it is the case that the sets $\{\mathscr{D}^0 z_\varepsilon : 0<\varepsilon<\varepsilon_0\}$ and $\{\mathscr{D}^1 z_\varepsilon : 0<\varepsilon<\varepsilon_0\}$ are

uniformly bounded and equicontinuous subsets of $C[a, b]$. By the Arzelà–Ascoli theorem and the compactness of Ω_0, there exists a sequence $\varepsilon_k \to 0$ $(k \to \infty)$, functions $z_1, z_2 \in C[a, b]$ and a point $(p_0, q_0) \in \Omega_0$ such that

$$\mathscr{D}^0 z_{\varepsilon_k} \to z_1, \qquad \mathscr{D}^1 z_{\varepsilon_k} \to z_2, \qquad (p_{\varepsilon_k}, q_{\varepsilon_k}) \to (p_0, q_0).$$

Therefore, for $a < t < b$,

$$z_1(t)u(t) = p_0 u_1(t) + q_0 u_2(t) + \int_a^t g(t, s)\bar{F}_N(s, z_1(s)u(s), z_2(s))\, ds, \tag{5.12}$$

$$z_2(t) = q_0 - p_0 + \int_a^t \frac{u(s)}{w(u_1, u_2)(s)} \bar{F}_N(s, z_1(s)u(s), z_2(s))\, ds. \tag{5.13}$$

But (5.12) and (5.13) imply $\mathscr{D}^1(uz_1) = z_2$; hence, let $y = uz_1$. Then,

$$Ly = \bar{F}_N(t, y(t), \mathscr{D}^1 y(t)), \qquad a < t < b,$$
$$(\mathscr{D}^0 y(a), \bar{\mathscr{D}}^1 y(a)) = (p_0, q_0) \in \Omega_0,$$

and since

$$\lim_{k\to\infty} \bar{\mathscr{D}}^1 z(b; p_{\varepsilon_k}, q_{\varepsilon_k}) = \lim_{k\to\infty} \left(p_{\varepsilon_k} - \int_{a_{\varepsilon_k}}^b \frac{u_2}{w(u_1, u_2)} \cdot \bar{F}_N(s, z_{\varepsilon_k}(s - \tau_{\varepsilon_k}), \mathscr{D}^1 z(s - \tau_{\varepsilon_k}))\, ds \right)$$
$$= p_0 - \int_a^b \frac{u_2}{w} \bar{F}_N(s, uz_1, z_2)\, ds = \bar{\mathscr{D}}^1 y(b)$$

and Ω_1 is compact,

$$(\mathscr{D}^0 y(b), \bar{\mathscr{D}}^1 y(b)) \in \Omega_1.$$

Clearly, (5.7) is satisfied; so Lemma 5 implies $\alpha \leqq y \leqq \beta$; hence, Lemma 3 implies $|\mathscr{D}^1 y| < N$; and so finally

$$Ly = f(t, y, y'), \tag{5.14}$$

which proves the theorem for this case.

Now consider the case when $\Omega_0 \cup \psi_{00} \cup \psi_{01}$ or $\Omega_1 \cup \psi_{10} \cup \psi_{11}$ are not connected. There are several possibilities here, and they are all handled similarly. We give the details for just one case, namely, assume $\Omega_0 \cup \psi_{01}$ is not connected but $\Omega_0 \cup \psi_{00}$ and $\Omega_1 \cup \psi_{10} \cup \psi_{11}$ are connected. Since $\Omega_0 \cup \Phi_0$ is connected by assumption, it must be the case that $\gamma = \Omega_0 \cap \{(\mathscr{D}^0\alpha(a), y) : \mathscr{D}^0\alpha(b) - J \leqq y \leqq \bar{\mathscr{D}}^1\alpha(a)\} \neq \varnothing$. Let $\mu = \sup\{y : (\mathscr{D}^0\alpha(a), y) \in \gamma\}$ and $\Omega_0^* = \Omega_0 \cup \{(\mathscr{D}^0\alpha(a), y) : \mathscr{D}^0\alpha(b) - J \leqq y \leqq \mu\}$. It is now the case that $\Omega_0^* \cup \psi_{00} \cup \psi_{01}$ is connected. Hence, by the part of the theorem already established, there exists a solution y of (5.14) such that $(\mathscr{D}^0 y(a), \bar{\mathscr{D}}' y(a)) \in \Omega_0^*$, $(\mathscr{D}^0 y(b), \bar{\mathscr{D}}' y(b)) \in \Omega_1$ and $\alpha \leqq y \leqq \beta$. We will show that necessarily $(\mathscr{D}^0 y(a), \mathscr{D}' y(a)) \in \Omega_0$. Suppose $\mathscr{D}^0 y(a) = \mathscr{D}^0\alpha(a)$ and $\mathscr{D}^0\alpha(b) - J \leqq \bar{\mathscr{D}}^1 y(a) \leqq \bar{\mathscr{D}}^1\alpha(a)$. Then

$$0 \leqq y(t) - \alpha(t) = \bar{\mathscr{D}}^1(y - \alpha)(a)u_2(t) + \int_a^t g(t, s)L(y - \alpha)(s)\, ds$$

or

$$0 \leqq \bar{\mathscr{D}}^1(y-\alpha)(a) + \lim_{t\to a^+} \frac{1}{u_2(t)} \int_a^t g(t,s)L(y-\alpha)(s)\,ds = \bar{\mathscr{D}}^1(y-\alpha)(a).$$

Hence, it must be the case that

$$\bar{\mathscr{D}}^1 y(a) = \bar{\mathscr{D}}^1 \alpha(a) = \mu,$$

and $(\mathscr{D}^0 y(a), \bar{\mathscr{D}}^1 y(a)) = (\mathscr{D}^0 \alpha(a), \bar{\mathscr{D}}^1 \alpha(a)) \in \Omega_0$.

6. Examples. Consider first the problem

$$y'' - y = r(t)yy', \quad y(0) = 0, \quad \lim_{t\to\infty} e^{-t}y(t) = 1, \tag{6.1}$$

where $r \in C[0, \infty)$ and $r \geqq 0$. Let $L = D^2 - 1$, $(u_1, u_2) = (e^{-t}, e^t - e^{-t})$, $u = e^t$, $\alpha = 0$ and $\beta = e^t$. Since

$$\frac{u^3(t)}{w^2(u_1, u_2)(t)} |F(t, y, z)| = \frac{e^{3t}}{4} |r(t)|\,|y|\,|y + 2e^{-t}z|,$$

the generalized Nagumo condition (1.6), as well as all the other conditions of the theorem, will be satisfied for $(\alpha, \beta) = (0, e^t)$ provided

$$r(t) = O(e^{-5t}) \quad \text{as } t \to \infty. \tag{6.2}$$

Thus, (6.1) has a solution if (6.2) holds. Similarly, one can conclude that

$$y'' = r(t)y, \qquad 0 \leqq t < \infty,$$

has a solution such that

$$\lim_{t\to\infty} t^{-1}y(t) = 1$$

provided

$$r \geqq 0 \quad \text{and} \quad r = O(t^{-4}) \quad \text{as } t \to \infty, \tag{6.3}$$

which indicates a certain lack of sharpness in the application of our results since

$$r = O(t^{-2-\varepsilon}) \quad \text{for some } \varepsilon > 0 \tag{6.4}$$

is all that is needed in this case. We will show in another paper on asymptotics how the general method can be adapted to obtain sharp asymptotic results for this and other problems without linearity restrictions.

Consider next the equation

$$t^2y'' + aty' + by = r(t)yy', \qquad 0 < t < 1, \tag{6.5}$$

where a and b are constants such that $(a-1)^2 > 4b$ and $r(t) = O(t^\mu)$, as $t \to 0+$. Then the Euler equation

$$Ly \equiv y'' + at^{-1}y' + bt^{-2}y = 0$$

is disconjugate in a neighborhood of 0 and has the fundamental set of solutions t^{λ_1}, t^{λ_2} where

$$2\lambda_2 = 1-\alpha+\sqrt{(a-1)^2-4b}, \qquad 2\lambda_1 = 1-a-\sqrt{(a-1)^2-4b}.$$

By taking $a(t)=0$ and

$$\beta(t) = Ct^{\lambda_1}(1+t^{\sigma}) \qquad (C>0),$$

where σ is chosen so that $0<\sigma<\min(\mu+\lambda_1-1, \sqrt{(a-1)^2-4b})$, the theorem implies that (6.5) always has a solution $y(t)$ such that

$$\lim_{t\to 0^+} t^{-\lambda_1} y(t) = C$$

provided

$$\lambda_1+\mu-1>0.$$

REFERENCES

[1] K. AKÔ, *Subfunctions for ordinary differential equations VI*, J. Fac. Sci. Univ. Tokyo Sect. I, 16 (1969), pp. 149–156.

[2] J. W. BEBERNES, AND R. FRAKER, *A priori bounds for boundary sets*, Proc. Amer. Math. Soc., 29 (1971), pp. 313–318.

[3] J. W. BEBERNES AND R. GAINES, *A generalized two point boundary value problem*, Proc. Amer. Math. Soc., 19 (1968), pp. 749–754.

[4] J. W. BEBERNES AND L. K. JACKSON, *Infinite interval boundary value problems for* $y''=f(x, y)$, Duke Math. J., 34 (1967), pp. 39–48.

[5] J. W. BEBERNES AND R. WILHELMSEN, *A general boundary value problem technique*, J. Differential Equations, 8 (1970), pp. 404–415.

[6] N. I. BRIŠ, *On boundary value problems for the equation* $\varepsilon y''=f(t, y, y')$ *for small* ε, Dokl. Akad. Nauk SSSR, 95 (1954), pp. 429–432.

[7] L. H. ERBE, *Nonlinear boundary value problems for second order ordinary differential equations*, J. Differential Equations, 7 (1970), pp. 459–472.

[8] T. C. LEE, *Boundary value problems for second order ordinary differential equations and applications*, Ph.D. thesis, Univ. Utah, Salt Lake City, 1975.

[9] L. FOUNTAIN AND L. K. JACKSON, *A generalized solution of the boundary value problem for* $y''=f(x, y, y')$, Pacific J. Math., 12 (1962), pp. 1251–1272.

[10] F. A. HOWES, *Singular perturbations and differential inequalities*, Ph.D. thesis, Univ. Southern Calif., Los Angeles, 1974.

[11] L. K. JACKSON, *Subfunctions and second-order ordinary differential inequalities*, Advances in Math., 2 (1968), pp. 307–368.

[12] H. W. KNOBLOCH, *Second order differential inequalities and a nonlinear boundary value problem*, J. Differential Equations, 5 (1969), pp. 55–71.

[13] J. S. MULDOWNEY, *On an inequality of Caplygin*, Proc. Roy. Irish Acad. Sect. A, to appear.

[14] J. S. MULDOWNEY AND D. WILLETT, *An elementary proof of the existence of solutions to second order nonlinear boundary value problems*, this Journal, 5 (1974), pp. 701–707.

[15] ———, *An intermediate value property for operators with applications to integral and differential equations*, Canad. J. Math., 26 (1974), pp. 27–41.

[16] M. NAGUMO, *Über die Differential gleichung* $y''=f(x, y, y')$, Proc. Phys. Math. Soc. Japan, 19 (1937), pp. 861–866.

[17] K. W. SCHRADER, *A note on second order boundary value problem*, Amer. Math. Monthly, 75 (1968), pp. 867–869.

[18] ———, *Boundary value problems for second order ordinary differential equations*, J. Differential Equations, 3 (1967), pp. 340–413.

[19] D. WILLETT, *A generalization of Caplygin's inequality with applications to singular boundary value problems*, Canad. J. Math., 25 (1973), pp. 1024–1039.

[20] ———, *Classification of second order linear differential equations with respect to oscillation*, Advances in Math., 3 (1969), pp. 594–623.

[21] ———, *Oscillation on finite intervals of second order linear differential equations*, Canad. Math. Bull., 14 (1971), pp. 539–550.

[22] ———, *Disconjugacy tests for singular linear differential equations*, this Journal, 2 (1971), pp. 536–545; *Errata*, this Journal, 3 (1972), p. 559.

SIAM J. MATH. ANAL.
Vol. 8, No. 5, October 1977

A NOTE ON POSITIVE REAL FUNCTIONS*

J. L. GOLDBERG AND J. L. ULLMAN†

Abstract. In this note we investigate properties of certain special classes of analytic functions—those which map the right half-plane into itself and the real axis into the real axis. These so-called "positive real" functions play a significant role in the synthesis of two-terminal, passive networks. Our concern is with the Taylor coefficients in the expansion about the point $z_0 = x_0 + iy_0$, $x_0 > 0$. If

$$f(z) = u + iv = \sum \alpha_n (z - z_0)^n,$$

then it is well known that $|\alpha_n| \leqq u_0 x_0^{-n}$, where $f(z_0) = u_0 + iv_0$. We derive results on $\arg \alpha_n$, one such being:

If f is a nonlinear positive real function and x is a real, positive number, then for every integer $N \geqq 1$ there exist integers $n \geqq N$ and $m \geqq N$ such that $f^{(n)}(x)$ and $f^{(m)}(x)$ have opposite signs.

1. Introduction. Suppose f is a nonconstant single-valued, analytic function in the open right half-plane. If f maps the open right half-plane into itself, we say f is a *positive* function (or more briefly; f is positive) and denote this by $f \in P$. If f is positive and if in addition maps the positive real axis into itself, we say f is *positive real* and denote this by $f \in PR$. For the purposes of notational convention, we formulate these definitions as follows.

DEFINITION 1.1. Let $z = x + iy$ and $f(z) = u(z) + iv(z)$. Then $f \in P$ if

(a) f is nonconstant, single-valued and analytic in $x > 0$

and

(b) $u > 0$ for $x > 0$.

DEFINITION 1.2. Let $f \in P$. Then $f \in PR$ if

(c) $f(z)$ is real for $z = x > 0$.

Besides their mathematical interest, the class of positive real functions has been extensively studied because of its importance in the synthesis of two-terminal passive networks (see [1], [2]).

Let $z_0 = x_0 + iy_0$, $x_0 > 0$ and $f(z_0) = u_0 + iv_0$. Then, if f is positive, f is analytic at z_0 and we have

$$f(z) = \sum_{n=0}^{\infty} \alpha_n (z - z_0)^n \tag{1.1}$$

valid in a circle at least as large as $|z - z_0| < x_0$. It was shown in [4] and [5] that

$$\left| \frac{f^n(z_0)}{n!} \right| = |\alpha_n| \leqq \frac{u_0}{x_0^n}, \qquad n = 1, 2, \cdots, \tag{1.2}$$

a constraint on the growth of the modulus of the derivatives of f. In this note we derive constraints on the argument of $f^{(k)}(z_0)$. As corollaries to the basic argument, we shall derive a number of interesting results on positive real functions—see, in particular, Theorems 3.3 and 3.4.

We conclude the Introduction with the statement of a well-known theorem on positive functions. The theorem has a long history and many proofs (see [3]).

* Received by the editors January 28, 1976, and in revised form September 8, 1976.

† Department of Mathematics, The University of Michigan, Ann Arbor, Michigan 48104.

THEOREM 1.3. *If $f \in P$, then*

$$\lim_{z\to\infty} f(z)/z = \lim_{z\to\infty} u(z)/x = \lim_{z\to\infty} f'(z) = A \geqq 0 \tag{1.3}$$

uniformly in the wedge, $|\arg z| \leqq \delta < \pi/2$. *Moreover, if f is nonlinear, then* $\tilde{f}(z) = f(z) - Az$ *is a positive function.*

We shall find it convenient to use this theorem for normalizing positive and positive real functions.

DEFINITION 1.4. If f is positive, then f is *normalized* if $\lim_{z\to\infty} f'(z) = A = 0$ for $|\arg z| \leqq \delta < \pi/2$.

Note that $\tilde{f}(z) = f(z) - Az$ is positive real (or constant) if f is positive real, since A is a real constant.

2. Some preliminaries and a fundamental lemma. Consider the linear fractional transformation

$$t = \frac{z - z_0}{z + \bar{z}_0} \tag{2.1}$$

and its inverse

$$z = \frac{z_0 + \bar{z}_0 t}{1 - t}, \tag{2.2}$$

where $z_0 = x_0 + iy_0$, $x_0 > 0$ and $\bar{z}_0 = x_0 - iy_0$. The first transformation maps the right half-plane onto the unit disk while the second transformation is a mapping of the unit disk onto the right half-plane. Write $f(z_0) = u_0 + iv_0$ and note that if f is positive, so is

$$F(z) = u_0^{-1}(f(z) - iv_0) = U + iV$$

since $U = u_0^{-1}u > 0$ for $x_0 > 0$ and $x > 0$.

Define $g(t)$ by the equation

$$g(t) = F\left(\frac{z_0 + \bar{z}_0 t}{1 - t}\right). \tag{2.3}$$

Then (a) $g(0) = F(z_0) = 1$ and (b) $g(t)$ is analytic in $|t| < 1$ and has positive real part there. From a theorem of Carathéodory [3],

$$g^{(n)}(0) = \frac{n!(2x_0)}{u_0} \sum_{k=1}^{n} \binom{n-1}{k-1} (2x_0)^{k-1} \frac{f^{(k)}(z_0)}{k!} \tag{2.4}$$

and

$$|g^{(n)}(0)| \leqq 2n!, \qquad n \geqq 1. \tag{2.5}$$

We are now in a position to prove the fundamental lemma. Recall that

$$f(z) = \sum_{n=0}^{\infty} \alpha_n (z - z_0)^n \tag{2.6}$$

for $|z - z_0| < x_0$, $z_0 = x_0 + iy_0$, $x_0 > 0$.

LEMMA 2.1. *If f is positive and if there exists a point z_0 in the open right half-plane and an integer $N \geqq 1$ such that the set $\{\alpha_k\}$, $k \geqq N$, lies in the same half-plane whose boundary is a line through the origin, then the numbers $\{\alpha_k\}$, $k \geqq N$, all lie on the boundary of this half-plane.*

Proof. Suppose $\alpha_k = r_k e^{i\theta_k}$, and suppose the half-plane of the hypothesis is defined by

$$\psi \leqq \theta_k \leqq \psi + \pi, \tag{2.7}$$

where we take $\theta_k = \psi$ if $\alpha_k = 0$. We write (2.4) in the form

$$\sum_{k=N}^{n} \binom{n-1}{k-1} (2x_0)^k r_k e^{i\theta_k} = \frac{u_0 g^{(n)}(0)}{n!} + \sum_{k=1}^{N-1} \binom{n-1}{k-1} (2x_0)^k (-\alpha_k) \tag{2.8}$$

for $1 < N < n$. If we multiply (2.8) by $e^{-i\psi}$, then take imaginary parts of both sides, then increase the right-hand side by replacing the imaginary part of $e^{-i\psi}(\alpha_k)$ by $|\alpha_k|$, we obtain with the aid of inequality (2.5),

$$\sum_{k=N}^{n} \binom{n-1}{k-1} (2x_0)^k r_k \sin(\theta_k - \psi) \leqq 2u_0 + \sum_{k=1}^{N-1} \binom{n-1}{k-1} (2x_0)^k |\alpha_k|. \tag{2.9}$$

Since $\psi \leqq \theta_k \leqq \psi + \pi$ for $k \geqq N$, $\sin(\theta_k - \psi) \geqq 0$ and thus each term on the left-hand side of (2.9) is nonnegative. Hence, for each j, $2 \leqq N \leqq j \leqq n$,

$$\begin{aligned} 0 &\leqq \binom{n-1}{j-1} (2x_0)^j r_j \sin(\theta_j - \psi) \\ &\leqq 2u_0 + \sum_{k=1}^{N-1} \binom{n-1}{k-1} (2x_0)^k |\alpha_k|. \end{aligned}$$

For fixed $j > N$, $1 \leqq k \leqq N-1$, this leads to

$$\begin{aligned} 0 &\leqq (2x_0)^j r_j \sin(\theta_j - \psi) \\ &\leqq \lim_{n\to\infty} \left\{ 2u_0 \binom{n-1}{j-1}^{-1} \right\} \\ &\quad + \sum_{k=1}^{N-1} (2x_0)^k |\alpha_k| \lim_{n\to\infty} \left\{ \binom{n-1}{k-1} \binom{n-1}{j-1}^{-1} \right\} \\ &= 0. \end{aligned}$$

Hence, $r_j \sin(\theta_j - \psi) = 0$. Thus either $\alpha_j = 0$ or $\theta_j - \psi$ is a multiple of 2π. In either case, α_j lies on the line through the origin with slope $\tan \psi$. This is the boundary of the half-plane.

Prompted by this lemma we introduce the notation $f \in P[z_0, \psi, N]$ to mean that $f \in P$ and there exists z_0 in the open right half-plane and an integer $N \geqq 0$ such that $\psi \leqq \arg \alpha_k \leqq \psi + \pi$ for all $k \geqq N$. The argument of zero is taken as ψ, by convention.

LEMMA 2.2. *If $f \in P[z_0, \psi, N]$, $N \geqq 2$, then $f \in P[z_0, \psi, 2]$.*

Proof. By Lemma 2.1, the numbers $\alpha_N, \alpha_{N+1}, \cdots$, all lie on a line through the origin. Hence $\alpha_{N-1}, \alpha_N, \alpha_{N+1}, \cdots$ lie in some half-plane with this line as boundary. Hence, by Lemma 2.1, the numbers $\alpha_{N-1}, \alpha_N, \alpha_{N+1}, \cdots$ lie on the boundary. That is, $f \in P[z_0, \psi, N-1]$. The argument may be repeated as long as $N>2$.

LEMMA 2.3. *If* $f \in P[z_0, \psi, N]$ *for* $z_0 = x_0 + iy_0$, $x_0 > 0$ *and* $N \geqq 2$, *then* $f \in P[z_0+s, \psi, 2]$ *for all real* $s > -x_0$.

Proof. By Lemma 2.2, $f \in P[z_0, \psi, 2]$. Suppose $-x_0 < s < x_0$. Since f is analytic in $x>0$,

$$f^{(k)}(z) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} \alpha_n (z - z_0)^{n-k}$$

for $k \geqq 2$ and for all z, $|z - z_0| < x_0$. Let $z = z_0 + s$. Then $(z - z_0)^{n-k} = s^{n-k}$ is real and $f^{(k)}(z_0+s)$ is the sum of co-linear vectors. Thus $f^{(k)}(z_0+s)$ lies on the same line as $\alpha_2, \alpha_3, \cdots$. A standard analytic continuation argument enables the removal of the restriction $s < x_0$ and this completes the proof.

Finally;

LEMMA 2.4. *If* f *is a normalized positive function such that* $f \in P[z_0, \psi, N]$, $N \geqq 2$, *then* $f \in P[z_0+s, \psi, 1]$ *for all* $s > -x_0$.

Proof. First note that $f \in P[z_0+t, \psi, 2]$ for all $t > -x_0$ by Lemma 2.4. Set $z_1 = z_0 + t$ and since f is analytic at z_1 (z_1 lies in the open right half-plane),

$$f^{(1)}(z) = f^{(1)}(z_1) + \sum_{n=2}^{\infty} n r_n e^{i\psi} (z - z_1)^{n-1} \tag{2.10}$$

for all z,

$$|z - z_1| \leqq t + x_0. \tag{2.11}$$

Here $f^{(n)}(z_1) = r_n e^{i\psi}$ and r_n may be negative. Fix z by setting $z = z_0 + s$, s real and chosen so that inequality (2.11) is satisfied. Thus

$$f^{(1)}(z_0+s) = f^{(1)}(z_0+t) + \sum_{n=2}^{\infty} n r_n e^{i\psi} (s-t)^{n-1} \tag{2.12}$$

valid for $|s-t| \leqq t + x_0$, $t > -x_0$. Since f is normalized, we conclude from Theorem 1.3 that $\lim_{t\to\infty} f^{(1)}(z_0+t) = 0$. But then

$$f^{(1)}(z_0+s) = e^{i\psi} \lim_{t\to\infty} \sum_{n=2}^{\infty} n r_n (s-t)^{n-1}$$

follows from (2.12). Thus $f^{(1)}(z_0+s)$ is a number on the line through the origin with slope $\tan\psi$. Finally, s is restricted only by $s > -x_0$ since for any s we may select t sufficiently large so that $|s-t| \leqq t + x_0$, $t > -x_0$.

3. The main theorems. To avoid a repetition of certain notational conventions we agree to write

$$f(z) = f(z_0) + \sum_{n=1}^{\infty} \alpha_n (z - z_0)^n,$$

where $z = x + iy$, $z_0 = x_0 + iy_0$, $x_0 > 0$. Denote by H the half-plane $H = \{w : \psi \leqq \arg w \leqq \psi + \pi\}$, where $\arg w = \psi$ if $w = 0$.

Our first theorem summarizes the content of the lemmas of § 2.

THEOREM 3.1. *If f is a normalized positive function and if for some z_0 there exists an integer $N \geqq 1$ such that $\{\alpha_k\}$, $k \geqq N$, lie in the half-plane H, then the numbers $\{\alpha_k\}$, $k \geqq 1$, all lie on the line through the origin with slope* $\tan \psi$.

Lemma 2.4 can also be written as a theorem of mappings of normalized positive functions.

COROLLARY 3.2. *If f satisfies the hypothesis of Theorem* 3.1, *then f maps the line $z_0 + s$ for $s > -x_0$ into the line*

$$f(z_0) + e^{i\psi}t, \quad t \text{ real.} \tag{3.1}$$

Proof. Since

$$f(z) = f(z_0) + \sum_{n=1}^{\infty} \alpha_n (z - z_0)^n$$

we have

$$f(z_0 + s) = f(z_0) + \sum_{n=1}^{\infty} \alpha_n s^n$$

and by Theorem 3.1, this latter equation may be written

$$f(z_0 + s) = f(z_0) + e^{i\psi}t,$$

where t is a real-valued function of s.

An immediate consequence of Theorem 3.1 is an interesting restriction on the signs of the derivatives of positive real functions at each real, positive z.

THEOREM 3.3. *Suppose f is a nonlinear positive real function and x is a real, positive number. Then, for every integer $N \geqq 1$, there exist integers n and m both larger than N such that $f^{(n)}(x)$ and $f^{(m)}(x)$ have opposite sign.*[1]

Proof. Let $\tilde{f}$ be the normalized function constructed from f so that $\tilde{f}(z) = f(z) - Az$. Since f is neither constant nor linear, $\tilde{f}$ is nonconstant. Also, $\tilde{f}(x)$ is real since A is real and f is positive real. Suppose the conclusion were false. Then since $\tilde{f}^{(k)}(x) = f^{(k)}(x)$ for $k > 1$, there would exist some N sufficiently large such that all the numbers $\{f^{(k)}(x)\}$, $k \geqq N > 1$, would have the same sign. In this event, the imaginary axis would be the boundary of a half-plane containing $\{\alpha_k\}$, $k \geqq N$, and hence from Theorem 3.1, $\alpha_1, \alpha_2, \cdots$, would be on the axis and real. Thus $\alpha_k = f^{(k)}(x)/k! = 0$, $k \geqq 1$. Hence $\tilde{f}(z) = f(z) - Az$ would be constant and f would either be linear or constant, neither of which is possible.

A similar argument generalizes the following idea: If f is positive and if the numbers $f(x_0), f^{(1)}(x_0), f^{(2)}(x_0), \cdots, x_0 > 0$ are all real, then by analytic continuation f is real for all real, positive z. Hence f is positive real. Contrast this with the conclusion of the next theorem.

THEOREM 3.4. *If f is positive and for some $x_0 > 0$ there is an integer $N \geqq 1$ such that $f(x_0), f^{(N)}(x_0), f^{(N+1)}(x_0), \cdots$, are all real, then f is positive real.*

Proof. Define $\tilde{f}(z) = f(z) - Az$ so that $\tilde{f}$ is normalized. Of course, $\alpha_k = f^{(k)}(x_0)/k!$ and we have $\tilde{f} \in P[x_0, 0, N]$. By Lemma 2.5, $\tilde{f} \in P[x_0, 0, 1]$ and thus $\{\tilde{f}^{(k)}(x_0)\}$, $k \geqq 1$, are real. Hence $\tilde{f}$ is positive real and so is f.

An application of Corollary 3.2 follows next.

[1] The derivatives of a positive real function at a positive number are all real.

THEOREM 3.5. *Suppose f is a normalized positive function $f \in P[z_0, \pi/2, N]$. Then f has bounded real part; $u \leqq 2u_0$.*

Proof. By Corollary 3.2, equation (3.1),

$$f(z_0+s) = f(z_0) + it(s), \tag{3.2}$$

where $t(s)$ is real. Thus f maps the horizontal line z_0+s, $s > -x_0$, into the vertical line $f(z_0)+it$, and by the principle of symmetry, the image of the open right half-plane must be symmetric with respect to the line $f(z_0)+it$. But then the real part of f cannot be unbounded. Indeed, $u \leqq 2u_0$ must hold.

The next theorem is proved as a consequence of Theorem 3.1 and concludes this note.

THEOREM 3.6. *Suppose f is positive and the numbers $\{\alpha_k\}$, $k \geqq N \geqq 1$, lie in the wedge, $\{w: \psi+\delta \leqq \arg w \leqq \psi+\pi-\delta, 0<\delta<\pi/2\}$. Then f is linear.*

Proof. Suppose $A \geqq 0$ is defined so that $\tilde{f}(z) = f(z) - Az$ is a normalized positive function. The wedge given in the hypothesis is the intersection of two half-planes whose boundaries are not co-incident. Theorem 3.1 implies, for each of these half-planes, that $\{\alpha_k\}$, $k \geqq 1$, lie on both boundaries simultaneously. The boundaries have only the origin in common. Thus, $0 = \alpha_1 = \alpha_2 = \cdots$, and $\tilde{f}$ is therefore a constant. Thus $f(z) = \gamma + Az$. Since f is not a constant, $A \neq 0$ and f must be linear.

REFERENCES

[1] R. BOTT AND R. J. DUFFIN, *Impedence synthesis without the use of transformers*, J. Appl. Phys., 20 (1949), p. 816.

[2] O. BRUNE, *Synthesis of a finite two-terminal network*, J. Math. and Phys., 10 (1931), pp. 191–235.

[3] J. L. GOLDBERG, *Functions with positive real part in a half-plane*, Duke Math. J., 29 (1962), pp. 333–340.

[4] ———, *Bounds on the derivatives of positive functions*, SIAM Rev., 8 (1966), pp. 343–345.

[5] J. WOLFE AND F. DEKOK, *Les fonctions holomorphes a partie réelle positive et l'integral de Stieltjes*, Bull. Soc. Math. France, 60 (1932), pp. 221–227.

SIAM J. MATH. ANAL.
Vol. 8, No. 5, October 1977

A VARIATIONAL APPROACH TO MULTIPARAMETER EIGENVALUE PROBLEMS FOR MATRICES*

PAUL BINDING AND PATRICK J. BROWNE†

Abstract. The variational theory of eigenvalues and eigenvectors is extended to the multiparameter problem $(T_r+\sum_{s=1}^{k}\lambda_s V_{rs})x_r=0$, $r=1,\cdots,k$, where T_r and V_{rs} are linear operators on finite-dimensional Hilbert spaces H_r. Appropriate variational problems are posed in $\bigoplus_{r=1}^{k}H_r$ and $\bigotimes_{r=1}^{k}H_r$ and give, for example, existence and reality of eigentuples and orthonormality of eigenvectors in an appropriate sense. The fact that the numerical range of an Hermitian matrix is the convex hull of its eigenvalues is directly generalized. An $\mathbb{R}^k$-valued generalized Rayleigh quotient is shown to possess analogues of constrained minimaxima and unconstrained saddle points, when evaluated at eigenvectors. Finally, dependence of T_r and V_{rs} on a parameter is investigated.

1. Introduction. It is the aim of this paper to discuss multiparameter eigenvalue problems of the type considered by Atkinson [1] from a variational point of view. Differential equation problems have been discussed in [9].

Let H_r denote the complex finite-dimensional Hilbert space $\mathbb{C}^{m_r}$ and let T_r and V_{rs} be Hermitian linear operators on H_r, $r,s=1,\cdots,k$. A *multiparameter eigenvalue problem* is formulated by asking for k-tuples $\lambda=(\lambda_1,\cdots,\lambda_k)$ of complex numbers and nonzero vectors $x_r\in H_r$, $r=1,2,\cdots,k$, such that

$$T_r x_r+\sum_{s=1}^{k}\lambda_s V_{rs}x_r=0, \qquad r=1,\cdots,k. \tag{1}$$

λ is then called an *eigenvalue* and $x=x_1\otimes\cdots\otimes x_k\in\bigotimes_{r=1}^{k}H_r$ a corresponding *eigenvector.* It is also possible to regard the eigenvector $\mathbf{x}=(x_1,\cdots,x_k)$ as a point in $\bigoplus_{r=1}^{k}H_r$.

We shall repose the problem (1) in variational terms in several ways and shall draw corresponding conclusions about existence and extremal properties of eigenvalues, geometric properties of eigenvectors and a type of numerical range of the system.

Using the $\bigoplus_{r=1}^{k}H_r$ setting for eigenvectors, one such variational formulation turns out to be the vector maximization of the k-vector of real numbers

$$(\det D_r(\mathbf{x}))_{r=1}^{k} \quad \text{subject to} \quad \det D_0(\mathbf{x})=1, \tag{2}$$

where

$$[D_0(\mathbf{x})]_{rs}=(V_{rs}x_r,x_r)=[D_i(\mathbf{x})]_{rs}, \qquad 0\neq i\neq s,$$

$$[D_i(\mathbf{x})]_{ri}=(T_r x_r,x_r), \qquad i,r,s=1,\cdots,k.$$

In this work, $a\leqq b$ where a,b are real k-vectors shall mean $a_r\leqq b_r$, $r=1,\cdots,k$. A vector $a\in S\subset\mathbb{R}^k$ is said to be a vector maximum of S if it exceeds all points of S with which it is comparable.

* Received by the editors August 21, 1975, and in revised form, April 19, 1976.

† Department of Mathematics and Statistics, University of Calgary, Calgary, Alberta, Canada T2N 1N4. The research of the authors was supported in part by the National Research Council of Canada, Grant No. A 9071 and Grant No. A 9073, respectively.

We shall assume that $D_0(\mathbf{x})$ is a positive definite matrix for any $\mathbf{x} = (x_1, \cdots, x_k) \in \bigoplus_{r=1}^{k} H_r$ with each $x_r \neq 0$. Note that this assumption implies that an eigenvalue $\lambda = (\lambda_1, \cdots, \lambda_k)$ has real entries λ_r: see for example, [1, Thm. 7.2.1, p. 117].

Before proceeding, we review a few well-known facts concerning the problem in the case $k = 1$. We shall omit subscripts for convenience. The problem becomes one of maximizing

$$-(Tx, x) \quad \text{subject to} \quad (Vx, x) = 1, \tag{3}$$

where, by assumption, T and V are Hermitian and V is positive definite. Then (3) can be solved by adjoining the constraint $(Vx, x) = 1$ with a Lagrange multiplier λ giving the first order condition $Tx + \lambda Vx = 0$ which is just (1) in case $k = 1$. Exactly m real possibilities $\lambda^1, \cdots, \lambda^m$ exist for λ corresponding to the various extreme values of $-(Tx, x)$ at eigenvectors $x^1, \cdots, x^m$. These eigenvectors are V-orthogonal in the sense that for $\lambda^i \neq \lambda^j$ we have $(Vx^i, x^j) = 0$. Further it is possible to form a V-orthonormal basis of H consisting of eigenvectors. Also T admits the spectral decomposition

$$Tx + \sum_{i=1}^{m} \lambda^i (Vx, x^i) Vx^i = 0.$$

Finally, if the λ^i are ordered by

$$\lambda^m \leqq \cdots \leqq \lambda^1, \tag{4}$$

then

$$\begin{aligned} \lambda^i &= \max \{-(Tx, x) | (Vx, x) = 1, (Vx, x^j) = 0, 1 \leqq j < i\} \\ &= \min \{\max \{-(Tx, x) | (Vx, x) = 1, (Vx, y^j) = 0\} | y^j \neq 0, 1 \leqq j < i\} \end{aligned}$$

with similar characterizations involving minima and maximinima. In particular, for any $y \neq 0$,

$$\lambda^2 \leqq \max \{-(Tx, x) | (Vx, x) = 1, (Vx, y) = 0\} \leqq \lambda^1,$$

and so on, leading to Rayleigh–Ritz approximation procedures. Suitable references for these facts are [5, Chap. 1] and [7, Chap. 2].

We shall obtain generalizations of these results to the case $k > 1$ in §§ 2 and 3. The key tool, now common in multiparameter theory (see Atkinson [1] or Browne [4]) is a second reformulation of (1) in the tensor product space $H = \bigotimes_{r=1}^{k} H_r$—a complex Hilbert space of finite dimension $M = \prod_{r=1}^{k} m_r$. This leads to a version of (2) in H where the constraint set is compact which it need not be for (2) as posed in $\bigoplus_{r=1}^{k} H_r$. In this multiparameter case, exactly M real eigenvalues exist. Although they cannot be ordered as in (4)—indeed there may be eigenvalues noncomparable in our sense of ordering of k-vectors—analogues of the extremal characterizations, along with Rayleigh–Ritz like formulas, do hold as will be seen.

In § 4 we obtain more quantitative information about the extremal nature of the eigenvalues. Specifically we set

$$\lambda_r(\mathbf{x}) = \det D_r(\mathbf{x}) / \det D_0(\mathbf{x}), \qquad r = 1, \cdots, k. \tag{5}$$

Unconstrained vector maximization of the k-tuple $(\lambda_r(\mathbf{x}))$ is a further reformulation of (2). Evidently this reduces to maximizing

$$\lambda(x) = -(Tx, x)/(Vx, x)$$

in case $k = 1$. Direct computation gives $\lambda_x = 0$, $\lambda_{xx} = -2(T + \lambda V)$, where subscripts denote differentiation at the corresponding eigenvector. In particular λ has a maximum (minimum) at $\lambda^1 (\lambda^m)$ and a saddle point at λ^i for $\lambda^1 > \lambda^i > \lambda^m$. Our results for the multiparameter case show that

$$\frac{\partial \lambda_r}{\partial x_i} = 0, \qquad \frac{\partial^2 \lambda_r}{\partial x_i \, \partial x_j} = 0, \qquad i \neq j,$$

and characterize $\partial^2 \lambda_r / \partial x_j^2$ at eigenvectors.

An example demonstrating the various possibilities for maxima and minima of eigenvalues will be given in § 5. We conclude with a short discussion of the situation in which the various operators depend continuously on a parameter.

2. The tensor product formulation. The maps T_r, V_{rs} induce maps $T_r^\dagger$, $V_{rs}^\dagger$ on H. For a decomposable tensor $x = x_1 \otimes \cdots \otimes x_k \in H$, $T_r^\dagger x = x_1 \otimes \cdots \otimes x_{r-1} \otimes T_r x_r \otimes x_{r+1} \otimes \cdots \otimes x_k$. $T_r^\dagger$ is then extended to all of H by linearity. $V_{rs}^\dagger$ is defined similarly. We now define operators $\Delta_0, \Delta_1, \cdots, \Delta_k$ on H by means of the formal determinantal expansion

$$\sum_{s=0}^{k} \alpha_s \Delta_s = \begin{vmatrix} \alpha_0 & \alpha_1 & \cdots & \alpha_k \\ T_1^\dagger & V_{11}^\dagger \cdots & & V_{1k}^\dagger \\ & \cdots & & \\ & \cdots & & \\ T_k^\dagger & V_{k1}^\dagger \cdots & & V_{kk}^\dagger \end{vmatrix},$$

the $\alpha_0, \cdots, \alpha_k$ being arbitrary complex numbers. Note that operators from different rows of this determinant commute. It is easily seen that the operators $\Delta_0, \cdots, \Delta_k$ are Hermitian on H with respect to the inner product $(\cdot, \cdot)$ induced by the inner products in the spaces H_r.

As is customary in multiparameter theory we shall adopt the definiteness hypothesis: Δ_0 *is positive definite.* Accordingly, we may define a new inner product on H by $[x, y] = (\Delta_0 x, y)$. We shall write $S = \{x \in H | [x, x] = 1\}$ and also $\Gamma_s = \Delta_0^{-1} \Delta_s$, $s = 1, \cdots, k$.

The eigenvalue problem (1) is equivalent to the simultaneous problems in H:

$$\Gamma_s x = \lambda_s x, \qquad s = 1, \cdots, k, \quad x \in H, \tag{6}$$

for decomposable x, see [1, § 6.8, pp. 111, 112].

The variational approach leads us to consider the $\mathbb{R}^k$-valued function Λ defined on nonzero elements of H by

$$(\Lambda x)_r = [\Gamma_r x, x]/[x, x], \qquad r = 1, \cdots, k. \tag{7}$$

DEFINITION. The *vectorial range* of the problem (1) is the set $\Lambda(S) \subseteq \mathbb{R}^k$.

Obviously $\Lambda(S) = \Lambda(H - \{0\})$, and indeed if K is any linear subspace of H, then $\Lambda(S \cap K) = \Lambda(K - \{0\})$. In case $k = 1$, $\Lambda : H \to \mathbb{R}$ is just the Rayleigh quotient whose "numerical" range $\Lambda(S)$ is the interval $[\lambda^m, \lambda^1]$ (see (4)).

We propose maximizing linear functionals over the vectorial range; that is, we shall investigate exposed faces of co $\Lambda(S)$, the real convex hull of $\Lambda(S)$.

LEMMA 1. *Let $\mu \in \mathbb{R}^k$ have unit norm. The problem of maximizing $\mu^T \Lambda x$ subject to $x \in S$ possesses solutions satisfying*

$$\sum_{r=1}^{k} \mu_r \Gamma_r x = \nu x, \tag{8}$$

$$\mu^T \Lambda x = \nu, \tag{9}$$

for some real ν. Further the set of solutions of (8) *for fixed μ and ν forms a subspace $K_{\mu\nu}$ of H, invariant for each Γ_s, $s = 1, \cdots, k$.*

Proof. S is compact since H is finite-dimensional. Λ is continuous on S and thus $\mu^T \Lambda$ is a real-valued continuous function on S. Accordingly, a maximizer $x \in S$ for $\mu^T \Lambda$ exists. The problem may be solved by using a Lagrange multiplier ν for the constraint $[x, x] = 1$. Equation (8) is then the first order maximization condition. Finally we note that, since $[x, x] = 1$,

$$\nu = [\nu x, x] = \left[\sum_{r=1}^{k} \mu_r \Gamma_r x, x\right] = \sum_{r=1}^{k} \mu_r (\Lambda x)_r = \mu^T \Lambda x.$$

Suppose x satisfies (8). We apply Γ_s to both sides of this equation and use $y_s = \Gamma_s x$ together with the commutativity of the operators $\Gamma_1, \cdots, \Gamma_k$ [1, Thm. 6.7.2, p. 110] to obtain

$$\sum_{r=1}^{k} \mu_r \Gamma_r y_s = \nu y_s.$$

This shows that $y_s \in K_{\mu\nu}$ and so $\Gamma_s K_{\mu\nu} \subseteq K_{\mu\nu}$.

We are now ready to establish the existence of eigenvalues and eigenvectors. If $\dim K_{\mu\nu} = 1$, then the corresponding vector x will be a solution of (6). In general, we shall simultaneously maximize enough functionals $\mu^T \Lambda$ over S so as to ensure a 1-dimensional space of solutions.

THEOREM 1. (i) *The problem* (6) *possesses at least one solution whose eigenvector e maximizes appropriate functionals $\mu^T \Lambda x$ over $x \in S$, Λe being the corresponding eigenvalue.*

(ii) *e may be taken decomposable in H and thus solves* (1).

Proof. From Lemma 1 we have $\dim K_{\mu\nu} \geqq 1$. Should equality hold for some μ and corresponding ν, then the proof is complete. If not, we select μ (and ν) so that $\dim K_{\mu\nu}$ is minimal. We shall write $K = K_{\mu\nu}$ and for any other μ' (and corresponding ν'), $K' = K_{\mu'\nu'}$.

First suppose that $K = H$, so $K' = H$ for all μ'. Then choosing μ' as the usual coordinate vectors in $\mathbb{R}^k$ in turn we reach (6) for all $x \in H$, $\lambda_1, \cdots, \lambda_k$ being the corresponding values of ν' for these choices of μ'. Hence any $x \in H$ is a solution of (6) with $\lambda_s = (\Lambda x)_s$, $s = 1, \cdots, k$.

We are left to consider $1 < \dim K < M$ and first isolate two cases.

Case 1. $K' \supseteq K$ for each μ' and corresponding ν'.

Again taking μ' as coordinate vectors in $\mathbb{R}^k$ we obtain (6) for all $x \in K$ and so any $x \in K$ is an eigenvector.

Case 2. $K' \cap K = \{0\}$ for each μ' and corresponding ν'.

Let $x' \in K'$ and write $x' = y + z$, where $y \in K$, $z \in K^{\perp}$ (the $[\cdot, \cdot]$ orthocomplement of K). Now we have by (8) with μ' and ν',

$$\sum_{r=1}^{k} \mu_r' \Gamma_r y + \sum_{r=1}^{k} \mu_r' \Gamma_r z = \nu' y + \nu' z.$$

Since K is an invariant subspace for each of the $[\cdot, \cdot]$-Hermitian operators Γ_r, so is $K^{\perp}$, and thus we see that

$$\sum_{r=1}^{k} \mu_r' \Gamma_r y = \nu' y, \qquad \sum_{r=1}^{k} \mu_r' \Gamma_r z = \nu' z.$$

In particular, this gives $y \in K'$ and so $y = 0$. Accordingly,

$$x' = z \in K^{\perp}. \tag{10}$$

We shall now take sequences μ^i, ν^i, x^i converging to μ, ν and some $x \in K$. Since the ν^i depend on μ^i we must digress slightly to ensure that μ^i can be chosen so that the ν^i do approach ν. Let F denote the face of co $\Lambda(S)$ normal to μ, so that

$$F = \text{co}\,\{\Lambda(x) | x \text{ maximizes } \mu^T \Lambda(x) \text{ subject to } x \in S\}.$$

Pick $x \in K$ so that $\lambda^e = \Lambda(x)$ is an exposed point of F; that is, so that for some $\xi \in \mathbb{R}^k$, the inequality

$$\xi^T(\lambda - \lambda^e) \geqq 0, \qquad \lambda \in F,$$

has only one solution, viz. $\lambda = \lambda^e$. Set

$$\mu^i = (\mu + i^{-1}\xi)/\|\mu + i^{-1}\xi\|$$

and let $x^i \in S$ be any maximizer of $\mu^{iT}\Lambda$.

Since S is compact we may assume that $x^i \to x^* \in S$ (passing to a subsequence if necessary). Let

$$\lambda^i = \Lambda(x^i) \to \lambda^* \in \text{co}\,\Lambda(S).$$

Now

$$\mu^{iT}\lambda^i \geqq \mu^{iT}\lambda \quad \forall \lambda \in \text{co}\,\Lambda(S)$$

because x^i is a maximizer of $\mu^{iT}\Lambda$, and hence

$$\mu^T\lambda^* \geqq \mu^T\lambda \quad \forall \lambda \in \text{co}\,\Lambda(S),$$

whence $\lambda^* \in F$ by definition. Thus addition of the inequalities

$$-\mu^T(\lambda^i - \lambda^e) \geqq 0, \qquad (\mu + i^{-1}\xi)^T(\lambda^i - \lambda^e) \geqq 0$$

yields

$$\xi^T(\lambda^i - \lambda^e) \geqq 0.$$

Letting $i \to \infty$ in this final equation we obtain $\lambda^* = \lambda^e$. The limiting versions of (8)

and (9) now show that $x^* \in K$ and $\nu^i \to \nu$. On the other hand, $x^* \in K^\perp \cap S$ follows from (10).

This contradiction shows that Case 2 cannot occur. We complete the proof of part (i) of the Theorem by applying the same reasoning to successively smaller invariant subspaces. Select μ' and ν' so that $\dim L$ is minimal where $L = K \cap K'$. Having eliminated Cases 1 and 2, we are left with

$$1 \leqq \dim L < \dim K.$$

We now use the above argument with L in place of K since K' and hence L is invariant under each Γ_r. If $\dim L = 1$, or more generally if $K'' \supset L$ for all μ'' and ν'', then (6) holds for all $x \in L$. If not, we take a third functional $(\mu'')^T \Lambda$ giving minimal $\dim (L \cap K'')$. Continuing in this way with at most k such functionals required we eventually produce an eigenspace.

(ii) So far then we have that a maximizing vector x satisfies $\Gamma_s x = \lambda_s x$. It now follows that

$$T_r^\dagger x + \sum_{s=1}^{k} \lambda_s V_{rs}^\dagger x = 0, \qquad r = 1, \cdots, k,$$

(see Atkinson [1, Thm. 6.8.1, p. 111]). That x may be taken decomposable and is thus an eigenvector follows from

$$0 \neq x \in \bigcap_{r=1}^{k} \ker\left(T_r^\dagger + \sum_{s=1}^{k} \lambda_s V_{rs}^\dagger\right) = \bigotimes_{r=1}^{k} \ker\left(T_r + \sum_{s=1}^{k} \lambda_s V_{rs}\right)$$

(see Atkinson [1, Thm. 4.7.2, p. 72]). This completes the proof.

As was pointed out earlier, (2) need not give a compact constraint set when posed in $\bigoplus_{r=1}^{k} H_r$: consider $k = 2$, $m_1 = m_2 = 1$, so that

$$\det \mathcal{D}_0(\mathbf{x}) = (V_{11} V_{22} - V_{12} V_{21})(x_1 x_2)^2 = 1,$$

the V_{ij} being real numbers in this case. This constraint gives two complex hyperbolae in the x_1–x_2-plane—not a compact constraint set.

THEOREM 2. *There are $M = \prod_{r=1}^{k} m_r$ possible eigenvalues (counted according to multiplicity) all real k-tuples. Eigenvectors corresponding to different eigenvalues are $[\cdot, \cdot]$-orthogonal. A $[\cdot, \cdot]$-orthonormal basis of H can be constructed from eigenvectors.*

Proof. The reality of eigenvalues is an easy consequence of our definiteness hypothesis and the Hermitian nature of the matrices involved (for a similar argument, see Browne [3, Thm. 1, p. 553]). If λ, μ are different eigenvalues and $x = x_1 \otimes \cdots \otimes x_k$, $y = y_1 \otimes \cdots \otimes y_k$ corresponding eigenvectors, then

$$T_r x_r + \sum_{s=1}^{k} \lambda_s V_{rs} x_r = 0,$$

$$T_r y_r + \sum_{s=1}^{k} \mu_s V_{rs} y_r = 0,$$

so that

$$\sum_{s=1}^{k} (\lambda_s - \mu_s)(V_{rs}x_r, y_r) = 0, \qquad r = 1, \cdots, k.$$

Hence $\det (V_{rs}x_r, y_r) = 0$; that is, $[x, y] = 0$.

The construction of an $[\cdot, \cdot]$-orthonormal basis can be carried out by the usual procedure of first selecting a decomposable x^1 solving the problem (6) with corresponding eigenvalue λ^1 and then repeating the maximization arguments of Theorem 1 subject to the additional constraint $[x, x^1] = 0$. Continuing in this way we produce M eigenvalues $\lambda^1, \lambda^2, \cdots, \lambda^M$ (counted according to multiplicity) and a system of $[\cdot, \cdot]$-orthogonal corresponding eigenvectors $x^1, \cdots, x^M$.

We now have the following eigenvector expansion theorem and Parseval equality.

COROLLARY 1. *Let $x^1, \cdots, x^M$ be $[\cdot, \cdot]$-orthonormal eigenvectors for the multiparameter problem* (1). *Then for any $x \in H$,*

(i) $$x = \sum_{i=1}^{M} [x, x^i]x^i,$$

(ii) $$[x, x] = \sum_{i=1}^{M} |[x, x^i]|^2.$$

In terms of projections we may state

COROLLARY 2. *Define operators P^i, Q^i, $i = 1, \cdots, M$, by $P^i x = [x, x^i]x^i$, $Q^i x = (\Delta_0 x, x^i)\Delta_0 x^i$, $i = 1, \cdots, M$. Then*

(i) $$I = \sum_{i=1}^{M} P^i,$$

(ii) $$\Delta_0 = \sum_{i=1}^{M} Q^i,$$

(iii) $$\Gamma_r = \sum_{i=1}^{M} \lambda_r^i P^i,$$

(iv) $$\Delta_r = \sum_{i=1}^{M} \lambda_r^i Q^i.$$

Thus we have produced an alternative approach to the theory of Atkinson [1, § 7.9, pp. 133–134].

3. The vectorial range. The vectorial range of the problem (1) is defined as $\Lambda(S)$ (see (7)), and in case $k = 1$, it is the convex hull of the eigenvalues. The proof of Theorem 1 shows that the eigenvalues possess certain extremal properties relative both to the set $\Lambda(S)$ and to certain real-valued functionals $\mu^T \Lambda$ defined on S. In this section we demonstrate that $\Lambda(S)$ is a convex polyhedron whose vertices are eigenvalues λ^i, so the corresponding functionals are those defined by the normals to $\Lambda(S)$ at λ^i.

Let I denote the index set $\{1, 2, \cdots, M\}$, J any subset thereof and $\sim J$ its complement in I. Further we set

$$\sigma_J = \{\lambda^j | j \in J\},$$

the $\lambda^j \in \mathbb{R}^k$ being obtained from Theorem 2. In particular, σ_I is called the *spectrum* of the problem (1). For a subset $G \subset H$, sp G will denote its linear span. Finally we shall use

$$S_J = \{x \in S | [x, x^j] = 0, j \in J\} = S \cap \text{sp}\,\{x^i | i \in \sim J\}.$$

Here, $x^1, \cdots, x^M$ represent a $[\cdot, \cdot]$-orthonormal basis of eigenvectors corresponding to $\lambda^1, \cdots, \lambda^M$ as per Theorem 2.

LEMMA 2. $\Lambda(S_J) = \text{co}\,(\sigma_{\sim J})$.

In particular,

$$\Lambda(S) = \text{co}\,(\sigma_I).$$

Proof. Let $x \in S$ be written $x = \sum_{i=1}^M \alpha^i x^i$. Then $\sum_{i=1}^M |\alpha^i|^2 = 1$ and $x \in S_J$ if and only if $\alpha^j = 0$ for $j \in J$. Thus for $x \in S_J$ we have

$$\Lambda(x) = \sum_{j=1}^M |\alpha^j|^2 \lambda^j = \sum_{\sim J} |\alpha^j|^2 \lambda^j \in \text{co}\,(\sigma_{\sim J}).$$

Conversely, if $\sum_{\sim J} \beta^j \lambda^j$ is a point in co $(\sigma_{\sim J})$ we take $x = \sum_{\sim J} \sqrt{\beta^j}\, x^j$. The above analysis shows that $x \in S_J$ and $\Lambda(x) = \sum_{\sim J} \beta^j \lambda^j$. This completes the proof.

The principal result of this section is a generalization of the minimax theorem which in case $k = 1$ admits the interpretation that a set of constraints of the form $[x, y^j] = 0$, $1 \leqq j \leqq p < m$, applied to the Rayleigh quotient cannot force its numerical range entirely to the left of λ^{p+1} (or to the right of λ^{m-p-1} corresponding to the maximin theorem). Put another way, the said range must intersect the interval $[\lambda^{p+1}, \lambda^1]$. We now show in the k-parameter case that for $y^j \in H$, $j = 1, \cdots, p < M$, and J a p-element subset of I, the constraints

$$[x, y^j] = 0, \qquad j = 1, \cdots, p, \tag{11}$$

cannot force the range of Λ "farther" from $\sigma_{\sim J}$ than do the constraints corresponding to choosing $y^j = x^j$, $j = 1, \cdots, p$.

We denote by G the set of those $x \in H$ satisfying (11) and put $Q = \Lambda(G \cap S)$.

THEOREM 3. *Let J be any subset of I containing p elements and let $z \in S_J$. Then Q intersects* co $\{\Lambda(z), \sigma_J\}$. *In particular if $J \subset K \subseteq I$ and K contains $p+1$ elements, then Q intersects* co σ_K.

Proof. G has dimension at least $M - p$ and so intersects

$$T = \text{sp}\,\{z, x^j | j \in J\}$$

in at least a line. Let such a line intersect S at t. Then arguing as in Lemma 2 we have $\Lambda(t) \in \text{co}\,\{\Lambda(z), \sigma_J\}$. However $\Lambda(t) \in \Lambda(G \cap T \cap S) \subseteq \Lambda(G \cap S) = Q$.

If the constraints (11) take a specially simple form in terms of the coefficients of x (the α^j of Lemma 2), the range Q may be convex; however, an example of nonconvex range will be given later.

Returning to the vector ordering in $\mathbb{R}^k$ we can restate the above results to give analogues of extremal characterizations of eigenvalues. Let $\lambda^{1,i}$, $i = 1, \cdots, n(1)$,

denote the noncomparable vector maxima of the spectrum listed in their lexicographic order. We term this collection the *first cycle of eigenvalues*. The *j-th cycle of eigenvalues*, written $\lambda^{j,i}$, $i = 1, \cdots, n(j)$, will be the set of noncomparable eigenvalues λ with the property that if μ represents another eigenvalue, then $\lambda \leqq \mu$ holds for at least one μ from the preceding cycle and $\lambda \leqq \mu$ cannot hold unless μ comes from an earlier cycle. That is, the jth cycle consists of the set of vector maxima of $\sigma_I - \bigcup_{i=1}^{j-1}$ (cycle i). Each cycle will be listed according to the lexicographic order of its members. This provides a listing $\lambda^{j,i}$, $i = 1, \cdots, n(j)$, $j = 1, 2, \cdots$, of all the eigenvalues. We shall denote the corresponding eigenvectors by $x^{j,i}$. As an immediate consequence of Lemma 2 we claim

COROLLARY 3. *Let $j \geqq 1$ and consider the vectorial range Λ subject also to the constraints $[x, x^{p,q}] = 0$, $q = 1, \cdots, n(p)$, $p = 1, \cdots, j-1$. Then the j-th cycle is a subset of the set V of vector maxima of Λ. Further, the vertices of V form a subset of the j-th cycle.*

If extra constraints of the form $[x, x^{j,r}] = 0$ for say $r = 1, \cdots, s < n(j)$ are also imposed, then the corresponding set Λ will have the remainder of the jth cycle as vector maxima but there may also be further vertices of the form $\lambda^{u,v}$ for $u > j$ among the vector maxima. We close this section with a direct analogue of the minimax theorem.

COROLLARY 4. *Let $p = \sum_{q=1}^{j-1} n(q)$. If the set Λ is subject to p extra constraints of the form* (11) *with Q as the corresponding vectorial range, then for each point $\lambda^{j,i}$ in the j-th cycle there is a vector maximum μ of Q with $\lambda^{j,i} \leqq \mu$.*

4. Differentiability of eigenvalues. It will now be convenient to return to the space $\bigoplus_{r=1}^{k} H_r$ and to recall (5) where we defined

$$\lambda(\mathbf{x}) = (\lambda_r(\mathbf{x}))_{r=1}^{k} = (\det D_r(\mathbf{x}) / \det D_0(\mathbf{x}))_{r=1}^{k}. \tag{5}$$

Evidently $\lambda(\mathbf{x}) = \Lambda(x_1 \otimes \cdots \otimes x_k)$ as defined in (7). We shall regard λ as a function of the k vector variables $x_1, \cdots, x_k$. λ_{jr}, λ_{jrs} will denote derivatives $\partial\lambda_j/\partial x_r$, $\partial^2\lambda_j/\partial x_r\, \partial x_s$ etc. Notice that $(\partial\lambda_j/\partial x_r)(\mathbf{x})$ is a linear map defined on H_r and taking values in $\mathbb{C}$ and as such $(\partial\lambda_j/\partial x_r)(\mathbf{x})$ can be regarded as a point in H_r. In similar fashion, λ_{jrr} can be regarded as a linear map defined on H_r and taking values in H_r.

THEOREM 4. *At any eigenvector $\mathbf{x}^i$, the function λ defined by* (5) *is infinitely differentiable. Further all first order, and all higher order mixed derivatives vanish and*

$$2\left(T_r + \sum_{j=1}^{k} \lambda_j(\mathbf{x}^i) V_{rj}\right) \delta_{rs} + \sum_{j=1}^{k} [D_0(\mathbf{x}^i)]_{rj} \lambda_{jss}(\mathbf{x}^i) = 0.$$

Proof. The differentiability of λ at an eigenvalue is clear from its definition. We write

$$U_r(\mathbf{x}) = T_r + \sum_{j=1}^{k} \lambda_j(\mathbf{x}) V_{rj} : H_r \to H_r, \qquad r = 1, \cdots, k,$$

$$v_{rs}(\mathbf{x}) = [D_0(\mathbf{x})]_{rs}, \qquad r, s = 1, \cdots, k.$$

Then (5) may be written

$$(x_r, U_r(\mathbf{x})x_r) = 0,$$

and differentiating both sides of this equation with respect to x_s we obtain

$$2U_r(\mathbf{x})x_r\delta_{rs} + \sum_{j=1}^{k} v_{rj}(\mathbf{x})\lambda_{js}(\mathbf{x}) = 0. \tag{12}$$

At $\mathbf{x} = \mathbf{x}^i$ we have $U_r(\mathbf{x}^i)x_r^i = 0$ and so

$$\sum_{j=1}^{k} v_{rj}(\mathbf{x}^i)\lambda_{js}(\mathbf{x}^i) = 0.$$

The nonsingularity of $D_0(\mathbf{x}^i)$ now yields

$$\lambda_{js}(\mathbf{x}^i) = 0.$$

Next we differentiate both sides of (12) with respect to x_t, $t \neq s$, to obtain at $\mathbf{x} = \mathbf{x}^i$,

$$\sum_{j=1}^{k} v_{rj}(\mathbf{x}^i)\lambda_{jst}(\mathbf{x}^i) = 0.$$

Thus at an eigenvalue the second and similarly all higher order mixed partial derivatives must vanish. Finally we differentiate both sides of (12) with respect to x_s to obtain at $\mathbf{x} = \mathbf{x}^i$,

$$2U_r(\mathbf{x}^i)\delta_{rs} + \sum_{j=1}^{k} v_{rj}(\mathbf{x}^i)\lambda_{jss}(\mathbf{x}^i) = 0.$$

This holds for $r, s = 1, \cdots, k$ and $i = 1, \cdots, M$. Higher order derivatives may be treated similarly. In particular, the invertibility of $D_0(\mathbf{x}^i)$ enables us to solve the last equation for $\lambda_{jss}(\mathbf{x}^i)$.

5. An example. We take $k = 2$, $m_1 = m_2 = 2$ and consider the 2-parameter system of equations for $x, y \in \mathbb{R}^2$, $\lambda, \mu \in \mathbb{R}$:

$$Ux = \left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \lambda\begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix} + \mu\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}\right]x = 0,$$

$$Vy = \left[\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} + \lambda\begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} + \mu\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\right]y = 0.$$

Here we shall use (λ, μ) for an eigenvalue. Calculation of the eigenvalues and eigenvectors is easily carried out via the characteristic equations

$$0 = \det\begin{pmatrix} -\lambda + 2\mu & 1 \\ 1 & -2\lambda - \mu \end{pmatrix} = \det\begin{pmatrix} -\mu & 0 \\ 0 & -\lambda - \mu \end{pmatrix}.$$

This gives the following system of 4 eigenvalues and eigenvectors:

$$(\lambda, \mu) = (\pm 1/\sqrt{2}, 0), \qquad x = \begin{pmatrix} \pm\sqrt{2} \\ 1 \end{pmatrix}, \qquad y = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$(\lambda, \mu) = (\pm 1/\sqrt{3}, \mp 1/\sqrt{3}), \qquad x = \begin{pmatrix} 1 \\ \pm\sqrt{3} \end{pmatrix}, \qquad y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The eigenvalues are displayed as A, B, C and D on the accompanying manutract. According to the classification of Corollary 3, A and B form the first cycle, C and D the second. Observe that no single eigenvalue is a vector maximum, also that $C \not\equiv B$. Theorem 1 indicates that the eigenvector $e_A = \begin{pmatrix} \sqrt{2} \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ corresponding to A maximizes the form

$$\alpha x_1 x_2 (y_1^2 + y_2^2) - \beta x_1 x_2 y_2^2$$

subject to

$$(x_1^2 + 2x_2^2)(y_1^2 + y_2^2) + (2x_1^2 - x_2^2) y_2^2 = 1$$

for any pair of vectors $\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ normal to the parallelogram $ABCD$ at A. This parallelogram is the vectorial range of our system; that is, it is the range of the function Λ (see (7)). If orthogonality to e_A (in the $[\,\cdot\,,\,\cdot\,]$ sense) is imposed on the function Λ, the resulting range becomes the triangle BCD, while if $[\,\cdot\,,\,\cdot\,]$-orthogonality to the eigenvector for B is also imposed, the range becomes the line segment CD, and so on.

According to Theorem 3, imposing $[\,\cdot\,,\,\cdot\,]$-orthogonality to any point in H will force the range to intersect every line joining pairs of the vertices A, B, C and D. In particular, this range cannot be interior to any of the triangles ABC, ABD, BCD and ACD.

We now give an example promised earlier of a nonconvex range arising from suitable orthogonality conditions. Let e_A, e_B, e_C and e_D be the eigenvectors corresponding to A, B, C and D. We consider the function Λ subject to conditions of orthogonality to the vectors $e_A + e_B + e_C - e_D$ and $e_A - 2e_D$. The elements (λ, μ) of this range, being in the parallelogram $ABCD$, may be expressed as

$$(\lambda, \mu) = \alpha^2 A + \beta^2 B + \gamma^2 C + \delta^2 D,$$

where α, β, γ, δ are real and satisfy

$$\alpha + \beta + \gamma = \delta, \quad \alpha = 2\delta \quad \text{and} \quad \alpha^2 + \beta^2 + \gamma^2 + \delta^2 = 1.$$

Eliminating α and δ we obtain an ellipse in (β, γ) coordinates and hence a quadratic equation in β^2 and γ^2. Since $\alpha^2 = 4\delta^2 = \frac{4}{5}(1 - \beta^2 - \gamma^2)$, it follows that the range is a curve. This curve is in fact closed, but we merely show its nonconvexity

by exhibiting the three points

$$P=(-(2\sqrt{3}+3\sqrt{2})/12, \sqrt{3}/6), \quad \text{corresponding to } \alpha=0;$$
$$Q=((9\sqrt{2}+2\sqrt{3})/36, -\sqrt{3}/18), \quad \text{corresponding to } \beta=0;$$
$$R=(\sqrt{2}/3, 0), \quad \text{corresponding to } \gamma=0.$$

It is easy to check that these points are not collinear.

Finally, we classify the extremal natures of A, B, C and D according to Theorem 4. For the pair A and C $(\pm 1/\sqrt{2}, 0)$ we have

$$U=\begin{pmatrix} \mp 1/\sqrt{2} & 1 \\ 1 & \mp\sqrt{2} \end{pmatrix}, \quad V=\begin{pmatrix} 0 & 0 \\ 0 & \mp 1/\sqrt{2} \end{pmatrix}, \quad D_0=\begin{pmatrix} -4 & 3 \\ 0 & -1 \end{pmatrix}.$$

Theorem 4 now gives

$$2\begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix}+\begin{pmatrix} -4 & 3 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \lambda_{xx} & \lambda_{yy} \\ \mu_{xx} & \mu_{yy} \end{pmatrix}=0,$$

where $\lambda_{xx}=\partial^2\lambda/\partial x^2$ at the appropriate eigenvector. Solving we obtain

$$\begin{pmatrix} \lambda_{xx} & \lambda_{yy} \\ \mu_{xx} & \mu_{yy} \end{pmatrix}=\frac{1}{2}\begin{pmatrix} U & 3V \\ 0 & 4V \end{pmatrix}.$$

For the case $A(1/\sqrt{2}, 0)$, U is negative semidefinite as is V. Hence we see that λ has a local maximum jointly in x and y while μ has a point of inflection in x and a maximum in y.

For the case $C(-1/\sqrt{2}, 0)$, both U and V are positive semidefinite so that the above conclusions hold with maximum replaced by minimum.

For the pair B and D $(\mp 1/\sqrt{3}, \pm 1/\sqrt{3})$ we have

$$U=\begin{pmatrix} \pm\sqrt{3} & 1 \\ 1 & \pm 1/\sqrt{3} \end{pmatrix}, \quad V=\begin{pmatrix} \mp 1/\sqrt{3} & 0 \\ 0 & 0 \end{pmatrix}, \quad D_0=\begin{pmatrix} -7 & -1 \\ -1 & -1 \end{pmatrix}.$$

Arguing as before we obtain

$$\begin{pmatrix} \lambda_{xx} & \lambda_{yy} \\ \mu_{xx} & \mu_{yy} \end{pmatrix}=\frac{1}{3}\begin{pmatrix} U & -V \\ -U & 7V \end{pmatrix}.$$

For the case $B(-1/\sqrt{3}, 1/\sqrt{3})$, U is positive semidefinite while V is negative semidefinite. Hence λ has a joint minimum while μ has a joint maximum.

For the case $D(1/\sqrt{3}, -1/\sqrt{3})$, U is negative semidefinite while V is positive semidefinite. Hence the conclusions for B hold here with maximum and minimum interchanged.

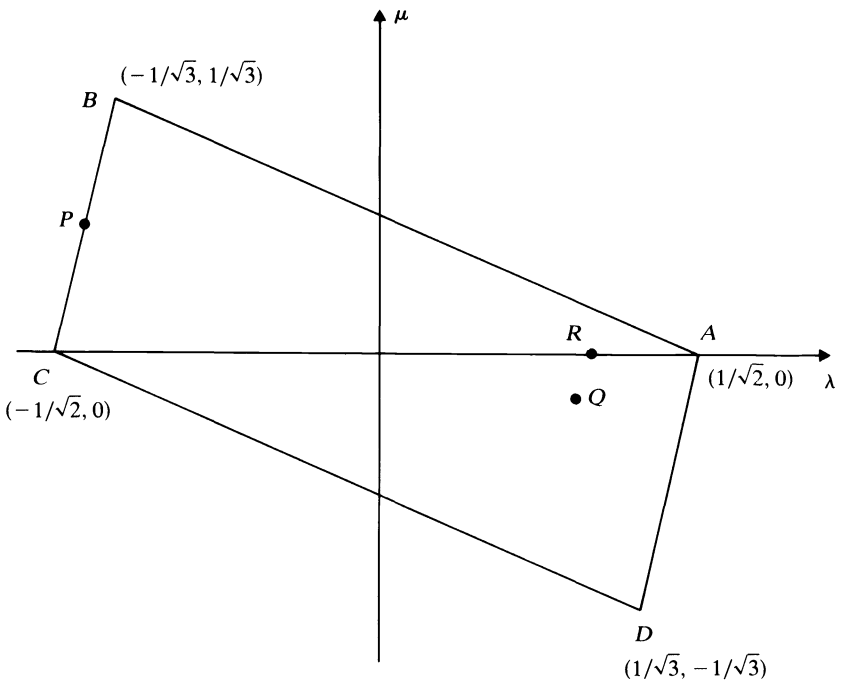


FIG. 1

6. Conclusion. In ordinary spectral theory, the variational approach extends to both infinite-dimensional operators and to those depending on a parameter. We hope to extend our theory to infinite-dimensional operators in a separate work. To conclude the present paper we present a brief discussion of the situation in which our operators depend continuously on the parameter t which for simplicity we take as ranging through a compact real interval containing zero.

THEOREM 5. *Let the operators T_r, V_{rs} depend continuously on t. Then eigenvalues depend continuously and eigenspaces upper semicontinuously on t.*

Proof. For notational ease, we treat only the case $t \to 0$. In the finite-dimensional case here, upper semicontinuity of eigenspaces means that if $x_t \in H$ is an eigenvector of the system corresponding to t and if $x_t \to x \in H$ as $t \to 0$, then x is an eigenvector for the system at $t = 0$.

Fix $\mu \in \mathbb{R}^k$, $\mu \neq 0$ and let x_t maximize $\mu^T \Lambda_t(x)$ over the $[\,\cdot\,,\,\cdot\,]_t$ unit ball S_t. Set

$$x_{0t} = x_t / [x_t, x_t]_0^{1/2}$$

to give $x_{0t} \in S_0$ and let $x^* \in S_0$ be any accumulation point of x_{0t} as $t \to 0$ so that without loss we take $x_{0t} \to x^*$. From this it readily follows that $x_t \to x^*$ by virtue of the continuous dependence of Δ_0 on t. Thus, as $t \to 0$,

$$\mu^T \Lambda_t(x_t) \to \mu^T \Lambda_0(x^*) \leqq \mu^T \Lambda_0(x_0).$$

Conversely, with

$$x_{t0} = x_0 / [x_0, x_0]_t^{1/2},$$

we obtain $x_{t0} \in S_t$ and $x_{t0} \to x_0$. Hence we have

$$\mu^T \Lambda_t(x_t) \geqq \mu^T \Lambda_t(x_{t0}) \to \mu^T \Lambda_0(x_0) \quad \text{as } t \to 0.$$

So far then, the vectorial range $\Lambda_t(S_t)$ is continuous in the sense of affine (support) functionals. This guarantees that each exposed point of $\Lambda_0(S_0)$ is the limit of exposed points of $\Lambda_t(S_t)$ as $t \to 0$, and also ensures that we may find a compact set $K \subset \mathbb{R}^k$ containing each numerical range $\Lambda_t(S_t)$.

The upper semicontinuity of eigenspaces is easily established, for suppose we have eigenvectors x_t such that $x_t \to x$ as $t \to 0$ through a sequence of values. By taking a subsequence if necessary we may assume that the corresponding eigenvalues λ_t converge to say $\lambda \in \mathbb{R}^k$. Now we have

$$\Gamma_{ts} x_t = \lambda_{ts} x_t, \qquad s = 1, \cdots, k,$$

and letting $t \to 0$ through this sequence we obtain

$$\Gamma_s x = \lambda_s x, \qquad s = 1, \cdots, k.$$

If we take $[x_t, x_t]_t = 1$ it follows that $[x, x]_0 = 1$ and so we see that x is an eigenvector.

A similar argument shows that any accumulation point of eigenvalues λ_t must be an eigenvalue for the $t = 0$ problem. Finally to show the dependence of the spectrum on t, let λ_0^i, $i = 1, \cdots, M$, be an enumeration of the eigenvalues for $t = 0$. Suppose there exist $\varepsilon > 0$ and a sequence of t values with $|\lambda_t - \lambda_0^i| > \varepsilon$, $i = 1, \cdots, M$, for at least one λ_t. These eigenvalues λ_t must have a point of accumulation λ^* which by the above is an eigenvalue for $t = 0$. However, we have $|\lambda^* - \lambda_0^i| \geqq \varepsilon$, $i = 1, \cdots, M$. This contradiction shows that $\forall \varepsilon > 0 \; \exists \, \delta(\varepsilon)$ such that

$$|t| < \delta \Rightarrow \sigma_t \subseteq \bigcup_{i=1}^{M} B(\lambda_0^i, \varepsilon).$$

Here, σ_t represents the spectrum at t and $B(\lambda_0^i, \varepsilon)$ represents the k-dimensional ball with center λ_0^i and radius ε. This is the desired eigenvalue continuity.

COROLLARY 5. *Let ν be a regular Borel measure defined on $\mathbb{R}$. If T_r, V_{rs} are ν-measurable functions of t, then a ν-measurable $[\cdot, \cdot]_t$-orthonormal basis of eigenvectors x_t^i, $i = 1, \cdots, M$, exists. Eigenvalues λ_t^i are also ν-measurable.*

Proof. The result is similar to Kaz's theorem [6, pp. 1341–1345]. We sketch a proof based on the ideas of [2, Thm. 3]. We require Vitali's theorem [6, Lemma 17, p. 1218] which states that measurable functions are continuous except on sets of arbitrarily small measure. Let us work then with a compact set on which T_r and V_{rs} are all continuous. We have seen that the set of eigenvectors is an upper semicontinuous function of t and the same is clearly true for the set of all $[\cdot, \cdot]_t$-orthonormal bases of eigenvectors. Using the selection theorem of [8, p. 398], we obtain a $[\cdot, \cdot]_t$-orthonormal basis $x_t^1, \cdots, x_t^M$, each x_t^i being a ν-measurable function of t. Now the eigenvalues are given by $\lambda_t^i = \Lambda_t(x_t^i)$, $i = 1, \cdots, M$, and Λ being continuous in t we obtain ν-measurable eigenvalues λ_t^i, $1 \leqq i \leqq M$. Vitali's theorem now shows that these considerations are sufficient to justify our claim.

Acknowledgment. This study was motivated by informal discussions with F. M. Arscott at the recent summer meeting of the Canadian Mathematical Congress.

REFERENCES

[1] F. V. ATKINSON, *Multiparameter Eigenvalue Problems*, vol. I, *Matrices and Compact Operators*, Academic Press, New York, 1972.

[2] P. BINDING AND P. J. BROWNE, *L^p Spaces generated by certain operator valued measures*, Canad. Math. Bull., to appear.

[3] P. J. BROWNE, *A multiparameter eigenvalue problem*, J. Math. Anal Appl., 38 (1972), pp. 553–568.

[4] ———, *Multi-parameter spectral theory*, Indiana Univ. Math. J., 24 (1974), pp. 249–257.

[5] R. COURANT AND D. HILBERT, *Methods of Mathematical Physics*, vol. I, Wiley–Interscience, New York, 1953.

[6] N. DUNFORD AND J. T. SCHWARTZ, *Linear Operators, Part II: Spectral Theory*, Wiley–Interscience, New York, 1963.

[7] S. H. GOULD, *Variational Methods for Eigenvalue Problems*, Oxford University Press, Oxford, 1966.

[8] K. KURATOWSKI AND C. RYLL-NARDZEWSKI, *A general theorem on selectors*, Bull. Acad. Polon. Sci. Ser. Mat. Astronom. Phys., 13 (1965), pp. 397–403.

[9] B. D. SLEEMAN, *Completeness and expansion theorems for a two-parameter eigenvalue problem in ordinary differential equations using variational principles*, J. London Math. Soc. (2), 6 (1973), pp. 705–712.

SIAM J. MATH. ANAL.
Vol. 8, No. 5, October 1977

REGULAR SINGULAR DIFFERENTIAL EQUATIONS WHOSE CONJUGATE EQUATION HAS POLYNOMIAL SOLUTIONS*

L. M. HALL†

Abstract. Consider the n-dimensional singular differential system defined by the operator $L: (Ly)(z) = z^p y'(z) + A(z)y(z)$, where z is a complex variable and p is a positive integer. The solvability of the nonhomogeneous system $Ly = g$ depends on the solutions of the homogeneous conjugate system, $L^*f = 0$, where L^* is the operator conjugate to L. We show that $L^*f = 0$ has polynomial solutions if the constant matrix in the series expansion of $A(z)$ has at least one nonpositive integer eigenvalue. Also, we show that if $L^*f = 0$ has a polynomial solution, then a finite number of the coefficients of $A(z)$ must satisfy certain properties. These results are then used to obtain a solvability condition for the nonhomogeneous Bessel equation of integer order.

1. Introduction. Let $A_{q,n}$ be the space of n-vector functions whose components are analytic in the open unit disc and q times continuously differentiable on the closed unit disc. A norm can be defined so that $A_{q,n}$ is a Banach space. Let p be a positive integer, let $A(z)$ be an $n \times n$ matrix with columns in $A_{0,n}$, and define the operator $L: A_{1,n} \to A_{0,n}$ by

$$Ly(z) = z^p y'(z) + A(z)y(z). \tag{1.1}$$

The following theorem, due to Grimm and Hall [2], states necessary and sufficient conditions for the nonhomogeneous system $Ly = g$, $g \in A_{0,n}$, to have a solution in $A_{1,n}$.

THEOREM A. *The system $Ly = g$ has a solution in $A_{1,n}$ if and only if*

$$\lim_{r \to 1^-} B(g, f; r) = 0 \tag{1.2}$$

for all f belonging to the conjugate space $A^*_{0,n}$ such that

$$\lim_{r \to 1^-} B(Ly, f; r) = 0 \tag{1.3}$$

for all y in $A_{1,n}$.

If $u(z) = \sum_{k=0}^{\infty} u_k z^k$ and $v(z) = \sum_{k=0}^{\infty} v_k z^k$ are n-vector functions analytic in the unit disc, $B(u, v; z)$ denotes the Hadamard product of u and v, i.e., $B(u, v; z) = \sum_{k=0}^{\infty} (u_k \cdot v_k) z^k$, where $u_k \cdot v_k = u_k^1 v_k^1 + u_k^2 v_k^2 + \cdots + u_k^n v_k^n$. A detailed treatment of the relationship between the Hadamard product and the space $A^*_{0,1}$ has been given by Taylor [5], and his results were extended to the vector case by Grimm and Hall [2].

Equation (1.3) characterizes $K(L^*)$, the kernel of the conjugate operator L^*, and (1.2) characterizes the annihilator of $K(L^*)$. In this paper we shall study systems for which (1.3) has polynomial solutions, and we shall also study the relationship between the regular singular property at $z = 0$ and the existence of polynomials in $K(L^*)$ for such systems.

* Received by the editors November 18, 1975.

† Department of Mathematics and Statistics, University of Nebraska—Lincoln, Lincoln, Nebraska 68508. This research was supported in part under National Science Foundation Grant MPS75-06368.

2. Preliminaries. We first rewrite (1.3) in a more useful form. Let $f(z) = \sum_{k=0}^{\infty} f_k z^k$, $A(z) = \sum_{k=0}^{\infty} A_k z^k$, and T denote transpose. Then (1.3) is equivalent to the following infinite system of equations:

$$\begin{aligned} &\sum_{k=0}^{\infty} f_k^T A_k = 0 \\ &\sum_{k=0}^{\infty} f_{k+1}^T A_k = -f_p^T \\ &\qquad\vdots \\ &\sum_{k=0}^{\infty} f_{k+l}^T A_k = -l f_{p+l-1}^T \\ &\qquad\vdots \end{aligned} \tag{2.1}$$

Also, we shall assume that the first nonzero coefficient in the power series expansion of $A(z)$ is in Jordan normal form. This can be done without loss of generality, and will facilitate several proofs.

3. Results for nilpotent A_0. In this section we assume that the matrix A_0 is nilpotent. In this case, if the rank of A_0 is r,

$$A_0 = \operatorname{diag}\{J_1, \cdots, J_{n-r}\},$$

where each J_i is an elementary Jordan matrix of dimension ρ_i, $\sum_{i=1}^{n-r} \rho_i = n$. We can arrange the matrices J_i so that $\rho_1 \geqq \rho_2 \geqq \cdots \geqq \rho_{n-r}$. Hence A_0 is nilpotent of index ρ_1, and each J_i is nilpotent of index ρ_i. Now define the $n \times n$ matrices $\tilde{J}_i$, $i = 1, \cdots, n-r$, as the matrices formed by replacing each elementary Jordan matrix in A_0 except J_i by the zero matrix of corresponding dimension. Clearly, each $\tilde{J}_i$ is nilpotent of index ρ_i. In case $\rho_i = 1$, define $\tilde{J}^{\rho_i - 1}$ to be the $n \times n$ matrix with a one in the $((\sum_{j=1}^{i-1} \rho_j)+1, \sum_{j=1}^{i} \rho_j)$th position, and zeros elsewhere.

THEOREM 3.1. *If A_0 is as given above, and $p = 1$, then there exist $n-r$ linearly independent vectors f_{0_i} such that the constant functions $f_i(z) = f_{0_i}$ belong to $K(L^*)$, $i = 1, \cdots, n-r$. Also, no other polynomials belong to $K(L^*)$ in this case.*

Proof. Since the rank of A_0 is r, there exist exactly $n-r$ linearly independent vectors f_{0_i} such that $f_{0_i}^T A_0 = 0$. These vectors can be written as

$$f_{0_i} = (0, \cdots, 0, 1, 0, \cdots, 0)^T,$$

where the 1 is in the $(\sum_{j=1}^{i} \rho_j)$th position. Therefore the functions $f_i(z) = f_{0_i}$ satisfy (2.1), $i = 1, \cdots, n-r$.

Now suppose that $f(z) = \sum_{k=0}^{N} f_k z^k$ belongs to $K(L^*)$ with $N \geqq 1$ and $f_N \neq 0$. Then from (2.1) f_N must satisfy $f_N^T A_0 = -N f_N^T$. This is impossible since A_0 is nilpotent. If we let $N = 0$ then we obtain the same vectors as before, and so the proof is complete.

THEOREM 3.2. *Let A_0 be as given above and let $p \geqq 2$. If there exists a nonnegative integer N such that, for $i = 1, \cdots, n-r$,*

$$\tilde{J}_i^{\rho_i - 1} A_k = 0, \qquad k = 1, \cdots, N, \quad k \neq p-1,$$

and

$$\tilde{J}_i^{\rho_i-1}[A_{p-1}+(N-p+1)I]=0 \quad (\text{if } N\geqq p-1),$$

then there exist $n-r$ *linearly independent vectors* f_{N_i} *such that* $f_i(z)=f_{N_i}z^N$, $i=1,\cdots,n-r$, *belongs to* $K(L^*)$.

Proof. As in the proof of Theorem 3.1 we can define the vectors

$$f_{N_i}=(0,\cdots,0,1,0,\cdots,0)^T, \qquad i=1,\cdots,n-r,$$

such that $f_{N_i}^TA_0=0$. If $N=0$, then we are done. If $N>0$, the condition $\tilde{J}_i^{\rho_i-1}A_k=0$ is equivalent to requiring that the $(\sum_{j=1}^{i}\rho_j)$th row of A_k is all zeros, $i=1,\cdots,n-r$, $k=1,\cdots,N$, $k\neq p-1$. Hence $f_{N_i}^TA_k=0$ for $i=1,\cdots,n-r$ and $k=1,\cdots,N$, $k\neq p-1$. Similarly, the condition $\tilde{J}_i^{\rho_i-1}[A_{p-1}+(N-p+1)I]=0$ implies that $f_{N_i}^TA_{p-1}=-(N-p+1)f_{N_i}^T$ for $i=1,\cdots,n-r$. Hence, the functions $f_i(z)=f_{N_i}z^N$ belong to $K(L^*)$, and the proof is complete.

The two preceding theorems do not, however, completely describe $K(L^*)$ unless $\dim K(L^*)=n-r$. Since we know from [2] that

$$\dim K(L^*)=n(p-1)+\dim K(L), \tag{3.1}$$

then when $p=1$, $\dim K(L^*)=n-r$ if and only if $\dim K(L)=n-r$, and when $p=2$, $\dim K(L^*)=n-r$ if and only if $r=0$ and $\dim K(L)=0$.

4. Results for systems with a regular singular point. In this section we drop the assumption that A_0 is nilpotent. However, if $p\geqq 2$ and (1.1) has a regular singular point at $z=0$, Harris [3] has shown that A_0 must be nilpotent, and so the results of § 3 will apply to such systems.

THEOREM 4.1. *Let* $f(z)=\sum_{k=0}^{N}f_kz^k$ *belong to* $K(L^*)$ *with* $N>p-1$ *and every component of* f_N *nonzero. Then* $A(z)$ *must satisfy*:

(i) $$A_k=0, \qquad k=0,\cdots,p-2,$$

(ii) $$A_{p-1}=-(N-p+1)I,$$

and

(iii)
$$\begin{aligned}
&\sum_{k=p}^{N}f_k^TA_k=(N-p+1)f_{p-1}^T\\
&\qquad\cdots\\
&\sum_{k=p}^{N}f_{k+N-p-1}^TA_k=2f_{N-2}^T\\
&f_N^TA_p=f_{N-1}^T.
\end{aligned} \tag{4.1}$$

Proof. We first remark that when $p=1$ condition (i) is vacuously satisfied. Since f is a solution of (2.1) we must have, for $p\geqq 2$, $f_N^TA_0=0$. Because A_0 is in Jordan normal form, each eigenvalue of A_0 must occur at least once as the only nonzero element of some column. But this means that every eigenvalue of A_0 is zero, since no component of f_N is zero and $f_N^TA_0=0$. The same argument can now be used again to show that every element of the superdiagonal of A_0 is zero and

that, consequently, $A_0 = 0$. The same argument yields $A_1 = A_2 = \cdots = A_{p-2} = 0$ and (i) is proved.

Counting from the bottom up, the first equation from (2.1) for this f with a nonzero right-hand side is

$$f_{N-p+1}^T A_0 + \cdots + f_N^T A_{p-1} = -(N-p+1) f_N^T. \tag{4.2}$$

Condition (i) implies that (4.2) is equivalent to $f_N^T[A_{p-1} + (N-p+1)I] = 0$, and the same argument as before yields $[A_{p-1} + (N-p+1)] = 0$, or $A_{p-1} = -(N-p+1)I$. This proves (ii).

The remaining equations from (2.1) for this f are equivalent to (4.1). Since f belongs to $K(L^*)$, these equations must be satisfied by $A_p, \cdots, A_N$ and so (iii) holds.

The next theorem provides conditions which guarantee that $K(L^*)$ contains only polynomials. Further, these polynomials will be constructed.

THEOREM 4.2. *In* (1.1) *let* $p = 1$ *and let* $A_0 = -NI$, *N a positive integer. Then*

(i) $K(L^*) = \{f_i(z)\}$, $i = 1, \cdots, n$, *where the* f_i *are linearly independent polynomials of degree N,*

(ii) $\dim K(L) = n$, *and* $z = 0$ *is an apparent singularity* (*see* [1]) *for* $Ly = 0$.

Proof. Let $\{f_{N_i}\}$, $i = 1, \cdots, n$, be an arbitrary set of n linearly independent constant vectors. We can now uniquely define the vectors $\{f_{k_i}\}$, $k = N-1, \cdots, 0$, and $i = 1, \cdots, n$ by the system (4.1). The functions $f_i(z) = \sum_{k=0}^{N} f_{k_i} z^k$, $i = 1, \cdots, n$, all satisfy (2.1), and so belong to $K(L^*)$. Since the f_i are linearly independent, $\dim K(L^*) \geqq n$. From (3.1) we have $\dim K(L^*) = \dim K(L)$. But $\dim K(L) \leqq n$, so $\dim K(L^*) = \dim K(L) = n$. This proves (i). To complete the proof we note that $\dim K(L) = n$ implies that every fundamental matrix for $Ly = 0$ is analytic at $z = 0$. Hence $z = 0$ is an apparent singularity as defined in [1].

The condition that A_0 be a multiple of the identity matrix in Theorem 4.2 is quite restrictive. In the next theorem we find that a weaker hypothesis still guarantees the existence of polynomials in $K(L^*)$. However, these polynomials no longer span $K(L^*)$.

THEOREM 4.3. *Let* A_0 *have a nonpositive integer eigenvalue, let* $-N$ *be the largest such eigenvalue, and let m be the number of linearly independent eigenvectors of* A_0^T *corresponding to* $-N$. *Then if* $z = 0$ *is a regular singular point for* $Ly = 0$, *there exist m linearly independent polynomials of degree N which belong to* $K(L^*)$.

Proof. If $p \geqq 2$, then, since $z = 0$ is a regular singular point, A_0 is nilpotent, and so we have $N = 0$. Hence, if the rank of A_0 is r, then $m = n - r$ and we can apply Theorem 3.2.

Assume $p = 1$. Then $Ly = 0$ always has a regular singular point at $z = 0$. Let $\{f_{N_i}\}$, $i = 1, \cdots, m$, be m linearly independent eigenvectors of A_0^T corresponding to the eigenvalue $-N$, and therefore satisfying $f_{N_i}^T A_0 = -N f_{N_i}^T$. Now successively define the vectors $\{f_{k_i}\}$, $k = N-1, \cdots, 0$ and $i = 1, \cdots, m$, by

$$f_{k_i}^T = \left(-\sum_{j=1}^{N-k} f_{(j+k)_i}^T A_j \right) (A_0 + kI)^{-1}. \tag{4.3}$$

These vectors are uniquely defined for each i in terms of f_{N_i} and $A(z)$ since $-N$ is

the largest nonpositive integer eigenvalue of A_0. Hence the m polynomials $f_i(z)=\sum_{k=0}^{N} f_{k_i} z^k$, $i=1,\cdots,m$, belong to $K(L^*)$ and the proof is complete.

We shall now give two theorems which provide necessary and sufficient conditions for $K(L^*)$ to contain a nontrivial polynomial in the regular singular and irregular singular cases, respectively.

THEOREM 4.4. *Let $z=0$ be a regular singular point for $Ly=0$. Then $K(L^*)$ contains a nontrivial polynomial if and only if A_0 has a nonpositive integer eigenvalue.*

Proof. Assume $K(L^*)$ contains a nontrivial polynomial of degree N. If $p=1$, the coefficient of z^N, call it f_N, must satisfy $f_N^T A_0=-Nf_N^T$. Hence $-N$ is an eigenvalue of A_0. If $p\geqq 2$, A_0 is nilpotent by the result of Harris mentioned before, and so zero is an eigenvalue of A_0.

The converse is a direct application of Theorem 4.3.

THEOREM 4.5. *Let $z=0$ be an irregular singular point for $Ly=0$. Then $K(L^*)$ contains a nontrivial polynomial if and only if A_0 is singular.*

Proof. Since $z=0$ is an irregular singular point, we must have $p\geqq 2$.

Assume $f(z)=\sum_{k=0}^{N} f_k z^k$ belongs to $K(L^*)$, with $N\geqq 0$ and $f_N\neq 0$. Then (2.1) implies that $f_N^T A_0=0$ and hence A_0 is singular.

Assume A_0 is singular. Then there exists a nonzero vector, f_0, such that $f_0^T A_0=0$. Let $f(z)=f_0$. The quantity f satisfies (2.1) and hence belongs to $K(L^*)$.

5. Examples. To illustrate some of the preceding results, we consider the following linear second order equation, with a, b, and g in $A_{0,1}$:

$$z^2y''+za(z)y'+b(z)y=g(z). \tag{5.1}$$

This equation clearly has a regular singular point at $z=0$ and, as a system, has the form

$$\begin{bmatrix} z & 0 \\ 0 & z \end{bmatrix}\begin{bmatrix} y^1 \\ y^2 \end{bmatrix}'+\begin{bmatrix} 0 & -1 \\ b(z) & a(z)-1 \end{bmatrix}\begin{bmatrix} y^1 \\ y^2 \end{bmatrix}=\begin{bmatrix} 0 \\ g(z) \end{bmatrix}, \tag{5.2}$$

where $y^1=y$ and $y^2=zy'$. Hence the matrix A_0 is given by

$$A_0=\begin{bmatrix} 0 & -1 \\ b(0) & a(0)-1 \end{bmatrix}. \tag{5.3}$$

A_0, as given in (5.3), is not in Jordan normal form. However, since the eigenvalues of a constant matrix are invariant under similarity transformations, (5.3) will be used to calculate the eigenvalues of A_0. If J is the Jordan normal form of A_0, then there exists a nonsingular matrix P such that $J=P^{-1}A_0P$. In the remainder of this section, L will be the operator corresponding to the system obtained from (5.2) after A_0 has been converted to Jordan normal form.

If $b(0)=0$ and $a(0)=1$, then A_0 will be nilpotent of index two, and we can apply Theorem 3.1. In this case, any constant two-dimensional vector whose first component is zero belongs to $K(L^*)$, and so, by Theorem A, if $g(0)\neq 0$, then (5.1) is not solvable in $A_{1,2}$.

If $a(z)=1$ and $b(z)=z^2-\nu^2$, then (5.1) becomes the nonhomogeneous Bessel equation of order ν. In this case the eigenvalues of A_0, from (5.3), are $-\nu$

and ν and so, if ν is a positive integer, $K(L^*)$ contains a polynomial of degree ν. Moreover, this polynomial spans $K(L^*)$ since, from (3.1), $\dim K(L^*) = \dim K(L)$ and $\dim K(L) = 1$ for Bessel's equation. We shall now construct this polynomial using Theorem 4.3 with $\nu = N$.

In (5.2) let $Y = (y^1, y^2)^T$, let $U = PY$, and multiply both sides of (5.2) by P^{-1}, where

$$P = \begin{bmatrix} 1 & 1 \\ \nu & -\nu \end{bmatrix}.$$

System (5.2) then becomes

$$zU' + \left\{ \begin{bmatrix} -\nu & 0 \\ 0 & \nu \end{bmatrix} + \frac{1}{2\nu} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} z^2 \right\} U = G, \tag{5.4}$$

where $G = (1/(2\nu))(g, -g)^T$. The operator equation $LU = G$ will now refer to (5.4), and we shall rename A_0 so that $A_0 = \operatorname{diag}(-\nu, \nu)$.

The one linearly independent eigenvector of A_0^T corresponding to $-\nu$ can be written as $f_\nu = (1, 0)^T$. Then (4.3) yields

$$f_k^T = \frac{1}{2\nu} f_{k+2}^T \begin{bmatrix} \dfrac{1}{\nu - k} & \dfrac{-1}{\nu + k} \\ \dfrac{-1}{\nu - k} & \dfrac{1}{\nu + k} \end{bmatrix}$$

or, equivalently,

$$f_{\nu-2j}^T = \frac{1}{\nu!} \left(\frac{1}{2}\right)^{2j} \left(\frac{(\nu - j)!}{j!}, \frac{-(\nu - j - 1)!}{(j-1)!} \right), \qquad j = 1, \cdots, \left[\frac{\nu}{2}\right]. \tag{5.5}$$

Hence, $K(L^*)$ is spanned by the polynomial

$$f(z) = \binom{z^\nu}{0} + \sum_{j=1}^{[\nu/2]} f_{\nu-2j} z^{\nu-2j},$$

where $f_{\nu-2j}$ is given in (5.5).

If we now apply Theorem A to the system (5.4) we see that (5.4) has a solution in $A_{1,2}$ if and only if ν is a positive integer, and

$$g_\nu + \sum_{j=1}^{[\nu/2]} \left(\frac{(\nu - j)! + j(\nu - j - 1)!}{\nu! j! 2^{2j}} \right) g_{\nu-2j} = 0. \tag{5.6}$$

Clearly, (5.4) has a solution in $A_{1,2}$ if and only if (5.1), with $a(z) = 1$ and $b(z) = z^2 - \nu^2$, has a solution in $A_{1,1}$.

Remark. Condition (5.6) corresponds to a condition given by Ibragimov and Kušnirčuk [4], who obtained their result by using the equivalence of the Bessel and Euler operators. They also gave a solvability condition for certain nonhomogeneous Euler equations which we can obtain by using Theorem 4.3 and Theorem A as we did with Bessel's equation.

REFERENCES

[1] E. A. CODDINGTON AND N. LEVINSON, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.

[2] L. J. GRIMM AND L. M. HALL, *An alternative theorem for singular differential systems*, J. Differential Equations, 18 (1975), pp. 411–422.

[3] W. A. HARRIS, JR., *Characterization of linear differential systems with a regular singular point*, Proc. Edinburgh Math. Soc., 18 (1972), No. 2, pp. 93–98.

[4] I. I. IBRAGIMOV AND I. F. KUŠNIRČUK, *On the equivalence of Bessel and Euler operators in spaces of functions analytic in a disk*, Soviet Math. Dokl., 15 (1974), pp. 29–33.

[5] A. E. TAYLOR, *Banach spaces of functions analytic in the unit circle, I, II*, Studia Math., 11 (1950), pp. 145–170; Studia Math., 12 (1951), pp. 25–50.

SIAM J. MATH. ANAL.
Vol. 8, No. 5, October 1977

MEAN VALUE INEQUALITIES*

A. M. FINK†

Abstract. We find the best possible constants $K_i = K_i(\varphi)$, $i = 1, 2$, for inequalities of the kind

$$f(t)\int_0^t \varphi(f(s))\,ds \leqq K_i\varphi(f(t))\int_0^t f(s)\,ds$$

when φ is a given positive function, valid for all functions f such that $f(0) = 0$ and either ($i = 1$) f is increasing and convex, or ($i = 2$) f is increasing.

1. Introduction. Burton [1] in a recent paper needed an inequality of the sort

$$(*) \qquad \int_0^t y(s)\int_0^s (y'(u))^2\,du\,ds \leqq Ky^3(t) + \text{l.o.t.}, \qquad 0 \leqq t < 1,$$

where l.o.t. means a polynomial in $y(t)$ of degree $\leqq 2$ whose coefficients may depend on y, and where y, y', and $y'' \to +\infty$ as $t \to 1^-$.

If it is assumed that $y, y', y'' \geqq 0$ and $y(0) = 0$, then $\int_0^s (y'(u))^2\,du \leqq (\int_0^s y'(u)\,du)y'(s) = y(s)y'(s)$. An integration yields (*) with $K = 1/3$ and l.o.t. $\equiv 0$. However, Burton needed the result (*) for some constant $K < 1/3$. Several examples show that $K = 2/9$ ought to work.

One may proceed formally, searching for alternate forms of (*). The inequality (*) may be written, assuming the limit exists, and $y(0) = 0$,

$$\lim_{t\to 1^-} \frac{\int_0^t y(s)\int_0^s (y'(u))^2\,du\,ds}{y^3(t)} \leqq K.$$

In this form l'Hôpital's rule suggests showing

$$\frac{\int_0^t (y'(u))^2\,du}{3y(t)y'(t)} \leqq K.$$

In this form (*) becomes ($K = 2/9$)

$$(**) \qquad \int_o^t (y'(u))^2\,du \leqq 2/3y(t)y'(t).$$

Now, if (**) is true and $y(t) \geqq 0$ with $y(0) = 0$, then (*) follows directly with $K = 2/9$ and l.o.t. $\equiv 0$.

Now (**) with $y' = f$ and $y(0) = 0$ becomes

$$(1) \qquad \int_0^t f^2(s)\,ds \leqq 2/3f(t)\int_0^t f(s)\,ds.$$

When written in the form

$$\int_0^t 3f^2(s)\,ds \leqq f(t)\int_0^t 2f(s)\,ds$$

* Received by the editors June 14, 1975, and in revised form May 26, 1976.

† School of Mathematics, University of Minnesota, Minneapolis, Minnesota 55455.

an integration under the integral sign is suggested. So, assuming $f'>0$, and $f(0)=0$ and using parts

$$\int_0^t 3f^2(s)\,ds = \int_0^t 3f^2(s)f'(s)\frac{ds}{f'(s)}$$
$$= \frac{f^3(t)}{f'(t)} + \int_0^t f^3(s)\frac{f''(s)}{[f'(s)]^2}\,ds.$$

Similarly

$$\int_0^t 2f(s)\,ds = \frac{f^2(t)}{f'(t)} + \int_0^t f^2(s)\frac{f''(s)}{[f'(s)]^2}\,ds.$$

Thus the inequality (1) is equivalent to

$$\int_0^t \frac{f''(s)}{[f'(s)]^2} f^2(s)[f(s)-f(t)]\,ds \leqq 0.$$

This is true provided $f''\geqq 0$ since $f(s)-f(t)\leqq 0$. Thus (1) is verified if $f(0)=0$, $f'>0$, and $f''\geqq 0$, and (*) is verified for $K=2/9$ and l.o.t. $\equiv 0$ when $y(0)=0$, $y'(0)=0$, $y''>0$, $y'''\geqq 0$. This special case implies the general result (*) for functions y with $y'''\geqq 0$. For if y is given let $z=y+at+b$ with a, b chosen so that z satisfies the special conditions on $[t_0, 1)$; applying (*) to z yields (*) with y and l.o.t. $\not\equiv 0$.

The above formulation suggests the problem of finding a general class of inequalities (1). If one writes (1) as

$$\frac{1}{f^2(t)}\int_0^t f^2(s)\,ds \leqq \frac{K}{f(t)}\int_0^t f(s)\,ds$$

one can think of replacing the square by an arbitrary positive function, say φ. In the terminology to be introduced, the above arguments show that $K_1(\varphi)=2/3$ and $K_2(\varphi)=1$ with $\varphi(x)=x^2$. Different proofs are given below.

2. Formulation of the problem. We will show that (1) is one of an interesting class of inequalities and that 2/3 is the best possible constant.

Specifically, let all functions be defined on $[0,\infty)$ and consider the classes

$$A_1 \equiv \{f; f(0)=0, f\in C^2, f'>0, f''\geqq 0\},$$
$$A_2 = \{f; f(0)=0, f\in C', f'>0\}, \text{ and}$$
$$B = \{\varphi; \varphi(0)=0, \varphi'(0+) \text{ exists}, \varphi>0 \text{ on } (0,\infty)\}.$$

Fix $\varphi\in B$. We consider the inequalities

$$\frac{1}{\varphi(f(t))}\int_0^t \varphi(f(s))\,ds \leqq K_i(\varphi)\frac{1}{f(t)}\int_0^t f(s)\,ds. \tag{2}$$

We let $K_i(\varphi)$ denote the best possible constants for which (2) holds for all $f\in A_i$. In this terminology, $K_1(x^2)=2/3$, while $K_2(x^2)=1$ is an obvious inequality. In general $K_1(\varphi)\leqq K_2(\varphi)$ since $A_1\subset A_2$.

3. Reformulation of the problem. The inequalities (2) can be viewed as some sort of averages over the range space. It is thus desirable to rewrite (2) to maintain this viewpoint more clearly. Let $u = f(t)$ and change the variables in the integrals by $v = f(s)$. One then gets the equivalent inequality

$$\int_0^u [u\varphi(v) - K_i(\varphi)v\varphi(u)]g(v)\,dv \leqq 0, \tag{3}$$

where $g(v) = Tf = [f^{-1}(v)]' = [f'(f^{-1}(v))]^{-1}$. The values u for which (3) is to hold depends on the range of f. Note that the domain of f is irrelevant here. It is clear that the sets A_i are transformed by the mapping T into the sets

$$\hat{A}_1 \equiv \{g \mid g \in C^1[0, \infty),\ g > 0,\ g' \leqq 0\}$$

and

$$\hat{A}_2 = \{g \mid g \in C[0, \infty),\ g > 0\}.$$

In fact the mapping is onto. Since $g \in \hat{A}_i$ is continuous, and positive one can solve the equation $g = (f^{-1})'$, $f^{-1}(0) = 0$ by $f^{-1}(x) = \int_0^x g(s)\,ds$ and observe that f^{-1} has an inverse function whose derivative is continuous and positive and if $g \in \hat{A}_1$, f^{-1} is concave so that f is convex.

Thus $K_i(\varphi)$ for inequalities (2) are also the best possible constants for which (3) holds for all $g \in \hat{A}_i$. We may assume the domain of these functions is $[0, \infty)$.

4. Characterizations of $K_i(\varphi)$. The problem of finding K_2 is straightforward.

LEMMA 1. *Let h be continuous. Then $\int_a^b h(u)g(u)\,du \leqq 0$ for all $g \in \hat{A}_2$ if and only if $h(u) \leqq 0$ on $[a, b]$.*

Proof. If $A \equiv \{u \mid h(u) \leqq 0\}$ is not $[a, b]$, then $k(x) = \operatorname{dist}(x, A) \geqq 0$ and $\int_a^b k(x)h(x)\,dx > 0$. For small $\varepsilon > 0$, $g(x) = k(x) + \varepsilon$ is in $\hat{A}_2$ and $\int_a^b h(u)g(u)\,du > 0$; thus $A = [a, b]$. The converse is trivial.

It now follows that $K_2(\varphi)$ is the smallest constant K such that

$$u\varphi(v) - Kv\varphi(u) \leqq 0, \qquad 0 \leqq v \leqq u < \infty. \tag{4}$$

THEOREM 1. *If $\varphi \in B$, then*

$$K_2(\varphi) = \sup_{u>0} \frac{u}{\varphi(u)} \sup_{0<v\leqq u} \frac{\varphi(v)}{v},$$

where $K_2(\varphi) = +\infty$ is interpreted that no constant exists for (2).

The K_1 inequality requires a different lemma.

LEMMA 2. *Let h be continuous on $[a, b]$. Then $\int_a^b h(u)g(u)\,du \leqq 0$ for all $g \in \hat{A}_1$ if and only if*

$$H(u) \equiv \int_a^u h(s)\,ds < 0 \quad \text{on} \quad [a, b].$$

Proof. If $g \in \hat{A}_1$, then by the mean value theorem, there is $\xi \in (a, b)$ such that

$$\begin{aligned}\int_a^b h(u)g(u)\,du &= g(a)\int_a^\xi h(u)\,du + g(b)\int_\xi^b h(u)\,du \\ &= [g(a) - g(b)]H(\xi) + g(b)H(b).\end{aligned}$$

Now $g(a) \geqq g(b) > 0$ and $H(\xi) \leqq 0$, $H(b) \leqq 0$ so $\int_a^b h(u)g(u)\,du \leqq 0$. Conversely, if $A \equiv \{x | H(x) > 0\}$ is nonempty, then $\int_a^b H\chi_A = \eta > 0$. Find a continuous function k such that $k(x) \leqq 0$ and $\int_a^b |k+\chi_A|^2 \leqq (\eta^2/4\{\int_a^b (H)^2\})$. Then

$$\int_a^b kH = \int_a^b (k+\chi_A)H - \int_a^b \chi_A H$$

$$\leqq \left(\int_a^b |k+\chi_A|^2\right)^{1/2}\left(\int_a^b (H)^2\right)^{1/2} - \eta \leqq -\eta/2.$$

Now let g be defined by $g'(x) = k(x)$ and $g(b) = \varepsilon > 0$. Then $g \in \hat{A}_1$ and

$$\int_a^b h(u)g(u)\,du = g(b)H(b) - \int_a^b H(u)g'(u)\,du$$

$$\geqq \varepsilon H(b) + \eta/2 > 0 \quad \text{if } \varepsilon \text{ is sufficiently small.}$$

Thus $A = \phi$ and $H \leqq 0$.

It now follows that $K_1(\varphi)$ is the smallest constant K such that

$$\int_0^v [u\varphi(s) - Ks\varphi(u)]\,ds \leqq 0, \qquad 0 \leqq v \leqq u < \infty.$$

Let $\Phi(u) = \int_0^u \varphi(s)\,ds$. Then this condition is $(\Phi(v)/v^2)\,(u/\varphi(u)) \leqq K$, $0 < v \leqq u$.

THEOREM 2. *Let $\varphi \in B$, then*

$$K_1(\varphi) = \sup_{u>0} \frac{u}{\varphi(u)} \sup_{0<v\leqq u} \frac{\Phi(v)}{v^2}.$$

If $K_1(\varphi) = +\infty$, then there is no inequality (2).

5. Corollaries and examples. The most accessible examples are afforded by the functions $\varphi(t) = t^\alpha$, $\alpha \geqq 1$. It follows that $K_2(t^\alpha) = 1$ and $K_1(t^\alpha) = 2(\alpha+1)^{-1}$. There are other classes of functions for which these constants can either be computed exactly or estimated. Again we let $\Phi(t) = \int_0^t \varphi(s)\,ds$.

We assume in this section that $\varphi \in B$.

COROLLARY 1. *If $\varphi(t)t^{-1}$ is nondecreasing (φ is convex for example), then $K_2(\varphi) = 1$. If $\varphi(t)t^{-1}$ is nonincreasing then $K_2(\varphi) = \varphi'(0)\lim_{t\to\infty}(t/\varphi(t))$.*

These follow directly from Theorem 1.

Example 1. Let $\varphi(t) = 1 + t - e^{-kt}$, $k > 0$; then $\varphi(t)t^{-1}$ is nonincreasing and $K_2(\varphi) = 1 + k$.

Example 2. Let $\varphi(t) = t(1+Bt)/(1+t)$, $0 < B < 1$. Then $\varphi(t)t^{-1}$ is nonincreasing and $K_2(\varphi) = B^{-1}$. Also $\Phi(t)t^{-2} = B/2 + (1-B)\sigma(t)$ • where $\sigma(t) = t^{-1} - t^{-2}\log(1+t)$. It can be shown that σ is decreasing. Therefore $\sup_{0<v\leqq u} v^{-2}\Phi(v) = \lim_{v\to 0^+} \Phi(v)v^{-2} = 1/2$. Thus $K_1(\varphi) = B^{-1}$.

Remark 1. Any positive number can be a $K_1(\varphi)$ since Example 2 gives any number $\geqq 1$ and $K_1(t^\alpha)$ give the numbers in $(0, 1]$.

COROLLARY 2. *For any φ, $K_2(\varphi) \geqq 1$ and any number in $[1, \infty)$ is a $K_2(\varphi)$.*

Proof. Example 2 gives the range of $K_2(\varphi)$ as including $[1, \infty)$. On the other hand, if $K_2(\varphi) \leqq 1$, then $(u/\varphi(u))\,(\varphi(v)/v) \leqq 1$, $0 < v < u$. This implies that $\varphi(t)t^{-1}$ is nondecreasing which by Corollary 1 implies that $K_2(\varphi) = 1$.

COROLLARY 3. *If φ is concave, then* $K_1(\varphi) = K_2(\varphi) = \varphi'(0+) \lim_{t\to\infty} (t/\varphi(t)) \geqq 1$.

Proof. If φ is concave then $\varphi(t)t^{-1}$ and $\Phi(t)t^{-2}$ are nonincreasing. The inequality follows from Corollary 2.

Note that Example 2 is a case in point.

COROLLARY 4. *If $\varphi(v)v^{-2}$ is nondecreasing, then* $K_1(\varphi) = 2 \sup_{t>0}[\Phi(t)/(t\varphi(t))]$.

In the situation of Corollary 4, the function $f(t) = t$ is an extremal, since when this f is put into (2), then K_1 is the smallest number for which that inequality holds.

It is worthwhile to look at the class of functions which satisfy the hypotheses of Corollary 4. For this purpose, consider the class

$$B_2 = \left\{\varphi \in B;\ \varphi \text{ is continuous, } (1/t)\int_0^t \varphi(s)\,ds \leqq \tfrac{1}{2}\varphi(t) \text{ for all } t>0\right\}.$$

Since $[\Phi(t)t^{-2}]' = t^{-3}[t\varphi(t) - 2\Phi(t)]$, $\varphi(t)t^{-2}$ is nondecreasing if and only if $\varphi \in B_2$.

THEOREM 3. *Let φ be continuous. Then $K_1(\varphi) \leqq 1$ if and only if $\varphi \in B_2$. Consequently, if $\varphi \in B_2$ then*

$$K_1(\varphi) = 2 \sup_{t>0} \frac{\Phi(t)}{t\varphi(t)}.$$

Proof. If $K_1(\varphi) \leqq 1$ then let $f(t) = t$ and (2) becomes the defining inequality of B_2. On the other hand, if $\varphi \in B_2$, then $\varphi(t)t^{-2}$ is nondecreasing and Corollary 4 applies. The defining inequality of B_2 gives $K_1(\varphi) \leqq 1$.

COROLLARY 5. *If φ is convex and continuous, then*

$$K_1(\varphi) = 2 \sup_{t>0} \frac{\Phi(t)}{t\varphi(t)} \leqq 1.$$

This follows from the fact that if φ is convex then $\varphi \in B_2$.

The idea of the set B_2 may be extended. Consider for $n \geqq 2$,

$$B_n \equiv \left\{\varphi \in B;\ \varphi \text{ is continuous, } \frac{1}{t}\int_0^t \varphi(s)\,ds \leqq 1/n\varphi(t) \text{ for all } t>0\right\}.$$

COROLLARY 6. *If φ is continuous, then $K_1(\varphi) \leqq 2/n$ if and only if $\varphi \in B_n$.*

The proof is a paraphrase of the proof of Theorem 3.

A related class is given by

$$D_n \equiv \{\varphi \mid \varphi \text{ is continuous and } 0 < \varphi(s) \leqq (s/t)^n \varphi(t) \text{ for } 0 < s \leqq t\}.$$

Observe that $\varphi \in D_n$ implies $\varphi(0) = 0$, and $\varphi'(0) = 0$ exists, so that $\varphi \in B$.

LEMMA 3. *Let φ be continuous, then*:

(i) *$\varphi \in B_n$ if and only if $\Phi \in D_n$.*

(ii) *$D_n \subset B_{n+1}$.*

(iii) *If $\varphi \in D_n$, then $\Phi \in D_{n+1}$.*

Proof. If $\varphi \in B_n$, then $\Phi(t) \leqq t/n\Phi'(t)$. One solves this differential inequality to get $\Phi \in D_n$. Conversely, if $\Phi \in D_n$ then $\Phi(t)t^{-n}$ is nondecreasing. Its derivative is nonnegative. This gives $\varphi \in B_n$. If $\varphi \in D_n$, then integrate the defining inequality

with respect to s from 0 to $u \leqq t$ to get

$$\frac{\Phi(u)}{u^2}\frac{t}{\varphi(t)} \leqq \left(\frac{u}{t}\right)^{n-1}\frac{1}{n+1} \leqq \frac{1}{n+1}.$$

This implies by Theorem 2 that $K_1(\varphi) \leqq 2(n+1)^{-1}$, so $\varphi \in B_{n+1}$. Now (i) and (ii) combine to give (iii).

THEOREM 4. *Suppose φ is $(n+1)$ times continuously differentiable with $\varphi^{(j)}(0) = 0$, $j = 0, \cdots, n-1$, $\varphi^{(n)}(0) \geqq 0$, and $\varphi^{(n+1)} > 0$ for $t > 0$. Then $\varphi \in D_n$ and $K_1(\varphi) \leqq 2(n+1)^{-1}$. If further $\varphi^{(n)}(0) \neq 0$, then $K_1(\varphi) = 2(n+1)^{-1}$.*

Proof. Consider the auxiliary function $\psi(u) = t^n\varphi(u) - u^n\varphi(t)$ on $0 \leqq u \leqq t$. Then $\psi^{(j)}(0) = 0$, $j = 0, \cdots, n-1$, $\psi(t) = 0$. Also $\psi^{(n)}(0) = t^n\varphi^{(n)}(0) - n!\varphi(t) = -[t^{n+1}/(n+1)]\varphi^{(n+1)}(\xi)$ for some $\xi > 0$ by Taylor's theorem. Thus $\psi^{(n)}(0) < 0$. Furthermore, $\psi^{(n+1)}(u) = t^n\psi^{(n+1)}(u) > 0$ on $(0, t)$. Thus ψ can have no more than $n+1$ zeros on $[0, t]$. Since $n+1$ are accounted for, ψ is nonzero on $(0, t)$. Since $\psi^{(n)}(0) < 0$, $\psi < 0$ on $(0, t)$ and $\varphi \in D_n$. Now $\varphi \in B_{n+1}$ so by Corollary 6, $K_1(\varphi) \leqq 2/(n+1)$ but by l'Hôpital's rule

$$\lim_{t\to 0^+}\frac{\Phi(t)}{t\varphi(t)} = \lim_{t\to 0^+}\frac{\varphi(t)}{\varphi(t)+t\varphi'(t)} = \lim_{t\to 0^+}\frac{\varphi^{(n)}(t)}{(n+1)\varphi^{(n)}(t)+t\varphi^{(n+1)}(t)} = \frac{1}{n+1}$$

if $\varphi^{(n)}(0) \neq 0$. So $K_1(\varphi) \geqq 2/(n-1)$.

Example 3. To construct an example for Theorem 4 let f be a function with $f^{(n)}(0) \geqq 0$ and $f^{(n+1)}(t) > 0$ on $(0, \infty)$. Define $\varphi(t) = f(t) - \sum_{j=0}^{n-1} f^{(j)}(0)t^j/j!$. In particular $\varphi_n(t) = e^t - \sum_{j=0}^{n-1} t^j/j! \in D_n$ and $k_1(\varphi_n) = 2(n+1)^{-1}$.

COROLLARY 7. *If $\varphi \in D_n$, then $K_2(\varphi) = 1$.*

Proof. The defining inequality gives $\sup_{0<v<u} \varphi(v)/v = \varphi(u)/u$.

6. Remark. It might be of interest to ask the question whether for a fixed $f \in A_i$, one can get an inequality of the form

$$\frac{1}{\varphi(f(t))}\int_0^t \varphi(f(s))\,ds \leqq M(f)\frac{1}{f(t)}\int_0^t f(s)\,ds \tag{5}$$

to hold for some class of φ's, say for $\varphi \in B$ and convex. The transformed problem becomes

$$u\int_0^u \varphi(v)g(v)\,dv \leqq M\varphi(u)\int_0^u vg(v)\,dv.$$

Consider u fixed and $M(u)$ the best constant. Then $M(f) = \sup_{u>0} M(u)$. But for u fixed, we may assume $\varphi(u) = 1$ since the inequality is homogeneous in φ. Then $M(u)$ is merely the $\sup_{\varphi\in B,\text{convex}} u\int_0^u \varphi(v)g(v)\,dv / \int_0^u vg(v)\,dv$. Note that $g > 0$ so that the sup is computed by taking the pointwise sup of the functions, i.e. $\varphi(t) = t/u$ in which case $M(u) = 1$ and $M(f) = 1$. In general, if $C \subset B$ and φ is required to be in C in computing M, then one needs to take the "maximal" element in C in the above argument, provided C is closed under scalar multiples. For example, if C is the concave functions,

$$M(u) = \frac{u\int_0^u g(v)\,dv}{\int_0^u vg(v)\,dv}$$

so that

$$M(f) = \sup_{u>0} M(u) = \sup_{t>0} \frac{f(t)\int_0^t ds}{\int_0^t f(s)\,ds} = \sup_{t>0} \frac{tf(t)}{F(t)},$$

where $F(t) = \int_0^t f(s)\,ds$. If $f \in A_1$, then

$$M^{-1}(f) = \inf_{t>0} \frac{F(t)}{tf(t)} \leqq \sup_{t>0} \frac{F(t)}{tf(t)} = \tfrac{1}{2}K_1(f).$$

REFERENCE

[1] T. A. BURTON, *Noncontinuation of solutions of differential equations of order n*, to appear.

SIAM J. MATH. ANAL.
Vol. 8, No. 5, October 1977

EXISTENCE AND UNIQUENESS FOR PERIODIC SOLUTIONS OF THE BENJAMIN–BONA–MAHONY EQUATION*

L. A. MEDEIROS AND G. PERLA MENZALA†

Abstract. We consider the problem $u_t - u_{xxt} + uu_x = 0$ in $-\infty < x, t < \infty$ with initial data at $t = 0$ which is 1-periodic and the boundary condition $u(x+1, t) = u(x, t)$ for all x, t; proving the existence and uniqueness of the solutions of such a problem. We use the semi-discrete approach together with the energy method.

1. Introduction. In this paper we shall prove an existence and uniqueness theorem for regular solutions of the equation

$$u_t + uu_x - u_{xxt} = 0 \tag{1.1}$$

on $-\infty < x < +\infty$, $t \geqq 0$ with the initial data

$$u(x, 0) = u_0(x), \qquad -\infty < x < \infty, \tag{1.2}$$

and the boundary condition

$$u(x+1, t) = u(x, t) \quad \text{for all } x \text{ and } t. \tag{1.3}$$

The equation (1.1) appears in fluid mechanics and was proposed by Benjamin–Bona–Mahony in the study of water waves (see [8]). It is usually called the BBM equation and is a modification of the Korteweg–de Vries equation $u_t + uu_x + u_{xxx} = 0$, which has been studied intensively in recent years.

To prove existence and uniqueness for regular solutions of problem (1.1)–(1.2)–(1.3), we use the differential-difference method as done by Sjöberg in [1] for the Korteweg–de Vries equation. However, in the case of BBM it is much simpler. More precisely, our main result is:

THEOREM 1. *Suppose $u_0 : \mathbb{R} \to \mathbb{R}$ = real numbers, with $u_0(x+1) = u_0(x)$ for all $x \in \mathbb{R}$, u_0 three times differentiable and d^3u_0/dx^3 Riemann square integrable on $0 \leqq x \leqq 1$. Then there exists one and only one function $u : \mathbb{R} \times [0, +\infty) \mapsto \mathbb{R}$ satisfying the following conditions:*

(i) *u has all derivatives in t and is twice continuously differentiable in x,*
(ii) *$u_t + uu_x - u_{xxt} = 0$ pointwise on $\mathbb{R} \times [0, \infty)$,*
(iii) *$u(x, 0) = u_0(x)$ for all $x \in \mathbb{R}$,*
(iv) *$u(x+1, t) = u(x, t)$ for all x and t.*

Some related results concerning (1.1) can be found in [2] and [6].

2. Differential-difference scheme. In this section we solve a discrete problem associated with (1.1), with discretization only on the space variable x. We will obtain a system of ordinary differential equations, which has a regular solution in t for all $t \geqq 0$. With the solution of this system, we shall prove in the next section the existence of solutions for (1.1) by discrete Fourier series and the Arzelá–Ascoli theorem.

* Received by the editors March 25, 1975, and in final revised form May 17, 1976.

† Instituto de Matemática, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil. This work was supported by FINEP, CEPG-UFRJ and by CBPF.

Let $n>0$ be a natural number and let us take $N=2n+1$ and consider a decomposition of the interval $[0, 1]$ in N equal parts, each one with length $h=1/N$. The point rh of $[0, 1]$ is represented by x_r for all $r=1, 2, \cdots, N$. The difference operators D_+, D_-, and D_0 are defined as follows: $hD_+u_N(x_r, t)=u_N(x_{r+1}, t)-u_N(x_r, t)$, $hD_-u_N(x_r, t)=u_N(x_r, t)-u_N(x_{r-1}, t)$, and $2hD_0u_N(x_r, t)=u_N(x_{r+1}, t)-u_N(x_{r-1}, t)$. For technical reasons, we write the (1.1) as $u_t+\frac{1}{3}[uu_x+(u^2)_x]-u_{xxt}=0$. Let us consider the following discrete problem: Find the grid function $u_N(x_r, t)$, $r=1, 2, \cdots, N$, $t\geqq 0$, such that the following conditions are satisfied:

$$\text{(a)}\quad \frac{\partial}{\partial t}u_N(x_r, t)+\frac{1}{3}[u_N(x_r, t)D_0u_N(x_r, t)+D_0u_N^2(x_r, t)]$$

$$-D_+D_-\frac{\partial}{\partial t}u_N(x_r, t)=0,$$

$$\text{(b)}\quad u_N(x_r, 0)=u_0(x_r),$$

$$\text{(c)}\quad u_N(x_{r+N}, t)=u_N(x_r, t),$$

for $r=1, 2, \cdots, N$.

The set (a), (b), and (c) is a system of differential-difference equations which has a solution on a time-interval (depending on N) which we will denote by $[0, t_N)$. This solution is defined at the point (x_r, t) of the grid for all t, $0<t<t_N$, and has derivatives of all orders with respect to t. By a priori estimates we shall prove that the solution exists on an interval (independent of N) which we will denote with $[0, T)$. It will be seen that the argument can be repeated indefinitely so establishing the existence of a solution over an arbitrary time-interval. First, we define a discrete inner product in the space of gridfunctions by

$$(u, v)_h=\sum_{r=1}^{N}\overline{u(x_r)}v(x_r)h$$

where $\bar{u}$ is the complex conjugate of u. The discrete norm is defined by $\|u\|_h^2=(u, u)_h$. By $(\cdot, \cdot)$ and $\|\cdot\|$ we represent the usual inner product and norm given by

$$(u, v)=\int_0^1 \bar{u}v\,dx \quad \text{and} \quad \|u\|^2=\int_0^1 |u|^2\,dx.$$

Observation 1. We have the following relations:

$$(u, D_+v)_h=-(D_-u, v)_h, \qquad (u, D_-v)_h=-(D_+u, v)_h,$$

and $(u, D_0v)_h=-(D_0u, v)_h$ for all N-periodic gridfunctions u, v.

LEMMA 2.1. *Let $u_N(x_r, t)$, $r=1, 2, \cdots, N$, be a solution of the discrete problem* (a), (b), (c) *on the interval* $[0, t_N)$. *Then there exists a constant $C>0$ independent of N such that*

$$\|u_N(\cdot, t)\|_h^2+\|D_-u_N(\cdot, t)\|_h^2<C$$

for all N and all t in $[0, t_N)$.

Proof. Taking the inner product of the equation (a) with $u_N(x_r, t)$ we obtain

$$\left(u_N, \frac{\partial u_N}{\partial t}\right)_h + \frac{1}{3}[(u_N, u_N D_0 u_N)_h + (u_N, D_0 u_N^2)_h] - \left(u_N, D_+ D_- \frac{\partial u_N}{\partial t}\right)_h = 0 \tag{2.1}$$

which is true for each t in $[0, t_N)$. By the definition of the discrete inner product $(\cdot, \cdot)_h$ it follows that $(u_N, u_N D_0 u_N)_h + (u_N, D_0 u_N^2)_h = 0$. Also, by Observation 1 above (2.1) reduces to

$$\frac{\partial}{\partial t}[\|u_N\|_h^2 + \|D_- u_N\|_h^2] = 0. \tag{2.2}$$

By integrating from 0 to $t < t_N$ we obtain

$$\|u_N(x_r, t)\|_h^2 + \|D_- u_N(x_r, t)\|_h^2 = \|u_0(x_r)\|_h^2 + \|D_- u_0(x_r)\|_h^2.$$

But, since $\|u_0(x_r)\|_h^2 \leqq 2\|u_0\|^2$ and $\|D_- u_0(x_r)\|_h^2 \leqq 2\|du_0/dx\|^2$ for small h (recall that u_0 has all the same hypotheses as in Theorem 1), the proof of the lemma is complete.

Observation 2. Since $D_+ D_- = D_- D_+$ and $2D_0 = D_+ + D_-$ it follows just be rewriting Lemma 2.1 that $\|D_+ u_N(\cdot, t)\|_h^2$ and $\|D_0 u_N(\cdot, t)\|_h^2$ are bounded by a constant on $[0, t_N)$ independent of N.

LEMMA 2.2. *Let u_N be as in the above Lemma* 2.1; *then there exist constants C_j, $j = 1, 2, 3, 4$, independent of N such that*

(i) $\|u_N(x_r, t)| < C_1$ *for all* r,

(ii) $\left\|\frac{\partial u_N}{\partial t}\right\|_h^2 + \left\|D_- \frac{\partial u_N}{\partial t}\right\|_h^2 < C_2$,

(iii) $\left\|D_+ D_- \frac{\partial u_N}{\partial t}\right\|_h^2 < C_3$,

(iv) $\left\|\frac{\partial^2 u_N}{\partial t^2}\right\|_h^2 + \left\|D_- \frac{\partial^2 u_N}{\partial t^2}\right\|_h^2 < C_4$

for all t *in* $[0, t_N)$.

Proof. To prove the estimate (i) we consider the discrete Sobolev inequality which states that, for each $\varepsilon > 0$ there exists $C = C(\varepsilon) > 0$ such that

$$\max_{1 \leqq r \leqq N} |u_N(x_r, t)| \leqq \varepsilon \|D_- u_N\|_h^2 + C(\varepsilon)\|u_N\|_h^2;$$

and from this inequality and Lemma 2.1 we obtain estimate (i). To prove (ii), let us consider the inner product of (a) with $\partial u_N/\partial t$ to get

$$\left\|\frac{\partial u_N}{\partial t}\right\|_h^2 + \left\|D_- \frac{\partial u_N}{\partial t}\right\|_h^2 \leqq \frac{1}{3}\left|\left(u_N D_0 u_N, \frac{\partial u_N}{\partial t}\right)_h\right| + \frac{1}{3}\left|\left(D_0 u_N^2, \frac{\partial u_N}{\partial t}\right)_h\right|. \tag{2.3}$$

We observe that

$$\begin{aligned}\frac{1}{3}\left|\left(u_N D_0 u_N, \frac{\partial u_N}{\partial t}\right)_h\right| &\leqq \frac{1}{3}\max_{1\leqq r\leqq N}|u_N(x_r,t)|\left(|D_0 u_N|, \left|\frac{\partial u_N}{\partial t}\right|\right)_h \\ &\leqq \frac{1}{3}\max_{1\leqq r\leqq N}|u_N(x_r,t)|\,\|D_0 u_N\|_h\left\|\frac{\partial u_N}{\partial t}\right\|_h \leqq C_\varepsilon + \frac{1}{18}\left\|\frac{\partial u_N}{\partial t}\right\|_h^2\end{aligned} \tag{2.4}$$

because of the Lemma 2.1(i), and the elementary inequality $2\alpha\beta \leqq \alpha^2+\beta^2$ for any pair of real numbers α and β. Similarly, we have

$$\begin{aligned}\frac{1}{3}\left|\left(D_0 u_N^2, \frac{\partial u_N}{\partial t}\right)_h\right| &\leqq 2\max_{1\leqq r\leqq N}|u_N(x_r,t)|\,\|D_0 u_N\|_h \frac{1}{3}\left\|\frac{\partial u_N}{\partial t}\right\|_h \\ &\leqq 4C_\varepsilon + \frac{1}{18}\left\|\frac{\partial u_N}{\partial t}\right\|_h^2\end{aligned} \tag{2.5}$$

where C_ε is positive constant. Now, (2.4) and (2.5) together with (2.3) complete the proof of (ii).

To prove (iii), we obtain from (a):

$$\left\|D_+D_-\frac{\partial u_N}{\partial t}\right\|_h \leqq \left\|\frac{\partial u_N}{\partial t}\right\|_h + \frac{1}{3}\|u_N D_0 u_N + D_0 u_N^2\|_h \tag{2.6}$$

which shows (iii) because the right side is bounded by a constant independent of N for all t in $[0, t_N)$. In fact, $\frac{1}{3}(u_N D_0 u_N + D_0 u_N^2) = u_N D_0 u_N$ which can be estimated in the norm $\|\cdot\|_h$ in terms of $\max_{1\leqq r\leqq N}|u_N(x_r,t)|$ and $\|D_0 u_N\|_h$.

Now, we take the derivative with respect to t of equation (a). To simplify the notation let us write $v_N = \partial u_N/\partial t$ and $w_N = \partial v_N/\partial t$. Then we have

$$w_N - D_+D_-w_N = -\tfrac{1}{3}v_N D_0 u_N - \tfrac{1}{3}u_N D_0 v_N - \tfrac{2}{3}D_0(u_N v_N). \tag{2.7}$$

Taking the inner product with w_N we get

$$\begin{aligned}\|w_N\|_h^2 + \|D_-w_N\|_h^2 &= -\tfrac{1}{3}(v_N D_0 u_N, w_N)_h - \tfrac{1}{3}(u_N D_p v_N, w_N)_h \\ &\quad -\tfrac{2}{3}(D_0(u_N v_N), w_N)_h.\end{aligned} \tag{2.8}$$

We observe that

$$\begin{aligned}\tfrac{1}{3}|(v_N D_0 u_N, w_N)_h| &\leqq \|v_N D_0 u_N\|_h \tfrac{1}{3}\|w_N\|_h \\ &\leqq \tfrac{1}{2}\|v_N D_0 u_N\|_h^2 + \tfrac{1}{18}\|w_N\|_h^2.\end{aligned} \tag{2.9}$$

Since $\|v_N D_0 u_N\|_h \leqq \max_{1\leqq r\leqq N}|v_N(x_r,t)|\,\|D_0 u_N\|_h$, it follows that $\|v_N D_0 u_N\|_h^2$ is bounded, because of the discrete Sobolev inequality and the previous discussion. Similarly

$$\tfrac{1}{3}|(u_N D_0 v_N, w_N)_h| \leqq \tfrac{1}{2}\|u_N D_0 v_N\|_h^2 + \tfrac{1}{18}\|w_N\|_h^2. \tag{2.10}$$

Also, $\|u_N D_0 v_N\|_h \leqq \max_{1\leqq r\leqq N}|u_N(x_r,t)|\,\|D_0 v_N\|_h$ which shows that $\|u_N D_0 v_N\|_h$ is bounded.

Finally, we observe that

$$\tfrac{2}{3}|(D_0(u_Nv_N),\, w_N)_h| \leqq \tfrac{1}{3}\|D_0(u_Nv_N)\|_h^2 + \tfrac{2}{9}\|w_N\|_h^2. \tag{2.11}$$

The first term of the right side of (2.11) is bounded because $2D_0 = D_+ + D_-$ and by the previous inequalities. Now, because of (2.9), (2.1), and (2.11) together with (2.8) the last part, (iv), of the lemma is proved.

LEMMA 2.3. *There exist constants C_5, C_6 and C_7, independent of N, such that*

$$\text{(i)} \quad \|D_+D_-^2 u_N\|_h < C_5,$$

$$\text{(ii)} \quad \|D_-D_0 u_N\|_h < C_6,$$

$$\text{(iii)} \quad \left\|D_+D_-^2 \frac{\partial u_N}{\partial t}\right\| < C_7$$

for all t in $[0, t_N)$.

Proof. In fact, let us apply D_- to equation (a) to get

$$-D_+D_-^2 \frac{\partial u_N}{\partial t} = -\frac{1}{3}D_-(u_ND_0u_N) - D_-\frac{\partial u_N}{\partial t}.$$

Therefore,

$$\left\|D_+D_-^2 \frac{\partial u_N}{\partial t}\right\|_h \leqq \left\|D_-\frac{\partial u_N}{\partial t}\right\|_h + \frac{1}{3}\|D_-(u_ND_0u_N) + D_-(D_0u_N^2)\|_h. \tag{2.12}$$

The first term of the right side is already bounded. To bound the second term, let us represent it by $\frac{1}{3}X$. We have, by the definition of $\|\cdot\|_h$,

$$X^2 = \sum_{r=1}^{N} |D_-(u_rD_0u_r) + D_-(D_0u_r^2)|^2 h$$

where we have written $u_r = u_N(x_r, t)$. Using the fact that $u_{r+1}D_0u_r + u_{r-1}D_0u_r = D_0u_r^2$ we have by direct calculation

$$\begin{aligned} |D_-(u_rD_0u_r) + D_-(D_0u_r^2)| &= |D_-(u_rD_0u_r) + D_-(u_{r+1}D_0u_r) + D_-(u_{r-1}D_0u_r)| \\ &= |u_{r-1}D_-D_0u_r + u_rD_-D_0u_r + D_+u_rD_0u_r \\ &\quad + D_-u_{r-1}D_0u_{r-1} + D_-u_rD_0u_r + u_{r-1}D_-D_0u_r|; \end{aligned}$$

therefore,

$$\begin{aligned} X^2 \leqq \text{const.} \left(\max_{1\leqq r\leqq N} |u_r|\right)^2 \sum_{r=1}^{N} &(|D_-D_0u_r|^2 + |D_-u_r|^2|D_0u_r|^2 \\ &+ |D_-u_{r-1}|^2|D_0u_{r-1}|^2)h. \end{aligned} \tag{2.13}$$

Now by using Lemmas 2.1–2.2 and the discrete Sobolev inequality we obtain (provided that $\|D^2u_N\|_h < \text{const.}$):

$$X^2 \leqq C_\varepsilon(\|D_-D_0u_N\|_h^2 + \|D_0u_N\|_h^2). \tag{2.14}$$

Note: by Sobolev's inequality $\max |D_-u_r| \leqq C_1(\varepsilon)$ for some $C_1(\varepsilon) > 0$. From (2.12) and (2.14) we observe that to bound $\|D_+D_-^2(\partial u_N/\partial t)\|_h$ we need to bound

$\|D_-D_0u_N\|_h$, but since $2D_0 = D_+ + D_-$, it follows that it is sufficient to bound $\|D_+D_-u_N\|_h$ and $\|D_-^2u_N\|_h$. Let us take the inner product of equation (a) with $-D_+D_-u_N$ to obtain

$$\begin{aligned}\frac{1}{2}\frac{\partial}{\partial t}[\|D_-u_N\|_h^2 + \|D_+D_-u_N\|_h^2] &\leqq \frac{1}{3}\|D_0u_N^2\|_h\|D_+D_-u_N\|_h \\ &\quad + \frac{1}{3}\|u_ND_0u_N\|_h\|D_+D_-u_N\|_h \\ &\leqq C + \|D_-u_N\|_h^2 + \|D_+D_-u_N\|_h^2.\end{aligned}$$

By integrating the above inequality from 0 to $t < t_N$ and then using Gronwall's inequality we obtain that $\|D_+D_-u_N\|_h^2$ is bounded, because of our hypotheses on u_0. Similarly, by taking the inner product of equation (a) with $D_+^2D_-^2u_N$ we obtain

$$\begin{aligned}&\frac{1}{2}\frac{\partial}{\partial t}[\|D_-^2u_N\|_h^2 + \|D_+D_-^2u_N\|_h^2 \\ &\quad \leqq \frac{1}{3}\|D_-(u_ND_0u_N) + D_-D_0u_N^2\|_h\|D_+D_-^2u_N\|_h \\ &\quad \leqq \frac{1}{6}[C_\varepsilon\|D_-D_0u_N\|_h^2 + \|D_0u_N\|_h^2] + \|D_+D_-^2u_N\|_h^2 \\ &\quad \leqq C + \|D_-^2u_N\|_h^2 + \|D_+D_-^2u_N\|_h^2\end{aligned}$$

from which it follows that $\|D_+D_-^2u_N\|_h$ is bounded, as is $\|D_-D_0u_N\|_h$ and the lemma is proved.

Observation 3. If we take the derivative with respect to t of equation (a), we can obtain an estimate for $\|D_+D_-(\partial^2u_N/\partial t^2)\|$ independent of N in $[0, t_N)$. We will use this fact in the next section.

By the a priori estimates of Lemma 2.2, the discrete Sobolev inequality and a standard theorem of continuation of solutions of ordinary differential equations (see [3, p. 47]) it follows that the solutions of the discrete problem exist for all t in $[0, T)$ independent of N, for all $0 < T < +\infty$.

3. Regular solutions. Let us consider the solution $u_N(x_r, t)$, $r = 1, 2, \cdots, N$, of the discrete problem. By $\Phi_N(x, t)$ we represent the discrete Fourier polynomial of u_N, that is

$$\Phi_N(x, t) = \sum_{k=-n}^{n} a_N(k, t)\exp(2k\pi ix)$$

where $a_N(k, t) = (\exp(2k\pi ix_r), u_N(x_r, t))_h$ and $N = 2n + 1$ with $Nh = 1$. We are going to prove that the family $\{\Phi_N\}$ satisfies the conditions of the Arzelá–Ascoli theorem. If this is so, it follows that there exists a subsequence which we still represent by $\{\Phi_N\}$ that converges in $\Omega = [0, 1]\times[0, T]$ to a function $u(x, t)$ which is the solution claimed in Theorem 1. All this argument is based on the lemmas below.

LEMMA 3.1. *There exist constants k_1 and $k_2 > 0$ such that*

$$\left\| \frac{\partial}{\partial x} \Phi_N(\cdot, t) \right\| < k_1 \quad \text{and} \quad \left\| \frac{\partial}{\partial t} \Phi_N(\cdot, t) \right\| < k_2$$

for all t in $[0, T)$ independent of N.

Proof. We have

$$\begin{aligned} D_+ \Phi_N(x, t) &= \sum_{k=-n}^{n} a_N(k, t) D_+ \exp(2k\pi i x) \\ &= \sum_{k=-n}^{n} a_N(k, t) z \exp(2k\pi i x) \end{aligned} \tag{3.1}$$

where $hz = \exp(2k\pi i h) - 1$. Also, we have

$$\begin{aligned} a_N(k, t) z &= (z \exp(2k\pi i x_r), u_N(x_r, t))_h \\ &= (D_+ \exp(2k\pi i x_r), u_N(x_r, t))_h. \end{aligned}$$

Therefore from (3.1) it follows that

$$\begin{aligned} \|D_+ \Phi_N(\cdot, t)\|^2 &= \sum_{k=-n}^{n} |(\exp(2k\pi i x_r), D_- u_N(x_r, t))_h|^2 \\ &\leqq \|D_- u_N(\cdot, t)\|_h^2 < \text{const.} \end{aligned} \tag{3.2}$$

for all $t \geqq 0$ and all N.

In [1] we find the proof of the following result: let τ_1, τ_2 be nonnegative integers with $\tau_1 + \tau_2 = \tau$. Then, if Φ_N is a Fourier polynomial, we have

$$\left(\frac{2}{\pi}\right)^{2\tau} \left\| \frac{\partial^\tau \Phi_N(\cdot, t)}{\partial x^\tau} \right\| \leqq \|D_+^{\tau_1} D_-^{\tau_2} \Phi_N(\cdot, t)\|^2 = \|D_+^{\tau_1} D_-^{\tau_2} \Phi_N(\cdot, t)\|_h^2$$

which proves the first part of the lemma, by taking $\tau_1 = 0$, $\tau_2 = \tau = 1$ and using (3.2). Now, since

$$\frac{\partial \Phi_N}{\partial t}(x, t) = \sum_{k=-n}^{n} b_N(k, t) \exp(2k\pi i x)$$

where $b_N(k, t) = (\exp(2k\pi i x_r), (\partial u_N / \partial t)(x_r, t))_h$. It is not difficult to show that a similar argument used in Lemma 2.2 and the above discussion gives us

$$\left\| \frac{\partial^2 \Phi_N}{\partial x \, \partial t}(\cdot, t) \right\| < \text{const.} \tag{3.3}$$

independent of N for all t in $[0, T)$. This clearly implies the second part of the lemma.

It is not difficult to prove that $\Phi_N(x, t)$ is bounded by a constant for all $t \geqq 0$, independent of N. The uniform bounds for Φ_N, $\partial \Phi_N / \partial t$, and $\partial \Phi_N / \partial x$ in the L^2-norm, imply that $\Phi_N(x, t)$ satisfies the conditions of Arzelá–Ascoli (see, for example, [7, p. 186]). It follows that there exists a subsequence, which we still represent by $\{\Phi_N\}$, such that $\Phi_N(x, t) \to u(x, t)$ uniformly on Ω. Now, we claim that $\partial \Phi_N / \partial t \to \partial u / \partial t$ uniformly on Ω. In fact, we already have uniform bounds for

$\partial\Phi_N/\partial t$, $\partial^2\Phi_N/\partial t^2$ and $\partial^2\Phi_N/\partial x\,\partial t$ in the L^2-norm. By using the same argument as above, the Arzelá–Ascoli theorem gives us $\partial\Phi_N/\partial t \to v$ uniformly on Ω. We now write

$$\Phi_N(x, t) = \Phi_N(x, 0) + \int_0^t \frac{\partial\Phi_N}{\partial s}(x, s)\, ds. \tag{3.4}$$

By passing to the limit in (3.4) as $N \to +\infty$ we get

$$u(x, t) = u_0(x) + \int_0^t v(x, s)\, ds. \tag{3.5}$$

Since v is continuous on Ω, we conclude that u is differentiable with respect to t and $\partial u/\partial t = v$. A similar argument shows that $\partial\Phi_N/\partial x \to \partial u/\partial t$ uniformly on Ω and also that $\partial^3\Phi_N/\partial x^2\partial t \to \partial^3 u/\partial x^2\partial t$ uniformly on Ω. Now it is trivial that u is a periodic solution of problem (1.1), (1.2), and (1.3). The uniqueness part of Theorem 1 is standard and quite easy. In fact, let us assume that u and v are solutions of (1.1), (1.2), and (1.3). Then the function $w = u - v$ satisfies the equation

$$w_t - w_{xxt} + wu_x + vw_x = 0$$

with $w(x, 0) = 0$. Multiplying by w and integrating we get

$$\frac{\partial}{\partial t}\int_0^1 [w^2 + w_x^2]\, dx \leqq \text{const.} \int_0^1 [w^2 + w_x^2]\, dx$$

from which it follows that $w \equiv 0$; therefore $u \equiv v$.

Acknowledgment. We would like to take this opportunity to express our gratitude to the referees of this Journal for a number of valuable criticisms, which have led to the clarification of various points.

REFERENCES

[1] A. SJÖBERG, *On the Korteweg–de Vries equation: Existence and uniqueness*, J. Math. Anal. Appl., 29 (1970), pp. 569–579.

[2] B. P. NEVES, *Sur un problème non linéaire d'évolution*, C.R. Acad. Sci. Paris Série A-18, T.281 (1975), pp. 231–232.

[3] E. A. CODDINGTON AND N. LEVINSON, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.

[4] D. H. PEREGINE, *Calculations of the development of an undular bore*, J. Fluid Mech., 25 (1966), Part 2, pp. 321–330.

[5] J. L. LIONS, *Quelques Méthodes de Résolution des Problèmes aux Limites Non Linéaires*, Dunod, Gauthier-Villars, Paris, 1969.

[6] M. M. MIRANDA, *Weak solutions of a modified KdV equation*. Bol. Soc. Brasil. Mat. to appear.

[7] S. GODOUNOV, *Equations de la Physique Mathématique*, Editions MIR, Moscow, 1973.

[8] T. B. BENJAMIN, J. L. BONA AND J. J. MAHONY, *Model equations for long waves in non-linear dispersive systems*, Philos. Trans. Roy. Soc. London Ser. A, 272 (1972), pp. 47–78.

[9] T. B. BENJAMIN, *Lectures on non linear wave motion*, Non linear Wave Motion, Lectures on Applied Mathematics, vol. 15, American Mathematical Society, Providence, RI, 1974.

SIAM J. MATH. ANAL.
Vol. 8, No. 5, October 1977

LAMÉ POLYNOMIALS OF LARGE ORDER*

B. A. HARGRAVE† AND B. D. SLEEMAN‡

Abstract. Lamé's equation

$$w''+\{h-N(N+1)k^2\,\mathrm{sn}^2\,z\}w(z)=0$$

is an example of a two parameter eigenvalue problem in ordinary differential equations. Here we present new results for the value of h when Nk and Nk' assume large, real values. Uniform asymptotic expansions for Lamé polynomials are also derived. All asymptotic solutions of Lamé's equation are presented in conjunction with a realistic error bound and constitute a uniform reduction of free variables.

The asymptotic expansions are derived for values of z on the rectangle bounded by the lines $\mathrm{Im}\,(z)=0$, K', $\mathrm{Re}\,(z)=0$, K. Some of the results presented here replace known nonuniform results whilst others appear to be entirely new.

1. Introduction and notation. The standard form of Lamé's equation is taken to be

$$w''+\{h-(2n+p)(2n+p+1)k^2\,\mathrm{sn}^2\,z\}w(z)=0. \tag{1.1}$$

Both n and p will always assume integer values. With this restriction, Lamé's equation admits a solution of the form

$$\mathrm{sn}^{\rho}\,z\,\mathrm{cn}^{\sigma}\,z\,\mathrm{dn}^{\tau}\,zP_n(\mathrm{sn}^2\,z), \tag{1.2}$$

where $P_n(\mathrm{sn}^2\,z)$ is a polynomial of degree n in $\mathrm{sn}^2\,z$, whenever h assumes an eigenvalue of a particular matrix.

In (1.2) ρ, σ and τ may take either the value zero or unity subject to the following constraint:

$$\rho+\sigma+\tau=p. \tag{1.3}$$

The general symbol for a Lamé polynomial is $E^m_{2n+p}(z,k)$, this symbol being prefixed by the letters s, c, d according as ρ, σ, τ are nonzero in (1.2). If all members of the latter set of parameters are zero then the symbol is prefixed by the letter u. Note that the *modulus* k of the full notation is normally suppressed, unless emphasis is to be laid on the value of the modulus.

The advantages of the above notation are

(i) n denotes the total number of zeros of the polynomial in

$$\{z\,|\mathrm{Im}\,(z)=0,\ \mathrm{Re}\,(z)\in(0,K)\}\cup\{z\,|\mathrm{Re}\,(z)=K,\ \mathrm{Im}\,(z)\in(0,K')\},$$

K and K' being the complete elliptic integrals of the first kind, associated with the modulus k. The integer n is known as the *order* of the Lamé polynomial.

(ii) m may take the values $0, 1, \cdots, n$ and is equal to the number of zeros of the polynomial in the real interval $(0, K)$.

(iii) The *species* of a Lamé polynomial depends on the number of factors preceding P_n in (1.2) and is defined to be equal to $p+1$.

* Received by the editors July 9, 1974, and in final revised form August 23, 1976.

† Department of Mathematics, University of Aberdeen, Aberdeen, Scotland. Now at Logica Ltd., London W1, England.

‡ Department of Mathematics, University of Dundee, Dundee, Scotland.

The *eigenvalue* h of Lamé's equation, corresponding to the polynomial $E^m_{2n+p}(z, k)$, is denoted by $h^m_{2n+p}(k)$. When there is no ambiguity, the eigenvalue is written simply as h.

From the general form (1.2) it is clear that Lamé polynomials satisfy boundary conditions of the form

$$E^{m(1-\rho)}_{2n+p}(0) = E^{m(1-\sigma)}_{2n+p}(K) = E^{m(1-\tau)}_{2n+p}(K+iK') = 0, \tag{1.4}$$

where $f^{(i)}(x_0)$ denotes the ith derivative of f evaluated at x_0. The convention that $f^{(0)}(x_0) = f(x_0)$ is adopted. This notation will occur frequently in the later sections and helps to simplify the expression for the boundary conditions and the matching coefficients $E^{m(\rho)}_{2n+p}(0)$, $E^{m(\sigma)}_{2n+p}(K)$, $E^{m(\tau)}_{2n+p}(K+iK')$. These *matching coefficients* are the nonzero values of the Lamé polynomial or its derivative at the respective points. It is clear from the boundary condition (1.4) that either the Lamé polynomial or its derivative vanishes at the points 0, K, $K+iK'$. It is also apparent from the form of the differential equation (1.1) that one of these quantities must be nonzero. The matching coefficients enable asymptotic expansions obtained in one region to be related to asymptotic expansions obtained in a second region, and hence they may be continued into this second region. Consequently the matching coefficients will be of frequent occurrence in later sections.

A transformation which is extremely useful in the theory of Lamé's equation is the simultaneous application of a transformation of independent variable

$$z = K + iK' - i\xi \tag{1.5a}$$

and Jacobi's imaginary transformation for elliptic functions. We shall refer to the transformation as *Jacobi's imaginary transformation for Lamé polynomials.* This transformation leads to results of the form

$$E^m_{2n+p}(z, k) = AE^{n-m}_{2n+p}(\xi, k'), \tag{1.5b}$$

where A is a multiplicative constant. If the function on the left hand side of (1.5b) has parameters ρ, σ and τ and the function on the right has parameters ρ', σ' and τ' then $\rho' = \tau$, $\sigma' = \sigma$ and $\tau' = \rho$. Consequently the polynomials in (1.5b) are of the same species. The eight cases of the Jacobi imaginary transformation for Lamé polynomials are given in full in § 13. A consequence of the above transformation is that

$$h^m_{2n+p}(k) = (2n+p)(2n+p+1) - h^{n-m}_{2n+p}(k'). \tag{1.6}$$

Lamé polynomials as defined above are determined except for a multiplicative constant, which may be specified by adopting a *normalization convention.* Three different conventions exist, each one being convenient to display certain properties of Lamé polynomials.

The particular convention which is suitable here is that Lamé polynomials are constrained to assume the value unity if they are nonzero at the origin, or if they are zero at the origin the derivative is unity. The condition may be expressed in the concise notational form

$$E^{m(\rho)}_{2n+p}(0) = 1. \tag{1.7}$$

The reader is referred to [1] for a complete exposition of the theory of Lamé's equation.

In §§ 6–11 the independent variable z in (1.1) will be finite assuming values on a rectangle. This rectangle bounded by the lines $\mathrm{Im}\,(z)=0, K'$, $\mathrm{Re}\,(z)=0, K$ will be called the *fundamental rectangle.* In the development we shall attempt to obtain asymptotic expansions for Lamé polynomials for all possible values of the argument on the fundamental rectangle. The sides of this rectangle will be denoted by $[0, K], [K, K+iK'], [iK', K+iK']$ and $[0, iK']$. This notation is used to avoid the necessity of writing both the real and imaginary parts of the interval. A notation similar to the above for closed intervals will be used for both open and semi-open intervals.

The choice of large parameter is conveniently taken to be χ, where

$$\chi = 2n+p+\tfrac{1}{2}. \tag{1.8}$$

The reason for this choice will be apparent in §§ 8 and 9. In all the following sections it is assumed that χk and $\chi k'$ are large.

2. Descriptive treatment of the problem. The material presented in the following sections constitutes a comprehensive treatment of the behavior of Lamé polynomials of large order. Present knowledge of the asymptotic behavior of such polynomials is restricted to the following cases (i), (ii) and (iii) below.

(i) If n is large and $m = O(1)$, Ince [7] derives an asymptotic series for the eigenvalue h, namely

$$h^m_{2n+p}(k) \sim (4m+2\rho+1)k\tilde{\chi} + O(1), \tag{2.1}$$

where

$$\tilde{\chi}^2 = (2n+p)(2n+p+1). \tag{2.2}$$

Hence in the variable $\tilde{\chi}$, the leading term of the coefficient of $w(z)$ in Lamé's equation has a double zero at the origin. Ince used the Liouville–Green approximation to obtain asymptotic expansions for Lamé polynomials, which are valid in the interval $(0, K]\cup[K, K+iK']$ provided $\tilde{\chi}\,\mathrm{sn}\,z$ is large.

(ii) If $n-m$ is $O(1)$, one may apply the Jacobi imaginary transformation for Lamé polynomials to derive corresponding results to those in (i). In this case Lamé's equation has a double turning point at $z=K+iK'$ and the Liouville–Green approximation is valid in $[0, K]\cup[K, K+iK')$ provided $\tilde{\chi}\,\mathrm{dn}\,z$ is large.

(iii) An interesting exceptional case arises for Lamé polynomials, which are identical except for a multiplicative factor on the intervals $[0, K]$ and $[K, K+iK']$. Such polynomials satisfy identical boundary conditions at $z=0$ and $z=K+iK'$. There is however no restriction on the boundary condition at $z=K$. Consequently this behavior may only occur for four types of Lamé polynomial. Furthermore $m=\frac{1}{2}n$ and $k^2=\frac{1}{2}$. For such polynomials

$$h^{n/2}_{2n+p}(1/\sqrt{2}) = \tfrac{1}{2}(2n+p)(2n+p+1). \tag{2.3}$$

This result was observed by Erdélyi [4]. However, asymptotic approximations for Lamé polynomials have not been obtained for this case.

In the light of the above knowledge, the following questions naturally arise.

(i) How does the turning point move as m increases, whilst k and n are fixed?

(ii) Given k, what is the ratio of m to n such that $z = K$ is a turning point of the differential equation (1.1)?

Obviously the location of the turning point at $z = K$ leads to a critical value of the ratio $m : n$. One may define m_c to be the value of m such that the turning point is located at $z = K$. There are five distinct configurations of turning points for Lamé polynomials, which may be conveniently illustrated by Figs. 1 to 5.

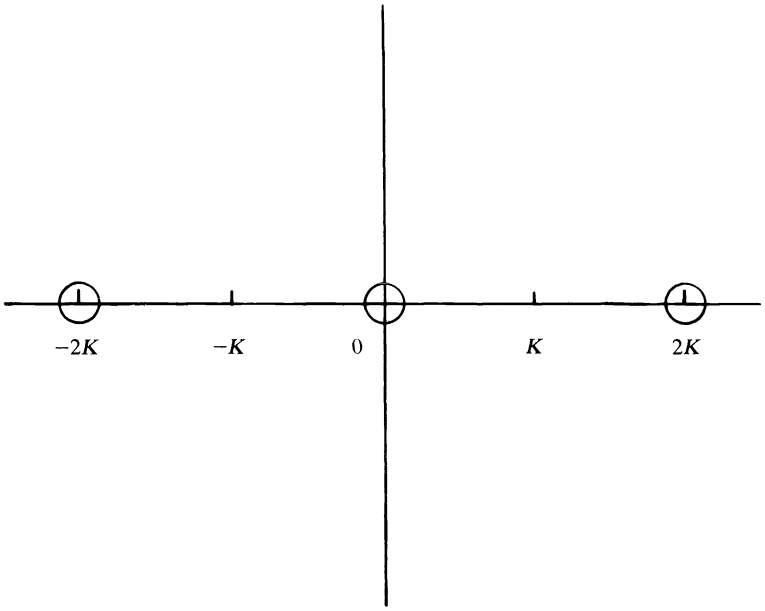


FIG. 1

In the five figures a dot enclosed by a circle denotes the location of a double turning point, whilst the center of a cross denotes the location of a simple turning point.

Figures 2 and 4 demonstrate the cases which occur most frequently, whilst Figs. 1, 3 and 5 portray cases of confluence, in which two simple turning points coalesce. Note that any line Im $(z) = r_1 K'$, or any line Re $(z) = r_2 K$ (r_1, r_2 integers) are lines of symmetry in all cases.

When m is $O(1)$, a confluent case (Fig. 1) applies. As m increases the two simple turning points separate and the turning points are located as in Fig. 2. As m approaches the critical value m_c, the simple turning points coalesce at $z = \pm K, \pm 3K, \pm 5K, \cdots$, as in Fig. 3. As m increases still further the turning points separate along the lines Re $(z) = \pm K, \pm 3K, \pm 5K, \cdots$, as in Fig. 4. Finally as m

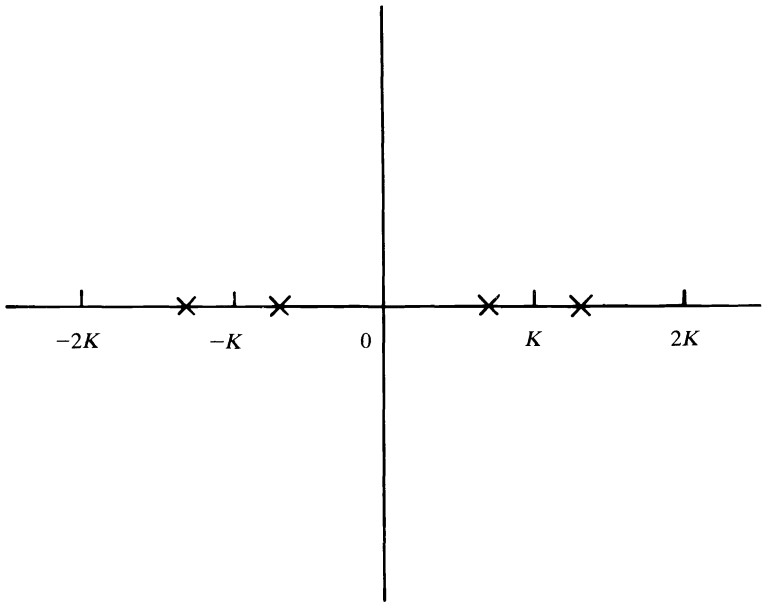


FIG. 2

approaches the value n the simple turning poles coalesce as in Fig. 5. In all cases double poles exist at the points $z \equiv iK' \bmod \{2K, 2iK'\}$.

The behavior of the solutions of a differential equation on opposite sides of turning points is often completely different. For simple turning points, the solution

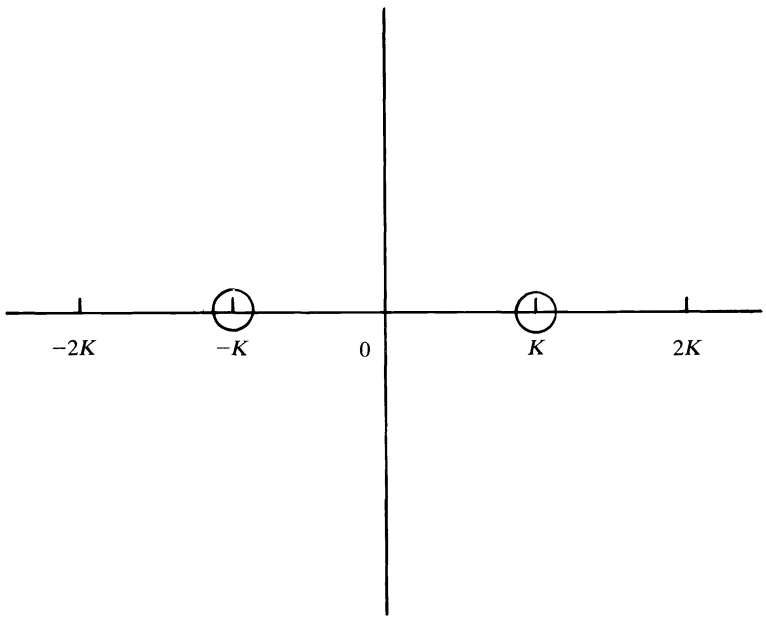


FIG. 3

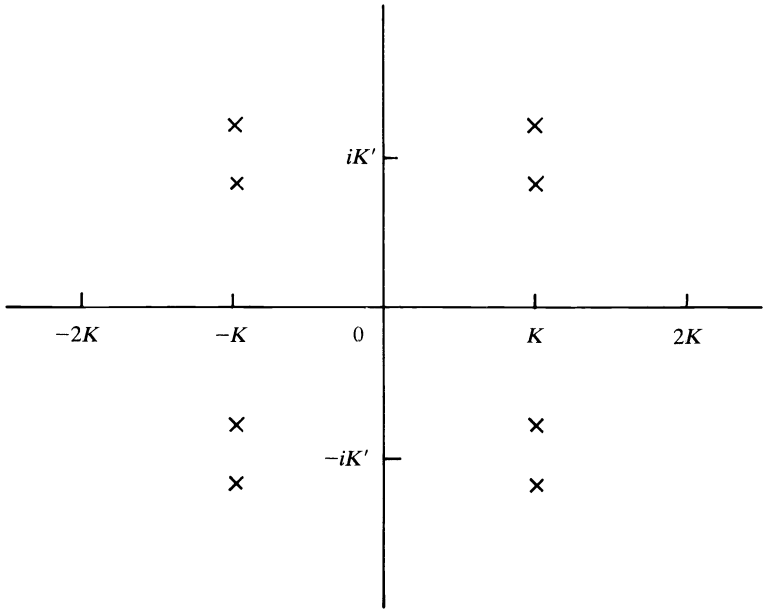


FIG. 4

on one side is oscillatory whilst on the other side exponential decay or growth occurs. The reason for this behavior is clearly that the solutions in this case may be expressed in terms of Airy functions which exhibit this kind of behavior. The sign of the coefficient of $w(z)$ in Lamé's equation determines the qualitative behavior

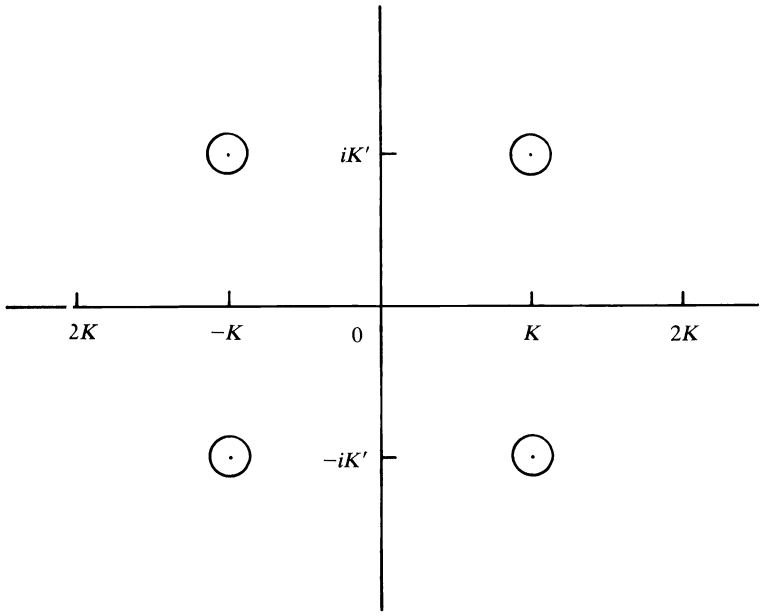


FIG. 5

of Lamé polynomials. For the case described by Fig. 2, in $(0, K)$ to the left of the turning point the coefficient of $w(z)$ is positive so that the solutions are oscillatory whilst to the right of the turning point, exponential behavior occurs. On the line $\mathrm{Re}\,(z) = K$, the solutions are oscillatory. Thus for the case of Fig. 1, the solutions on the real line will behave exponentially away from the origin whilst for the case of Fig. 3 the solutions on both the real line and the line $\mathrm{Re}\,(z) = K$ will be oscillatory. When the turning point is located on the line $\mathrm{Re}\,(z) = K$, the real line will be an interval possessing oscillatory solutions, whilst the part of the line $\mathrm{Re}\,(z) = K$ between $z = K$ and the turning point will have solutions which behave exponentially. Between the turning point and $K + iK'$ the solutions will again be oscillatory.

On account of the Jacobi imaginary transformation for Lamé polynomials, it is always possible to associate a Lamé polynomial with a turning point on the line $\mathrm{Im}\,(z) = K$ with a Lamé polynomial which has a turning point on the real line. Thus, when one calculates asymptotic expansions for Lamé polynomials it is only necessary to discuss the case of turning points on the real line.

In §§ 3 and 4 eigenvalues of Lamé's equation are discussed. In the former section certain integrals which occur in connection with both eigenvalues and Lamé polynomials are also investigated.

In § 5 basic properties of the solutions of the approximating equation are introduced. These solutions occur frequently throughout §§ 6, 7, 8 and 9, in which asymptotic solutions of Lamé's equations are calculated on the four sides of the fundamental rectangle. In §§ 10 and 11 the properties of the approximating functions are used to enable the constant multipliers, which characterize Lamé polynomials, to be determined.

In § 12, the range of validity of the results of §§ 10 and 11 is extended beyond the fundamental rectangle. Further reference is made to the Jacobi imaginary transformation for Lamé polynomials in order that Lamé polynomials with turning points on $\mathrm{Re}\,(z) = K$ may be identified. Some special cases are discussed in § 13 and conclusions are presented in § 14.

3. The eigenvalue condition for *h*. For the purpose of obtaining eigenvalues of Lamé's equation it is often advantageous to reformulate the three point boundary value problem in the complex plane, (1.1) and (1.4), as a two parameter problem, whose independent variables assume real values. The reformulated problem is

$$w_1''(x) + \{h - \chi^2 k^2 \operatorname{sn}^2 x + \tfrac{1}{4}k^2 \operatorname{sn}^2 x\} w_1(x) = 0, \tag{3.1a}$$

$$w_1^{(1-\rho)}(0) = w_1^{(1-\sigma)}(K) = 0,$$

$$w_2''(y) + \{h' - \chi^2 k'^2 \operatorname{sn}^2 y + \tfrac{1}{4}k'^2 \operatorname{sn}^2 y\} w_2(y) = 0, \tag{3.1b}$$

$$w_2^{(1-\tau)}(0) = w_2^{(1-\sigma)}(K') = 0,$$

where all the elliptic functions occurring in (3.1b) have modulus k', and $h' = \chi^2 - (\frac{1}{4} + h)$. Throughout this section elliptic functions with argument y have modulus k'. The former equation is Lamé's equation on the real line with appropriate boundary conditions, whilst the latter is obtained by applying the Jacobi imaginary transformation to Lamé's equation on the line $\mathrm{Re}\,(z) = K$.

Classical Sturm–Liouville theory [2, Chap. 10] may be applied to equations (3.1a, b) in order to deduce that

$$h^i_{2n+p} \geqq 0, \quad h'^i_{2n+p} \geqq 0, \quad h^i_{2n+p} < h^{i+1}_{2n+p}, \quad \forall i.$$

Hence $0 \leqq h^m_{2n+p} \leqq \chi^2 - \frac{1}{4}$ for all integers m, which implies that all the turning points of Lamé's equation lie on the lines $\mathrm{Im}\,(z) \equiv 0 \bmod \{2K'\}$ or $\mathrm{Re}\,(z) \equiv K \bmod \{2K\}$. Thus two of the intervals, namely $[0, iK']$ and $[iK', K+iK']$, on which asymptotic expansions for Lamé polynomials are to be constructed, are in general free from turning points. The exceptional case occurs when the turning point on the real line is located at the origin or the turning point on the line $\mathrm{Re}\,(z) = K$ is located at $z = K + iK'$.

In order to answer the first question posed in § 2, it is necessary to consider the coefficient of $w(z)$ in Lamé's equation. This is $(h - \chi^2 k^2 \,\mathrm{sn}^2 z + \frac{1}{4} k^2 \,\mathrm{sn}^2 z)$. The dominant terms of the coefficient are, except in the neighborhood of turning points, the first two. These may be conveniently rewritten as

$$h - \chi^2 k^2 \,\mathrm{sn}^2 z = \chi^2 k^2 (\alpha^2 - \mathrm{sn}^2 z), \tag{3.2}$$

for some real α.

Consider the union I of the two intervals $[0, K]$ and $[K, K+iK']$. The function $\mathrm{sn}\, z$ is monotonic increasing as z moves from the origin to $K+iK'$ remaining in the set I. Hence there exists at most one $z_0 \in I$ such that

$$\alpha = \mathrm{sn}\, z_0. \tag{3.3}$$

The upper and lower bounds for h given above imply that

$$0 < \alpha < 1/k. \tag{3.4}$$

Condition (3.4) is sufficient to ensure that a z_0 satisfying (3.3) exists. Consequently Lamé's equation possesses a unique turning point in I. Provided $|\alpha - 1|$ is sufficiently large, it is possible to use a direct method for the calculation of eigenvalues, whilst for small values of $|\alpha - 1|$ it is preferable to calculate solutions of the differential equation and deduce the eigenvalues from these solutions. The range of orders of magnitude for $|\alpha - 1|$, which enables eigenvalues to be calculated directly, will be obtained later in this section. All results for the case $1 < \alpha < 1/k$ may be deduced by similar methods to those employed for the case $0 \leqq \alpha < 1$, so that we may concentrate on the latter case giving only the results for the former case. The aim of the discussion is the establishment of a condition α such that (3.1a, b) admit solutions which are Lamé polynomials.

Let us suppose that $0 \leqq \alpha < 1$, then there exists a $z_0 \in [0, K)$ such that

$$\alpha = \mathrm{sn}\, z_0.$$

As this value of z_0 is unique in I, it is clear that for all $z \in [K, K+iK']$,

$$\alpha^2 \neq \mathrm{sn}^2 z,$$

since $\mathrm{sn}\, z$ is nonnegative for $z \in I$. This is precisely the condition for the Liouville–Green approximation to lead to uniformly valid solutions of Lamé's equation in $[K, K+iK']$. This is equivalent to saying that it is possible to obtain an asymptotic

series for the eigenvalues, from the modified Prüfer angular equation corresponding to (3.1b). The modified Prüfer substitution is defined by [2, p. 267]

$$w_2(y) = R(y)(h' - \chi^2 k'^2 \operatorname{sn}^2 y + \tfrac{1}{4}k'^2 \operatorname{sn}^2 y)^{-1/4} \cos \phi(y),$$
$$w_2'(y) = R(y)(h' - \chi^2 k'^2 \operatorname{sn}^2 y + \tfrac{1}{4}k'^2 \operatorname{sn}^2 y)^{1/4} \sin \phi(y),$$

and the resulting angular equation is

$$\phi' = -(h' - \chi^2 k'^2 \operatorname{sn}^2 y + \tfrac{1}{4}k'^2 \operatorname{sn}^2 y)^{1/2} - \frac{(\chi^2 - \frac{1}{4})k'^2 \operatorname{sn} y \operatorname{cn} y \operatorname{dn} y}{2(h' - (\chi^2 - \frac{1}{4})k'^2 \operatorname{sn}^2 y)} \sin 2\phi(y). \tag{3.5}$$

The boundary conditions for this equation are

$$\phi(0) = \tfrac{1}{2}\tau\pi, \qquad \phi(K') = -(n-m)\pi - \tfrac{1}{2}\sigma\pi.$$

These conditions ensure that the solution has $(n-m)$ zeros in the open interval $(0, K')$. Integration of (3.5) leads to

$$[\phi]_0^{K'} = -\chi \int_0^{K'} \left(\operatorname{dn}^2 y - \alpha^2 k^2 - \frac{1}{4\chi^2} \operatorname{dn}^2 y\right)^{1/2} dy - \frac{1}{2}\int_0^{K'} \frac{(1 - \frac{1}{4}\chi^{-2})k'^2 \operatorname{sn} y \operatorname{cn} y \operatorname{dn} y}{(\operatorname{dn}^2 y - \alpha^2 k^2 - (1/4\chi^2)k'^2 \operatorname{sn}^2 y)} \sin 2\phi(y)\, dy. \tag{3.6}$$

On account of the restriction on α, we may infer that ϕ' has no stationary points and that

$$k'^2 \operatorname{sn} y \operatorname{cn} y \operatorname{dn} y \Big/ \left(\operatorname{dn}^2 y - \alpha^2 k^2 - \frac{1}{4\chi^2} \operatorname{dn}^2 y\right)$$

is finite throughout the region of integration. The latter integral will thus have an integrand of $O(1)$ provided α is bounded away from unity as $\chi \to \infty$, and the dominant contribution to the latter integral is $O(1/\chi)$ as the integrand vanishes at the end points [12, p. 97]. The former integrand may be expanded as a binomial series and integrated to produce the result

$$[(n-m) + \tfrac{1}{2}(\sigma+\tau)]\pi = \chi \int_0^{K'} (\operatorname{dn}^2 y - \alpha^2 k^2)^{1/2}\, dy + O\left(\frac{1}{\chi}\right). \tag{3.7}$$

The behavior of this integral is of particular interest as α approaches unity from below. Suppose that

$$1 - \alpha = O(\chi^{-\beta}), \qquad \beta > 0.$$

The neglected term in the first integral in (3.6) is $O(1/\chi^{1-(\beta/2)})$, whilst the second integral is $O(\chi^{2\beta-1})$ [12, p. 98]. Thus (3.7) is valid with error term $o(1)$ provided $\beta < \frac{1}{2}$.

In terms of the original variable z, where

$$z = K + iK' - iy \quad \text{and} \quad \operatorname{sn}(z, k) = k^{-1} \operatorname{dn}(y, k'),$$

$$[n - m + \tfrac{1}{2}(\sigma+\tau)]\pi = \chi k \int_K^{K+iK'} (\operatorname{sn}^2 z - \alpha^2)^{1/2}\, dz + o(1). \tag{3.8}$$

provided $1 - \alpha = O(\chi^{-\beta})$, $\beta \in [0, \frac{1}{2})$.

The integral of $k(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}$ will prove to be of critical importance in later sections. If the function $k(\alpha^2-\mathrm{sn}^2\,\xi)^{1/2}$ is integrated around the contour γ of Fig. 6, one may apply Cauchy's theorem to deduce that

$$k\int_\gamma (\alpha^2-\mathrm{sn}^2\,\xi)^{1/2}\,d\xi=0.$$

The contributions to this integral are real if

$$\text{(i)}\quad 0\leqq\xi<z_0,$$
$$\text{(ii)}\quad K\leqq\xi\leqq K+iK',$$
$$\text{(iii)}\quad |\xi-iK'|=\varepsilon,$$

and pure imaginary otherwise. Thus

$$k\int_0^{z_0}(\alpha^2-\mathrm{sn}^2\,\xi)^{1/2}\,d\xi+k\int_K^{K+iK'}(\alpha^2-\mathrm{sn}^2\,\xi)^{1/2}\,d\xi+k\int_{iK+\varepsilon}^{iK'-i\varepsilon}(\alpha^2-\mathrm{sn}^2\,\xi)^{1/2}\,d\xi=0 \tag{3.9}$$

and

$$k\int_{z_0}^{K}(\alpha^2-\mathrm{sn}^2\,\xi)^{1/2}\,d\xi+k\int_{K+iK'}^{iK'+\varepsilon}(\alpha^2-\mathrm{sn}^2\,\xi)^{1/2}\,d\xi+k\int_{iK'-i\varepsilon}^{0}(\alpha^2-\mathrm{sn}^2\,\xi)^{1/2}\,d\xi=0.$$

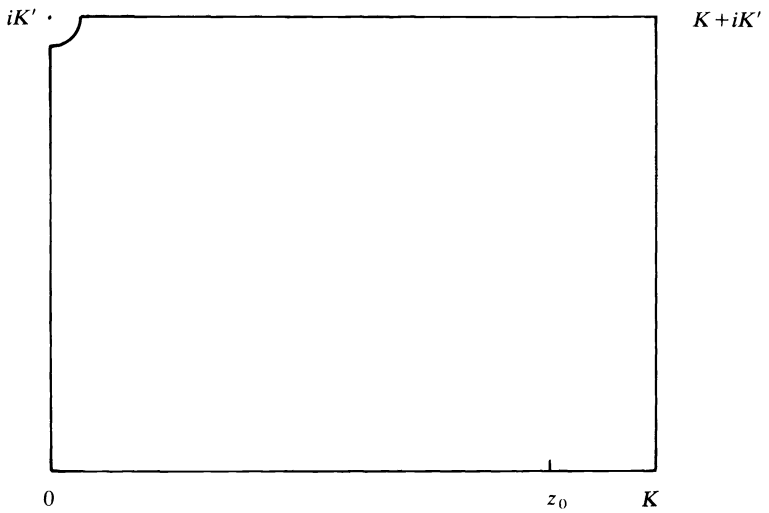


FIG. 6

This latter equation may be rewritten in terms of real valued integrals as

$$k\int_{z_0}^{K}(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi+k\int_{K+iK'}^{iK'+\varepsilon}(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi$$
$$=k\int_{0}^{iK-i\varepsilon}(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi. \tag{3.10}$$

In the above integrals, the branch of $(\alpha^2-\mathrm{sn}^2\,\xi)^{1/2}$ in the regions on which $(\alpha^2-\mathrm{sn}^2\,\xi)$ is negative is the one which is consistent with the limiting case $\alpha=1$, i.e. $(\alpha^2-\mathrm{sn}^2\,\xi)^{1/2}=\mathrm{cn}\,\xi$, when $\alpha=1$.

Consequently $\arg(\alpha^2-\mathrm{sn}^2\,\xi)^{1/2}=\frac{1}{2}\pi$, in the latter two integrands in (3.9). On account of the pole at $\xi=iK'$, (3.9) simplifies to

$$k\int_{0}^{z_0}(\alpha^2-\mathrm{sn}^2\,\xi)^{1/2}\,d\xi+k\int_{K}^{K+iK'}(\alpha^2-\mathrm{sn}^2\,\xi)^{1/2}\,d\xi=\frac{\pi}{2}. \tag{3.11}$$

The results (3.8) and (3.11) imply that

$$\chi k\int_{0}^{z_0}(\alpha^2-\mathrm{sn}^2\,\xi)^{1/2}\,d\xi=\tfrac{1}{2}\chi\pi-[n-m+\tfrac{1}{2}(\sigma+\tau)]\pi+o(1),$$

i.e.,

$$\chi k\int_{0}^{z_0}(\alpha^2-\mathrm{sn}^2\,\xi)^{1/2}\,d\xi=m\pi+\tfrac{1}{2}\rho\pi+\tfrac{1}{4}\pi+o(1), \tag{3.12}$$

provided $\beta\in[0,\frac{1}{2})$. Thus if the turning point is located on the real line, for fixed n, k, (3.12) describes the movement of the turning point for different choices of m, until m is such that the turning point is located in a neighborhood, of radius $O(\chi^{-1/2})$, of the point $z=K$. The relation (3.12) is the condition to be satisfied by the parameter h, related to α by (3.2), if Lamé's equation is to admit a polynomial solution satisfying the relevant boundary conditions and possessing m zeros in $(0, K)$ and $n-m$ zeros in $(K, K+iK')$ with the above restriction on the location of the turning point. The integral on the left hand side of (3.12) is a monotonic increasing function of α. Thus for a given Lamé polynomial n, m and k are known, and the parameter α is determined uniquely by the condition (3.12).

Observe also that as $\alpha\to 1$,

$$\int_{z_0}^{K}(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi=O(\chi^{-2\beta}),\quad \text{if } \beta>0. \tag{3.13}$$

This result may be seen immediately by replacing $\mathrm{sn}\,\xi$ by its Taylor series expansion in the neighborhood of $\xi=K$.

The results corresponding to (3.8), (3.10) and (3.12) for the case $1<\alpha<k^{-1}$ are

$$\chi k\int_{0}^{K}(\alpha^2-\mathrm{sn}^2\,\xi)^{1/2}\,d\xi=m\pi+\tfrac{1}{2}(\rho+\sigma)\pi+o(1), \tag{3.14a}$$

$$\chi k\int_{z_0}^{K+iK'}(\alpha^2-\mathrm{sn}^2\,\xi)^{1/2}\,d\xi=(n-m)\pi+\tfrac{1}{2}\tau\pi+\tfrac{1}{4}\pi+o(1) \tag{3.14b}$$

and

$$k\int_K^{z_0}(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi+k\int_{K+iK'}^{iK'+\varepsilon}(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi$$
$$=k\int_0^{iK-i\varepsilon}(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi. \tag{3.15}$$

4. The critical ratio of m to n for fixed values of n and k. We have established conditions in the previous section for the determination of eigenvalues when the turning point is either located on the real line or located on the line $\mathrm{Re}\,(z)=K$. Both conditions are subject to the restriction that the turning point must not be too close to the point $z=K$. If the turning point is located at $z=K$, we know a priori that h has the asymptotic behavior

$$h\sim\chi^2k^2.$$

Consequently the leading term of the asymptotic series for h is known once a condition constraining the turning point to be located at $z=K$ has been established. This condition will take the form of a critical value of the ratio $m:n$. Let this critical value be denoted by C. C will of course be a function of the modulus k.

The critical case arises when

$$h-\chi^2k^2\,\mathrm{sn}^2\,K=0,\quad \text{i.e. } \alpha=1.$$

For a given Lamé polynomial m, n and k are known. Here we wish to keep n and k fixed and allow m to assume various integer values between zero and n.

Consider $I(\alpha)=\int_0^{z_0}(\alpha^2-sn^2\,\xi)^{1/2}\,d\xi$, $\alpha\in[0,1]$, where as before $\mathrm{sn}\,z_0=\alpha$. As a consequence of the restriction on α, $z_0\in[0,K]$. Observe that $I'(\alpha)$ is positive so that I is a monotonic increasing function of α. Observe also that

$$I(0)=0,\qquad I(1)=\int_0^K\mathrm{cn}\,\xi\,d\xi=\frac{1}{k}\sin^{-1}k.$$

Thus for $0\leqq\alpha<1$, i.e. the case of a turning point in the interval $[0,K)$,

$$0\leqq\frac{C\pi}{2k}<\frac{1}{k}\sin^{-1}k. \tag{4.1}$$

This follows from (3.12). For $\alpha\in[1,1/k]$, define

$$J(\alpha)=\int_{z_0}^{K+iK'}(\alpha^2-\mathrm{sn}^2\,\xi)^{1/2}\,d\xi,$$

where $\mathrm{sn}\,z_0=\alpha$, and $\arg(\alpha^2-\mathrm{sn}^2\,\xi)^{1/2}=-\frac{1}{2}\pi$ in $[z_0,K+iK']$. The restriction on α implies that $z_0\in[K,K+iK']$. Now $J'(\alpha)<0$, so that J is a monotonic decreasing function of α. Also,

$$J(1)=\frac{1}{k}\sin^{-1}k',\qquad J\left(\frac{1}{k}\right)=0.$$

Thus for a turning point in $(K, K+iK']$,

$$0 \leqq \frac{(1-C)}{2k}\pi < \frac{1}{k}\sin^{-1} k'. \tag{4.2}$$

This follows from (3.14b). Now if α is equal to unity, it is not possible for C to satisfy either (4.1) or (4.2). If C were to lie in either interval the turning point would no longer be at $z = K$ and α would not be unity. Also C is restricted by the fact that $0 \leqq m \leqq n$ so that $0 \leqq C \leqq 1$. Thus the only possible value for C for the case $\alpha = 1$ is the limiting value of (4.1) and (4.2), i.e.,

$$C = \frac{2}{\pi}\sin^{-1} k. \tag{4.3}$$

This predicted value for C agrees not only with the known results mentioned in § 2, but also with numerically computed results [5] for eigenvalues of Lamé polynomials of large order.

5. Solutions of the approximating equations. These solutions, which are constructed in the following sections, are exponential functions, trigonometric functions or parabolic cylinder functions. As different forms for standard solutions of the parabolic cylinder equations exist, we take this opportunity to specify the functions that will be used in the later sections.

In the notation of Miller [9], [10], the standard form of the parabolic cylinder equation is

$$w'' = (a + \tfrac{1}{4}x^2)w. \tag{5.1}$$

This equation has solutions $U(a, \pm x)$ and $\bar{U}(a, \pm x)$, both of which may be defined in terms of the confluent hypergeometric function ${}_1F_1$ by

$$\begin{aligned} U(a, \pm x) = {} & \pi^{1/2}2^{-(2a+1)/4}\, e^{-x^2/4}\, {}_1F_1(\tfrac{1}{2}a + \tfrac{1}{4}; \tfrac{1}{2}; \tfrac{1}{2}x^2)/\Gamma(\tfrac{3}{4} + \tfrac{1}{2}a) \\ & \mp \pi^{1/2}2^{-(2a-1)/4}\, e^{-x^2/4} x\, {}_1F_1(\tfrac{1}{2}a + \tfrac{3}{4}; \tfrac{1}{2}; \tfrac{1}{2}x^2)/\Gamma(\tfrac{1}{4} + \tfrac{1}{2}a), \end{aligned} \tag{5.2a}$$

$$\begin{aligned} \bar{U}(a, \pm x) = {} & \pi^{-1/2}2^{-(2a+1)/4}\Gamma(\tfrac{1}{4} - \tfrac{1}{2}a)\sin(\tfrac{3}{4}\pi - \tfrac{1}{2}a\pi)\, e^{-x^2/4}\, {}_1F_1(\tfrac{1}{2}a + \tfrac{1}{4}; \tfrac{1}{2}; \tfrac{1}{2}x^2) \\ & \mp \pi^{-\frac{1}{2}}2^{-(2a-1)/4}\Gamma(\tfrac{3}{4} - \tfrac{1}{2}a)\sin(\tfrac{5}{4}\pi - \tfrac{1}{2}a\pi)\, e^{-x^2/4} x\, {}_1F_1(\tfrac{1}{2}a + \tfrac{3}{4}; \tfrac{1}{2}; \tfrac{1}{2}x^2). \end{aligned} \tag{5.2b}$$

When the sign of the $\frac{1}{4}x^2$ term in (5.1) is negative new standard solutions must be defined. The equation is

$$w'' = (a - \tfrac{1}{4}x^2)w \tag{5.3}$$

and the standard solutions are $W(a, \pm x)$, defined by

$$\begin{aligned} W(a, \pm x) = {} & 2^{-3/4}|\Gamma(\tfrac{1}{4} + \tfrac{1}{2}ia)/\Gamma(\tfrac{3}{4} + \tfrac{1}{2}ia)|^{1/2}\, e^{ix^2/4}\, {}_1F_1(\tfrac{1}{2}ia + \tfrac{1}{4}; \tfrac{1}{2}; -\tfrac{1}{2}ix^2) \\ & \mp 2^{-1/4}|\Gamma(\tfrac{3}{4} + \tfrac{1}{2}ia)/\Gamma(\tfrac{1}{4} + \tfrac{1}{2}ia)|^{1/2}\, e^{ix^2/4} x\, {}_1F_1(\tfrac{1}{2}ia + \tfrac{3}{4}; \tfrac{3}{2}; -\tfrac{1}{2}ix^2). \end{aligned} \tag{5.4}$$

In (5.2) and (5.4) the usual convention, i.e. that the upper or lower sign should be used throughout, is adopted. The reader is referred to Olver [11], [13] for the properties of the functions defined by (5.2) and (5.4).

In later sections auxiliary functions, related to parabolic cylinder functions, occur in the error terms. These functions are defined in [13]. In [13], Olver uses the general symbols E, M and N to denote these auxiliary functions for both (5.1) and (5.3). We modify this notation giving each function a suffix, unity for functions associated with (5.1) and two for the functions associated with (5.3). A further notational change from [13] is that the function

$$\{1+\exp(2\pi a)\}^{1/2}-\exp(\pi a)$$

is denoted by $l(a)$.

6. Solutions of Lamé's equation on the real line. The solutions, which are investigated here, correspond to the case $0\leqq\alpha\leqq 1$. Thus Lamé's equation possesses a turning point, which we shall assume to be located at $z=z_0$, in the interval $[0, K]$. Lamé's equation also possesses turning points at $z=-z_0$ and $z=2K-z_0$, so that if either $z_0\to 0$ or $z_0\to K$, two turning points are coalescing. For $z_0\in(0, K)$, examination of Lamé's equation shows that the turning point is simple. The behavior of Lamé polynomials in the two cases $z_0\to 0$ and $z_0\to K$ is unfortunately diverse as the comparison equation is different for each of these cases. Thus it is necessary to consider solutions in each of these cases separately. The theory [13] that is used is however applicable to simple turning points, which may coalesce to form a double turning point. It is therefore possible to cover the case $0\leqq\alpha<1$ with one application of the theory and the case $0<\alpha\leqq 1$ with a second application of the theory. This would result in a duplication of the results for $\alpha\in(0, 1)$, so we restrict attention to the following cases:

(a) $0\leqq\alpha\leqq\frac{1}{2}$,

(b) $\frac{1}{2}\leqq\alpha\leqq 1$.

In §§ 6 to 9 the two cases will be treated separately.

Case (a). In order to use the theory of [13], preliminary transformations of both dependent and independent variable are required to reduce Lamé's equation to standard form. Lamé's equation may be written as

$$w''(z)=\{\chi^2k^2\,\mathrm{sn}^2 z-h-\tfrac{1}{4}k^2\,\mathrm{sn}^2 z\}w(z), \tag{6.1a}$$

i.e.

$$w''(z)=[\chi^2k^2\{\mathrm{sn}^2 z-\alpha^2\}-\tfrac{1}{4}k^2\,\mathrm{sn}^2 z]w(z), \tag{6.1b}$$

where χ is defined by (1.8).

A new dependent variable is introduced by

$$W(\zeta_1)=\left(\frac{dz}{d\zeta_1}\right)^{-1/2}w(z), \tag{6.2a}$$

whilst the new independent variable ζ_1 is given by

$$\left(\frac{dz}{d\zeta_1}\right)^2k^2(\mathrm{sn}^2 z-\alpha^2)=\zeta_1^2-\nu_1^2, \tag{6.2b}$$

where ν_1 is not yet specified.

This choice of transformation is made since the leading term on the right hand side of (6.1b) is negative between the turning points located at $\pm z_0$. On account of the symmetry of the problem about the origin, it is only necessary to consider the behavior of W and ζ_1 for positive ζ_1. This enables us to simplify a little the method of [13].

The parameter ν_1 is at present arbitrary. It may be specified by stipulating that $\zeta_1(0)=0$ and $\zeta_1(z_0)=\nu_1$, ζ_1 being considered as a function of z. Thus on rearranging (6.2b) and integrating

$$k\int_0^{z_0}(\alpha^2\,\mathrm{sn}^2\,t)^{1/2}\,dt=\int_0^{\nu_1}(\nu_1^2-\tau^2)^{1/2}\,d\tau,$$

or

$$\nu_1^2=\frac{4k}{\pi}\int_0^{z_0}(\alpha^2-\mathrm{sn}^2\,t)^{1/2}\,dt. \tag{6.3}$$

The dependence of ζ_1 upon z may be expressed by means of the following two integral relations:

$$k\int_0^{z}(\alpha^2-\mathrm{sn}^2\,t)^{1/2}\,dt=\int_0^{\zeta_1}(\nu_1^2-\tau^2)^{1/2}\,d\tau \quad \text{for } z\leqq z_0, \tag{6.4a}$$

and

$$k\int_{z_0}^{z}(\mathrm{sn}^2\,t-\alpha^2)^{1/2}\,dt=\int_{\nu_1}^{\zeta_1}(\tau^2-\nu_1^2)^{1/2}\,d\tau \quad \text{for } z\geqq z_0. \tag{6.4b}$$

Observe that relations (6.4a, b) imply that ζ_1 is a continuous, increasing function of z in $[0, K]$. The transformations (6.2a, b) lead to the comparison equation

$$\frac{d^2W}{d\zeta_1^2}=\{\chi^2(\zeta_1^2-\nu_1^2)+\psi_1(\chi,\nu_a,\zeta_1)\}W(\zeta_1), \tag{6.5}$$

where

$$\psi_1(\chi,\nu_1,\zeta_1)=-\frac{1}{4}\left(\frac{dz}{d\zeta_1}\right)^2k^2\,\mathrm{sn}^2\,z+\left(\frac{dz}{d\zeta_1}\right)^{1/2}\frac{d^2}{d\zeta_1^2}\left(\frac{dz}{d\zeta_1}\right)^{-1/2}. \tag{6.6}$$

In order to construct error bounds a positive valued function on $(-\infty,\infty)$, which is $O(x)$ as $x\to\pm\infty$, is required. A suitable function is thus

$$\Omega_1(x)=\Gamma(\tfrac{1}{2}+\tfrac{1}{2}\chi\nu_1^2)/M_1^2(-\tfrac{1}{2}\chi\nu_1^2,|x|).$$

The error control function may now be defined as

$$F_1(\chi,\nu_1,\zeta_1)=\int\frac{\psi_1(\chi,\nu_1,\zeta_1)}{\Omega_1(\zeta_1\sqrt{2\chi})}\,d\zeta_1$$

with associated variational operator [12, p. 27]

$$\mathscr{V}_{a,b}(F_1)=\int_a^b\frac{|\psi_1(\chi,\nu_1,\zeta_1)|}{\Omega_1(\zeta_1\sqrt{2\chi})}\,d\zeta_1.$$

The main theorem of [13] may now be applied to (6.5). This theorem states that two continuous, twice differentiable, independent solutions of (6.5) are given by

$$W_{11}(\chi, \nu_1, \zeta_1) = U(-\tfrac{1}{2}\chi\nu_1^2, \zeta_1\sqrt{2\chi}) + \varepsilon_{11}(\chi, \nu_1, \zeta_1), \tag{6.7a}$$

$$W_{12}(\chi, \nu_1, \zeta_1) = \bar{U}(-\tfrac{1}{2}\chi\nu_1^2, \zeta_1\sqrt{2\chi}) + \varepsilon_{12}(\chi, \nu_1, \zeta_1), \tag{6.7b}$$

where

$$\frac{|\varepsilon_{11}(\chi, \nu_1, \zeta_1)|}{M_1(-\tfrac{1}{2}\chi\nu_1, \zeta_1\sqrt{2\chi})}, \frac{|\partial\varepsilon_{11}(\chi, \nu_1, \zeta_1)/\partial\zeta_1|}{(2\chi)^{1/2}N_1(-\tfrac{1}{2}\chi\nu_1, \zeta_1\sqrt{2\chi})} \leqq \frac{[\exp\{\tfrac{1}{2}(\chi/\pi)^{1/2}\mathscr{V}_{\zeta_1,\hat{\zeta}_1}(F_1)\}-1]}{E_1(-\tfrac{1}{2}\chi\nu_1^2, \zeta_1\sqrt{2\chi})}, \tag{6.8a}$$

and

$$\frac{|\varepsilon_{12}(\chi, \nu_1, \zeta_1)|}{M_1(-\tfrac{1}{2}\chi\nu_1^2, \zeta_1\sqrt{2\chi})}, \frac{|\partial\varepsilon_{12}(\chi, \nu_1, \zeta_1)/\partial\zeta_1|}{(2\chi)^{1/2}N_1(-\tfrac{1}{2}\chi\nu_1^2, \zeta_1\sqrt{2\chi})} \leqq E_1(-\tfrac{1}{2}\chi\nu_1^2, \zeta_1\sqrt{2\chi})\left[\exp\left\{\frac{1}{2}\left(\frac{\pi}{\chi}\right)^{1/2}\mathscr{V}_{0,\zeta_1}(F_1)\right\}-1\right]. \tag{6.8b}$$

In the above formulas $\hat{\zeta}_1 = \zeta_1(K)$.

On returning to the original independent variable, one finds that solutions of Lamé's equation have the general form

$$\left(\frac{dz}{d\zeta_1}\right)^{1/2}[A_{11}W_{11}(\chi, \nu_1, \zeta_1) + A_{12}W_{12}(\chi, \nu_1, \zeta_1)], \tag{6.9}$$

where A_{11}, A_{12} are constants. As Lamé polynomials are real valued on the real line and all the functions in (6.9) are also real, it follows that these constants will be real for Lamé polynomials. As the function $\psi_1(\chi, \nu_1, \zeta_1)$ is bounded for $\zeta_1 \in [0, \hat{\zeta}_1]$ the variation of F_1 will also be bounded. One may evaluate the error control function, replacing the auxiliary function M_1 by Airy functions [13]. It may be shown that the variation is $O(\chi^{-1/2})$. Consequently the inequalities for the error terms have a factor of $O(1/\chi)$ on the right hand side, which is sufficient to ensure that the error terms are satisfactorily small.

Case (b). The turning point z_0 of Lamé's equation may be close to $z = K$. There is a corresponding turning point at $2K - z_0$, which may also be close to $z = K$. Between these turning points the leading term of the coefficient of $w(z)$ in (6.1b) is positive. Hence the following transformations are used:

$$W(\zeta_2) = \left(\frac{dz}{d\zeta_2}\right)^{-1/2} w(z), \tag{6.10a}$$

and

$$\left(\frac{dz}{d\zeta_2}\right)^2 k^2(\alpha^2 - \mathrm{sn}^2 z) = \zeta_2^2 - \nu_2^2. \tag{6.10b}$$

The parameter ν_2 is once again at our disposal: on this occasion is it specified by the requirement that $\zeta_2(K)=0$ and $\zeta_2(z_0)=\nu_2$. Thus from (6.10b),

$$k\int_{z_0}^{K}(\mathrm{sn}^2\, t-\alpha^2)^{1/2}\,dt=\int_0^{\nu_2}(\nu_2^2-\tau^2)^{1/2}\,d\tau, \tag{6.11}$$

i.e.

$$\nu_2^2=\frac{4k}{\pi}\int_{z_0}^{K}(\mathrm{sn}^2\, t-\alpha^2)^{1/2}\,dt. \tag{6.12}$$

Once again on account of symmetry it will not be necessary to consider negative values of ζ_2. The relation between ζ_2 and z may be deduced from (6.10b) to be

$$k\int_{z}^{K}(\mathrm{sn}^2\, t-\alpha^2)^{1/2}\,dt=\int_0^{\zeta_2}(\nu_2^2-\tau^2)^{1/2}\,d\tau \quad \text{for } z\geqq z_0 \tag{6.12a}$$

and

$$k\int_{z}^{z_0}(\alpha^2-\mathrm{sn}^2\, t)^{1/2}\,dt=\int_{\nu_2}^{\zeta_2}(\tau^2-\nu_2^2)^{1/2}\,d\tau \quad \text{for } z\leqq z_0. \tag{6.12b}$$

Thus ζ_2 is a continuous, decreasing function of z for $z\in[0, K]$. The transformations (6.10a, b) lead to the comparison equation

$$\frac{d^2W}{d\zeta_2^2}=\{\chi^2(\nu_2^2-\zeta_2^2)+\psi_2(\chi, \nu_2, \zeta_2)\}W \tag{6.13}$$

where

$$\psi_2(\chi, \nu_2, \zeta_2)=-\frac{1}{4}\left(\frac{dz}{d\zeta_2}\right)^2 k^2\,\mathrm{sn}^2\, z+\left(\frac{dz}{d\zeta_2}\right)^{1/2}\frac{d^2}{d\zeta_2^2}\left(\frac{dz}{d\zeta_2}\right)^{-1/2}. \tag{6.14}$$

A suitable function for definition of the error control function is

$$\Omega_2(x)=\frac{1}{M_2^2(\frac{1}{2}\chi\nu_2^2, x)}.$$

The error control function is

$$F_2(\chi, \nu_2, \zeta_2)=\int\frac{|\psi_2(\chi, \nu_2, \zeta_2)|}{\Omega_2(\zeta_2\sqrt{2\chi})}\,d\zeta_2.$$

The main theorem of [13] may now be applied to (6.13). This equation, consequently, has two continuous, twice differentiable, independent solutions given by

$$W_{13}(\chi, \nu_2, \zeta_2)=l^{-1/2}(\tfrac{1}{2}\chi\nu_2^2)\,W(\tfrac{1}{2}\chi\nu_2^2, \zeta_2\sqrt{2\chi})+\varepsilon_{13}(\chi, \nu_2, \zeta_2), \tag{6.15a}$$

$$W_{14}(\chi, \nu_2, \zeta_2)=l^{1/2}(\tfrac{1}{2}\chi\nu_2^2)\,W(\tfrac{1}{2}\chi\nu_2^2, -\zeta_2\sqrt{2\chi})+\varepsilon_{14}(\chi, \nu_2, \zeta_2), \tag{6.15b}$$

where

$$\frac{|\varepsilon_{13}(\chi, \nu_2, \zeta_2)|}{M_2(\frac{1}{2}\chi\nu_2^2, \zeta_2\sqrt{2\chi})}, \frac{|\partial\varepsilon_{13}(\chi, \nu_2, \zeta_2)/\partial\zeta_2|}{(2\chi)^{1/2}N_2(\frac{1}{2}\chi\nu_2^2, \zeta_2\sqrt{2\chi})} \tag{6.16a}$$
$$\leqq\left[\exp\left\{\left(\frac{1}{2\chi}\right)^{1/2}\mathcal{V}_{\zeta_2,\hat{\zeta}_2}(F_2)\right\}-1\right]\Big/E_2(\tfrac{1}{2}\chi\nu_2^2, \zeta_2\sqrt{2\chi})$$

and

$$\frac{|\varepsilon_{14}(\chi, \nu_2, \zeta_2)|}{M_2(\frac{1}{2}\chi\nu_2^2, \zeta_2\sqrt{2\chi})}, \frac{|\partial\varepsilon_{14}(\chi, \nu_2, \zeta_2)/\partial\zeta_2|}{(2\chi)^{1/2}N_2(\frac{1}{2}\chi\nu_2^2, \zeta_2\sqrt{2\chi})} \tag{6.16b}$$
$$\leqq E_2(\tfrac{1}{2}\chi\nu_2^2, \zeta_2\sqrt{2\chi})\left[\exp\left\{\left(\frac{1}{2\chi}\right)^{1/2}\mathcal{V}_{0,\zeta_2}(F_2)\right\}-1\right].$$

In the above formulas $\hat{\zeta}_2 = \zeta_2(0)$. Thus in terms of the original independent variable, the general solution of Lamé's equation may be expressed as

$$\left(\frac{dz}{d\zeta_2}\right)^{1/2}[A_{13}W_{13}(\chi, \nu_2, \zeta_2)+A_{14}W_{14}(\chi, \nu_2, \zeta_2)], \tag{6.17}$$

where A_{13}, A_{14} are constants. As $(dz/d\zeta_2)$ is negative and Lamé polynomials are real in $[0, K]$, then A_{13}, A_{14} corresponding to Lamé polynomials will be pure imaginary. One may also observe that as $\psi_2(\chi, \nu_2, \zeta_2)$ is bounded for $\zeta_2 \in [0, \hat{\zeta}_2]$, the variation of F_2 is $O(1/\chi^{1/2})$ and consequently the right hand sides of both inequalities for the error terms have factors of $O(1/\chi)$. This implies that the bounds for the error terms are satisfactorily small.

7. Solutions of Lamé's equation on the line Re $(z)=K$.

Case (a). The turning point is bounded away from $z = K$ as $\chi \to \infty$, so that the interval $[K, K+iK']$ is free from turning points. For a region free from turning points it is natural to use the Liouville–Green approximation as in [12, p. 222]. Observe that although z is complex in this region the functions of z occurring in (6.1b) are real. The results of the Liouville–Green method may be applied immediately. However it is first convenient to adopt a notation for some functions of z which occur frequently. In (6.1b) define

$$f(\chi, z) = k^2(\mathrm{sn}^2 z - \alpha^2), \tag{7.1a}$$

$$g(z) = -\tfrac{1}{4}k^2 \mathrm{sn}^2 z. \tag{7.1b}$$

Thus equation (6.1b) may be written as

$$w''(z) = \{\chi^2 f(\chi, z) + g(z)\}w(z). \tag{7.1c}$$

In (7.1c), $g(z)$ is $O(1)$ so that the equation has the correct form for the Liouville–Green approximation. Consequently (7.1c) has two twice differentiable, independent solutions given by

$$W_{21}(z)=\{f(\chi,z)\}^{-1/4}\exp\left\{\chi\int_K^z f^{1/2}(\chi,t)\,dt\right\}\{1+\varepsilon_{21}(\chi,z)\},\tag{7.2a}$$

$$W_{22}(z)=\{f(\chi,z)\}^{-1/4}\exp\left\{-\chi\int_K^z f^{1/2}(\chi,t)\,dt\right\}\{1+\varepsilon_{22}(\chi,z)\},\tag{7.2b}$$

where

$$\begin{aligned}&|\varepsilon_{2j}(\chi,z)|,\ \chi^{-1}f^{-1/2}(\chi,z)|\varepsilon'_{2j}(\chi,z)|\\ &\qquad\leqq\exp\left\{\frac{1}{\chi}\mathscr{V}_{a_j,z}(G)\right\}-1, \qquad j=1,2,\end{aligned}\tag{7.2c}$$

where

$$G(\chi,z)=\int\left\{\frac{1}{f^{1/4}}\frac{d^2}{dz^2}\left(\frac{1}{f^{1/4}}\right)-\frac{g}{f^{1/2}}\right\}dz,\tag{7.2d}$$

and $a_1=K$, $a_2=K+iK'$ and $\mathscr{V}$ is the variational operator.

It is easily verified that the integrand of (7.2d) is purely real. Note also that the solutions (7.2a, b) are oscillatory as the integrand $f^{1/2}(\chi,t)$ is real but the region of integration is in the imaginary direction. The right hand side of relation (7.2c) is $O(\chi^{-1})$ which provides a realistic error bound.

The solutions (7.2a, b) allow us to express the general form of solutions of Lamé's equation on $[K, K+iK']$, as

$$A_{21}W_{21}(z)+A_{22}W_{22}(z),\tag{7.3}$$

where A_{21}, A_{22} are complex constants.

Case (b). In this situation Lamé's equation has two turning points at z_0, $2K-z_0$ which are arbitrarily close to K as $z_0\to K$. For the same reason as in Case (a) above it is possible to apply the theory of [13] for the case of two turning points arbitrarily close to the interval of definition of the differential equation.

Similar transformations to (6.10) are made, the new independent variable being ζ_3 and the sign of ν_2^2 changing to positive in (6.10b). Integration of this latter equation shows that the choice of ν_2 is consistent with the conditions $\zeta_3(K)=0$ and $\zeta(z_0)=i\nu_2$, since

$$k\int_K^{z_0}(\alpha^2-\mathrm{sn}^2\,t)^{1/2}\,dt=\frac{i\pi\nu_2^2}{4}=\int_0^{i\nu_2}(\tau^2+\nu_2^2)^{1/2}\,d\tau,$$

the argument of $(\alpha^2-\mathrm{sn}^2\,t)^{1/2}$ being $-\frac{1}{2}\pi$ in this interval (cf. § 3).

The main theorem of [13] may be applied to deduce that the general solution of Lamé's equation is of the form

$$\left(\frac{dz}{d\zeta_3}\right)^{1/2}\{A_{23}W_{23}(\chi,\nu_2,\zeta_3)+A_{24}W_{24}(\chi,\nu_2,\zeta_3)\},\tag{7.4}$$

where A_{23}, A_{24} are constants. In (7.4),

$$W_{23}(\chi, \nu_2, \zeta_3) = W(-\tfrac{1}{2}\chi\nu_2^2, \zeta_3\sqrt{2\chi}) + \varepsilon_{23}(\chi, \nu_2, \zeta_3),$$
$$W_{24}(\chi, \nu_2, \zeta_3) = W(-\tfrac{1}{2}\chi\nu_2^2, \zeta_3\sqrt{2\chi}) + \varepsilon_{24}(\chi, \nu_2, \zeta_3),$$

where the error terms have a factor $O(\chi^{-1})$ and are consequently realistically small (cf. § 6, Case (b)).

8. Solutions of Lamé's equation on $[K+iK', iK']$. The derivation of solutions for Lamé's equation in this region is simpler than the same problem for the other three intervals. Observe that the turning point of Lamé's equation is located on the real line for all values of $\alpha \in [0, 1]$. Thus the line $\operatorname{Im}(z) = K'$ is free from turning points in both cases and the Liouville–Green method may be applied.

Case (a). Recall the formulation of Lamé's equation described by equations (7.1a, b, c). Equation (7.1c) has two twice differentiable, independent solutions

$$W_{31}(z) = \{f(\chi, z)\}^{-1/4} \exp\left\{\chi \int_{K+iK'}^{z} f^{1/2}(\chi, t)\, dt\right\}\{1 + \varepsilon_{31}(\chi, z)\}, \tag{8.1a}$$

$$W_{32}(z) = \{f(\chi, z)^{-1/4} \exp\left\{-\chi \int_{K+iK'}^{z} f^{1/2}(\chi, t)\, dt\right\}\{1 + \varepsilon_{32}(\chi, z)\}, \tag{8.1b}$$

where

$$|\varepsilon_{31}(\chi, z)|, \frac{|\varepsilon'_{31}(\chi, z)|}{2\chi f^{1/2}(\chi, z)} \leqq \exp\left\{\frac{1}{2\chi} \mathscr{V}_{K+iK', z}(G)\right\} - 1 \tag{8.2a}$$

and

$$|\varepsilon_{32}(\chi, z)|, \frac{|\varepsilon'_{32}(\chi, z)|}{2\chi f^{1/2}(\chi, z)} \leqq \exp\left\{\frac{1}{2\chi} \mathscr{V}_{iK', z}(G)\right\} - 1 \tag{8.2b}$$

with G defined by (7.2d).

There are two remarks which should be made concerning the above expressions. First of all $f^{1/2}(\chi, z)$ is real for $z \in [K+iK', iK']$, so that one solution will be exponentially increasing and one exponentially decreasing as z moves away from $K+iK'$. Secondly, the differential equation (7.1c) possesses a double pole at $z = iK'$. It is therefore necessary to show that these asymptotic solutions are valid in the neighborhood of this pole.

The justification of this assertion presents the first illustration that the choice of large parameter χ made in (1.8) was correct. Conditions for $\mathscr{V}_{K+iK', z}(G)$ and $\mathscr{V}_{iK', z}(G)$ to be convergent are given by [12, p. 205]. The dominant function $f(\chi, z)$ has a double pole at $z = iK'$. The condition for convergence is that the function $g(z)$ may be expanded as a power series of the following form:

$$g(z) = \frac{-1}{4(iK'-z)^2}\left\{1 + \sum_{s=1}^{\infty} g_s(iK'-z)^s\right\}. \tag{8.3}$$

As $g(z) = -\frac{1}{4}k^2 \operatorname{sn}^2 z$ and $\operatorname{sn} z$ has a simple pole with residue $1/k$ at $z = iK'$, (8.3) is satisfied.

Thus the terms on the right hand side of (8.2a, b) are both $O(1/\chi)$ and the error terms are uniformly valid for $z \in [iK', K+iK']$. Furthermore (8.1a, b) give asymptotic representations for the dominant and recessive solutions in the neighborhood of $z = iK'$. The general solution of Lamé's equation in this region is thus

$$A_{31}W_{31}(z)+A_{32}W_{32}(z), \tag{8.4}$$

where A_{31}, A_{32} are constants. As Lamé polynomials are real in this region if $\sigma+\tau \neq 1$, A_{31}, A_{32} corresponding to Lamé polynomials will be real if this condition is satisfied and pure imaginary otherwise.

Case (b). The form of the solutions will be identical to those of (8.1a, b), (8.2a, b) and for the same reasons as in Case (a) the asymptotic representations will be valid at $z = iK'$. However, the constant multipliers of W_{31}, W_{32} corresponding to Lamé polynomials will necessarily be different due to the differences in the behavior of the Lamé polynomials between Cases (a) and (b) in the regions $[0, K]$ and $[K, K+iK']$. Thus the general solution of Lamé's equation will be

$$A_{33}W_{31}(z)+A_{34}W_{32}(z), \tag{8.5}$$

where the constants A_{33}, A_{34} are governed by σ and τ as in Case (a) above.

9. Solutions of Lamé's equation on $[0, iK']$.

Case (a). Lamé's equation does not have a turning point on the line $\mathrm{Re}\,(z) = 0$ in general. However, if α is allowed to approach zero two turning points located at $\pm z_0$ are coalescing at the origin so that in the limiting case $\alpha = 0$, Lamé's equation has a turning point of second order at $z = 0$. As $\alpha \to 0$ the function $\chi^2 f(\chi, z)$ of (7.1c) is no longer dominant on the whole of the interval $[0, iK']$ and the Liouville–Green approximation does not lead to uniformly valid asymptotic expansions. Consequently it is appropriate to use the theory of [13] for Case (a).

A further complication on the line $\mathrm{Re}\,(z) = 0$ is the presence of the pole at $z = iK'$. However, due to the correct choice of large parameter χ, it is possible to show that the error terms are bounded at the pole, and consequently the solutions obtained from the method of [13] provide asymptotic representations of the dominant and recessive solutions in the neighborhood of the pole.

Introduce the following transformations of dependent and independent variable:

$$W(\zeta_4) = \left(\frac{dz}{d\zeta_4}\right)^{-1/2} w(z) \tag{9.1a}$$

and

$$\left(\frac{dz}{d\zeta_4}\right)^2 k^2(\mathrm{sn}^2\, z - \alpha^2) = \nu_1^2 + \zeta_4^2, \tag{9.1b}$$

as the dominant term in (6.1b) is positive for $\mathrm{Re}\,(z) = 0$. The parameter ν_1 is defined by the relation (6.3). The choice of ν_1 is sufficient to ensure that

$$\zeta_4(0) = 0 \quad \text{and} \quad \zeta_4(z_0) = -i\nu_1,$$

since

$$k\int_0^{z_0}(\alpha^2-\operatorname{sn}^2\xi)^{1/2}\,d\xi=\frac{\pi\nu_1^2}{4}\quad\text{by (6.3)}$$

and

$$i\int_0^{-i\nu_1}(\nu_1^2+\tau^2)^{1/2}\,d\tau=i\left[-\frac{i\pi\nu_1^2}{4}\right]=\tfrac{1}{4}\pi\nu_1^2.$$

The relation between ζ_4 and z may thus be expressed as

$$k\int_0^{z}(\alpha^2-\operatorname{sn}^2\xi)^{1/2}\,d\xi=i\int_0^{\zeta_4}(\tau^2+\nu_1^2)^{1/2}\,d\tau. \tag{9.2}$$

Thus, as z increases on $\operatorname{Re}(z)=0$, ζ_4 is real and increasing. Furthermore ζ_4 is an analytic function of z, for $z\in[0, iK']$, the derivative $d\zeta_4/dz$ having argument $-\pi/2$.

Thus as z increases on $\operatorname{Re}(z)=0$, ζ_4 is real and increasing. Furthermore ζ_4 is

$$\frac{d^2W}{d\zeta_4^2}=\{\chi^2(\zeta_4^2+\nu_1^2)+\psi_4(\chi,\nu_1,\zeta_4)\}W, \tag{9.3}$$

where

$$\psi_4(\chi,\nu_1,\zeta_4)=-\frac{1}{4}\left(\frac{dz}{d\zeta_4}\right)^2k^2\operatorname{sn}^2 z+\left(\frac{dz}{d\zeta_4}\right)^{1/2}\frac{d^2}{d\zeta_4^2}\left(\frac{dz}{d\zeta_4}\right)^{-1/2}. \tag{9.4}$$

The function Ω_4 of the error control function may be conveniently specified as

$$\Omega_4(x)=\frac{2\pi^{1/2}}{\Gamma(\frac{1}{2}+\frac{1}{2}\chi\nu_1^2)\,|U(\frac{1}{2}\chi\nu_1^2,x)U(\frac{1}{2}\chi\nu_1^2,-x)|}.$$

This function is nonnegative and has the correct behavior as $x\to\pm\infty$. The error control function for this case is

$$F_4(\chi,\nu_1,\zeta_4)=\int\frac{|\psi_4(\chi,\nu_1,\zeta_4)|}{\Omega_4(\zeta_4\sqrt{2\chi})}\,d\zeta_4.$$

One may apply the main theorem of [13] to deduce that, for $\zeta_4\in[0,\infty)$, (9.3) has two continuous, twice differentiable, independent solutions, which are given by

$$W_{41}(\chi,\nu_1,\zeta_4)=U(\tfrac{1}{2}\chi\nu_1^2,\zeta_4\sqrt{2\chi})+\varepsilon_{41}(\chi,\nu_1,\zeta_4) \tag{9.5a}$$

and

$$W_{42}(\chi,\nu_1,\zeta_4)=U(\tfrac{1}{2}\chi\nu_1^2,-\zeta_4\sqrt{2\chi})+\varepsilon_{41}(\chi,\nu_1,-\zeta_4), \tag{9.5b}$$

where

$$|\varepsilon_{41}(\chi, \nu_1, \pm\zeta_4)| \leqq U(\tfrac{1}{2}\chi\nu_1^2, \pm\zeta_4\sqrt{2\chi})\left[\exp\left\{\frac{1}{\chi^{1/2}}\mathscr{V}_{\pm\zeta_4,\infty}(F_4)\right\}-1\right], \tag{9.6a}$$

$$\left|\frac{\partial\varepsilon_{41}(\chi, \nu_1, \pm\zeta_4)}{\partial\zeta_4}\right| \leqq (2\chi)^{1/2} l_1^{1/2}(\tfrac{1}{2}\chi\nu_1^2) U'(\tfrac{1}{2}\chi\nu_1^2, \pm\zeta_4\sqrt{2\chi}) \cdot\left[\exp\left\{\frac{1}{\chi^{1/2}}\mathscr{V}_{\pm\zeta_4,\infty}(F_4)\right\}-1\right], \tag{9.6b}$$

and

$$l_1(a) = 1 + \sup_{x\in(-\infty,\infty)}\left\{\frac{U'(a,-x)U(a,x)}{U'(a,x)U(a,-x)}\right\}.$$

In the above expressions, the variation of F_4 is defined over an infinite range of values of ζ_4, since $\zeta_4 \to \infty$ as $z \to iK'$. The integral will be convergent provided that

$$\psi_4(\chi, \nu_1, \zeta_4) = o(1) \quad \text{as } \zeta_4 \to \pm\infty. \tag{9.7}$$

The convergence of the integral is also sufficient to ensure that the error bounds are realistic and the two solutions of (9.5) provide asymptotic representations of the dominant and recessive solutions in the neighborhood of the pole.

To show that this is in fact the case, it is only necessary to consider the case of positive ζ_4. The result for negative values of ζ_4 follows immediately by a symmetry argument. The proof of this result is facilitated by considering both the Schwarzian derivative and the relation between ζ_4 and z in more explicit form.

Introduce the function $p(\chi, z_0, z)$ in which the first two arguments are usually suppressed, by

$$p(\chi, z_0, z) = k^2(\mathrm{sn}^2 z - \alpha^2)/(z^2 - z_0^2). \tag{9.8}$$

Then, for $z \in [0, iK')$, $p(z)$ is negative and in particular

$$p(z) = \frac{1}{(z^2-z_0^2)|(iK'-z)|^2}\{1 + O(|iK'-z|^2)\} \quad \text{as } z \to iK'.$$

Denoting the derivative $dz/d\zeta_4$ by $\dot{z}$,

$$\dot{z}^{1/2}\frac{d^2}{d\zeta_4}(\dot{z}^{-1/2}) = \left[\left\{\frac{p''(z)}{p(z)} - \frac{5}{16}\frac{p'^2(z)}{p^2(z)}\right\}(z^2-z_0^2)^2 - \frac{p'(z)}{4p(z)}z(z^2-z_0^2) - \tfrac{1}{4}(3z^2+2z_0^2)\right]\frac{\zeta_4^2+\nu_1^2}{p(z)(z^2-z_0^2)^3} + O\left(\frac{1}{\zeta_4^2}\right). \tag{9.9}$$

On account of the behavior of the function $p(z)$ in the neighborhood of iK',

$$\frac{p''(z)}{p(z)} = \frac{6}{|iK'-z|^2}\{1 + O(|iK'-z|^2)\},$$

and

$$\frac{p'(z)}{p(z)} = \frac{2}{|iK'-z|}\{1 + O(iK'-z|^2)\}$$

so that the right hand side of (9.9) is

$$\frac{1}{4}\frac{1}{|iK'-z|^2}\{1+O(|iK'-z|)\}\frac{\zeta_4^2+\nu_1^2}{p(z)(z^2-z_0^2)}+O\left(\frac{1}{\zeta_4^2}\right).$$

Once again applying the result for $p(z)$ as $z \to iK'$, we see that this reduces to

$$\tfrac{1}{4}(\zeta_4^2+\nu_1^2)\{1+O(|iK'-z|)\}+O\left(\frac{1}{\zeta_4^2}\right).$$

Now

$$\begin{aligned}\dot{z}^2 g(z) &= \frac{\zeta_4^2+\nu_1^2}{k^2(\operatorname{sn}^2 z-\alpha^2)}(-\tfrac{1}{4}k^2\operatorname{sn}^2 z)\\ &= -\tfrac{1}{4}(\zeta_4^2+\nu_1^2)\{1+O(|iK'-z|^2)\}.\end{aligned}$$

Hence

$$\psi_4(\chi, \nu_1, \zeta_4) = O\left(\zeta_4^2|iK'-z|+\frac{1}{\zeta_4^2}\right). \tag{9.10}$$

In order to show that ψ_4 is $o(1)$ as $\zeta_4 \to \infty$, one needs to know the relation between ζ_4 and z and $z \to iK'$. In (9.2) α and ν_1 are both $O(1)$ so that we may deduce that the dominant term on the left hand side of (9.2) is

$$-k\int_{z_1}^{z} i\operatorname{sn}\zeta\, d\zeta = -i\ln\frac{1}{|iK'-z|}+O(1),$$

for some finite $z_1 \in (0, iK')$, such that $|\operatorname{sn} z_1| > \alpha$.

The negative sign appears as $(\alpha^2-\operatorname{sn}^2\zeta)^{1/2}$ is positive in this region whilst $\operatorname{sn}\zeta$ is pure imaginary with argument $\pi/2$. The right hand side of (9.2) is

$$i\left[\tfrac{1}{2}\nu_1^2\ln\frac{2\zeta_4}{\nu_1}+\tfrac{1}{2}\zeta_4^2\right]+O(1).$$

Thus on exponentiating (9.2) and comparing the large terms one obtains

$$C|iK'-z| = \zeta_4^{-\nu_1^2/2}\exp\{-\tfrac{1}{2}\zeta_4^2\}\{1+o(1)\}, \tag{9.11}$$

and it follows that the order term on the right hand side of (9.10) may be replaced by $O(\zeta_4^{-2})$ and (9.7) holds. Consequently the general solution of Lamé's equation in the region $[0, iK']$ may be expressed as

$$\left(\frac{dz}{d\zeta_4}\right)^{1/2}[A_{41}W_{41}(\chi, \nu_1, \zeta_4)+A_{42}W_{42}(\chi, \nu_1, \zeta_4)], \tag{9.12}$$

where A_{41}, A_{42} are constants. On account of the result proved for $\psi_4(\chi, \nu_1, \zeta_4)$ as $\zeta_4 \to \infty$, it follows that all error bounds are realistic, containing a factor of $O(\chi^{-1})$, and are uniformly valid for $z \in [0, iK']$.

Case (b). A much simpler problem ensues when the turning point of Lamé's equation is bounded away from the origin. The Liouville–Green approximation may be used to deduce that two independent solutions of Lamé's equation are

given by

$$(9.13a)\quad W_{43}(z)=\{-f(\chi,z)\}^{-1/4}\exp\left\{i\chi\int_0^z\{-f(\chi,t)\}^{1/2}\,dt\right\}\{1+\varepsilon_{43}(\chi,z)\}$$

and

$$(9.13b)\quad W_{44}(z)=\{-f(\chi,z)\}^{-1/4}\exp\left\{-i\chi\int_0^z\{-f(\chi,t)\}^{1/2}\,dt\right\}\{1+\varepsilon_{44}(\chi,z)\}.$$

As before the error terms are uniformly valid for $z\in[0,iK']$, possessing a factor of $O(\chi^{-1})$. Thus the solution of Lamé's equation in this region has the general form

$$(9.14)\qquad A_{43}W_{32}(z)+A_{44}W_{44}(z),$$

where A_{43}, A_{44} are constants,

10. Identification of Lamé polynomials for Case (a). The identification of Lamé polynomials consists of the determination of the constants occurring as multiplicative factors of solutions of Lamé's equation. In each region a pair of such factors occurs. The criteria available for the determination of these factors are of two types. First of all there are boundary conditions to be satisfied, which enable us to relate one factor to the second member of the pair. If there are two boundary conditions to be satisfied on one interval the fact that h assumes an eigenvalue ensures that the second boundary condition is satisfied.

Each Lamé polynomial has associated parameters ρ, σ and τ, which determine the boundary conditions. These conditions are of the form (1.4). The quantities, which are nonzero at the points 0, K, $K+iK'$, introduced in § 1 as matching coefficients, are used for the determination of the remaining multiplicative constant associated with a Lamé polynomial. In the region $[0,K]$, the normalization condition (1.7) must be satisfied at the origin. In the other regions the matching coefficients are used to continue the asymptotic formulas from one region to the next.

There are several methods available for checking certain aspects of the result. These checks will be valid for all three cases. First of all the solution is continued around the basic rectangle in an anticlockwise direction, so that one may check that the Lamé polynomial and its derivative have the correct value on returning to the origin. The continuation of a Lamé polynomial from $(iK',K+iK')$ to the interval $[0,iK')$ is complicated by the existence of a singularity of the Lamé polynomial at $z=iK'$. Stokes' phenomenon occurs at this pole, the phase of the Lamé polynomial being discontinuous. In order to calculate the phase discontinuity consider a typical Lamé polynomial at $z_1=iK'+\varepsilon$ and at $z_2=iK'-i\varepsilon$, where ε is some small, real, positive number. The Lamé polynomials will behave as their term of highest power of $\mathrm{sn}^2 z$, so that

$$E_{2n+p}^m(z_1)\sim A\,\mathrm{sn}^\rho z_1\,\mathrm{cn}^\sigma z_1\,\mathrm{dn}^\tau z_1\,\mathrm{sn}^{2n}z_1=\frac{A'}{\varepsilon^{2n+p}},$$

and

$$E^m_{2n+p}(z_2) \sim A \operatorname{sn}^\rho z_2 \operatorname{cn}^\sigma z_2 \operatorname{dn}^\tau z_2 \operatorname{sn}^{2n} z_2 = \frac{A'}{(-i\varepsilon)^{2n+p}},$$

for some constants A and A'. Consequently,

$$\arg\{E^m_{2n+p}(z_2)\} \sim \arg\{E^m_{2n+p}(z_1)\} + n\pi + \tfrac{1}{2}p\pi, \tag{10.1a}$$

and

$$|E^m_{2n+p}(z_2)| \sim |E^m_{2n+p}(z_1)|. \tag{10.1b}$$

The asymptotic relations (10.1a, b) are the required formulas for the connection of Lamé polynomials on these two intervals.

A second check on our results is that the behavior of the Lamé polynomials in the neighborhood of $z = iK'$ should be such that the dominant term is $O((1/|iK' - z|)^{2n+p})$ as $z \to iK'$.

The normalization condition implies that a Lamé polynomial is positive in some deleted neighborhood of the origin. Thus it is possible to predict the sign of a Lamé polynomial in certain regions:

$$\operatorname{sgn}\{E^{m(\sigma)}_{2n+p}(K)\} = (-1)^{m+\sigma} \tag{10.2a}$$

$$\arg\{E^{m(\tau)}_{2n+p}(K+iK')\} = [\tfrac{1}{2}(\tau-\sigma)\pi + n\pi], \tag{10.2b}$$

$$\arg\{E^m_{2n+p}(z)\} = [-\tfrac{1}{2}(\tau+\sigma)\pi + n\pi], \qquad z \in (iK', K+iK') \tag{10.2c}$$

$$\arg\{E^m_{2n+p}(z)\} = \tfrac{1}{2}\rho\pi, \qquad z \in (0, iK'). \tag{10.2d}$$

When the multiplicative constants are being calculated, only the leading term of the asymptotic series for the constants will be calculated. As error bounds are available for the solutions, elementary error analysis shows that the relative error in the constant, due to the neglected terms in the asymptotic series, is the same as the relative error in the solution at the point at which the constant is evaluated. For example, if $w(x)$ is the solution with error $\varepsilon(x)$ and the multiplicative constant A is to be calculated by the condition that $Aw(x_0) = \eta$, then if A has error $\hat{\varepsilon}$,

$$(A + \hat{\varepsilon})\{w(x_0) + \varepsilon(x_0)\} = \eta,$$

then

$$\hat{\varepsilon} w(x_0) + A\varepsilon(x_0) = 0 \quad \text{and} \quad \left|\frac{\hat{\varepsilon}}{A}\right| = \left|\frac{\varepsilon(x_0)}{w(x_0)}\right|.$$

As all error terms calculated are satisfactorily small, the relative error being $O(1/\chi)$, all multiplicative constants will have the same order of accuracy.

For the region $[0, K]$, Lamé polynomials are described by expression (6.9). The parabolic cylinder functions occurring in this expression have a first argument which may be large and negative or which tends to zero as $\nu_1 \to 0$. In both cases there exists a finite interval of the real line extending to $z = K$ on which $U(-\frac{1}{2}\chi\nu_1, \zeta_1\sqrt{2\chi})$ is exponentially decreasing and $\bar{U}(-\frac{1}{2}\chi\nu_1, \zeta_1\sqrt{2\chi})$ is exponentially increasing. Furthermore neither these functions nor their derivatives may have a zero when $z = K$. Therefore to satisfy the boundary condition at $z = K$

($\zeta_1 = \hat{\zeta}_1$) the factors $A_{11}W_{11}$ and $A_{12}W_{12}$ must be of the same order of magnitude at $z = K$. Thus away from this point, it follows from the behavior of the parabolic cylinder functions mentioned above that $A_{11}W_{11}$ dominates $A_{12}W_{12}$.

In order to satisfy the boundary condition we must have

$$A_{11}U(-\tfrac{1}{2}\chi\nu_1^2, \hat{\zeta}_1\sqrt{2\chi}) + A_{12}\bar{U}(-\tfrac{1}{2}\chi\nu_1^2, \hat{\zeta}_1\sqrt{2\chi}) = 0 \quad \text{if } \sigma = 1$$

or

$$A_{11}U'(-\tfrac{1}{2}\chi\nu_1^2, \hat{\zeta}_1\sqrt{2\chi}) + A_{12}\bar{U}'(-\tfrac{1}{2}\chi\nu_1^2, \hat{\zeta}_1\sqrt{2\chi}) = 0 \quad \text{of } \sigma = 0.$$

Thus replacing the parabolic cylinder functions by the leading term of their asymptotic expansions [11] and combining the two cases, we obtain

$$\begin{aligned} A_{12} \sim A_{11}(-1)^{\sigma} 2^{-\chi\nu_1^2 - 1/2}\pi^{1/2}\{\Gamma(\tfrac{1}{2}+\tfrac{1}{2}\chi\nu_1^2)\}^{-1} \\ \cdot e^{-\chi\nu_1^2/2}\nu^{\chi\nu_1^2}\chi^{\chi\nu_1^2/2} \exp\left\{-2\chi\int_{\nu_1}^{\hat{\zeta}_1}(\tau^2-\nu_1^2)^{1/2}\,d\tau\right\}. \end{aligned} \tag{10.3}$$

At the origin, the first term of (6.9) dominates so that A_{11} may be evaluated by matching the value of the Lamé polynomial with the normalization condition. However, this condition involves a sine or cosine function with an argument $\frac{1}{4}\pi + \frac{1}{2}\chi\nu_1^2$. It is possible to evaluate this quantity by recalling the result (3.12) and simultaneously verify that the eigenvalue of the Lamé polynomial satisfies the boundary condition at the origin.

The parameter ν_1 is defined by the equation (6.3),

$$\nu_1^2 = \frac{4k}{\pi}\int_0^{z_0}(\alpha^2 - \mathrm{sn}^2\, t)^{1/2}\,dt.$$

Thus from (3.12),

$$\chi\nu_1^2 = \frac{4}{\pi}[m + \tfrac{1}{2}\rho + \tfrac{1}{4}]\pi = 4m + 2\rho + 1.$$

The leading term of the asymptotic representation at the origin yields the following results:

$$\begin{aligned} E^m_{2n+p}(0) = A_{11}\left(\frac{\nu_1}{\pi k\alpha}\right)^{1/2} 2^{m+\rho/2}\Gamma(\tfrac{1}{2}+\tfrac{1}{2}\rho+m) \\ \cdot [\sin(m+\tfrac{1}{2}\rho+\tfrac{1}{2})\pi + O(\chi^{-1})] \end{aligned} \tag{10.4a}$$

and

$$\begin{aligned} E^{m'}_{2n+p}(0) = -A_{11}\left(\frac{\alpha k\chi}{\pi\nu_1}\right)^{1/2} 2^{m+1+\rho/2}\Gamma(1+\tfrac{1}{2}\rho+m) \\ \cdot [\sin(m+\tfrac{1}{2}\rho+1)\pi + O(\chi^{-1})]. \end{aligned} \tag{10.4b}$$

The presence of the term $\sin[(m+\frac{1}{2}\rho+\frac{1}{2})\pi]$ in (10.4a) indicates that the Lamé polynomial has m zeros in the interval $(0, K)$. Note also that if ρ is equal to zero the expression on the right hand side of (10.4b) is zero and if ρ is unity the expression on the right hand side of (10.4a) is zero. Thus the boundary condition

at the origin is satisfied by the leading term of the Lamé polynomial. The matching coefficient at the origin is also determined from the expressions (10.4a, b).

If $\rho = 0$,

$$E_{2n+p}^{m}(0) \sim A_{11}\left(\frac{\nu_1}{\pi k\alpha}\right)^{1/2} 2^m \Gamma(m+\tfrac{1}{2})(-1)^m,$$

and if $\rho = 1$,

$$E_{2n+p}^{m'}(0) \sim A_{11}\left(\frac{\alpha k\chi}{\pi\nu_1}\right)^{1/2} 2^{m+3/2}\Gamma(m+\tfrac{3}{2})(-1)^m.$$

Thus in order to satisfy the normalization condition

$$E_{2n+p}^{m(\rho)}(0) = 1,$$

it is necessary that

$$A_{11} \sim (-1)^m \frac{\pi^{1/2}}{2^{m+3\rho/2}\chi^{\rho/2}}\left(\frac{\nu_1}{k\alpha}\right)^{\rho-1/2} \frac{1}{\Gamma(m+\rho+\frac{1}{2})}. \tag{10.5}$$

The behavior of the leading term of a Lamé polynomial on the real line has now been completely determined. The value of the Lamé polynomial or its derivative at $z = K$ is required for matching purposes. These terms are

$$E_{2n+p}^{m}(K) \sim A_{11} 2^{3/4-\chi\nu_1^2/4}\chi^{\chi\nu_1^2/4-1/4} e^{-\chi\nu_1^2/4}\nu_1^{\chi\nu_1^2/2} k^{-1/2}(1-\alpha^2)^{-1/4}$$
$$\cdot \exp\left\{-\chi\int_{\nu_1}^{\hat{\xi}_1} (\tau^2-\nu_1^2)^{1/2}\,d\tau\right\} \quad \text{if } \sigma = 0$$

and

$$E_{2n+p}^{m'}(K) \sim -A_{11} 2^{3/4-\chi\nu_1^2/4}\chi^{\chi\nu_1^2/4+3/4} k^{1/2}(1-\alpha^2)^{1/4} e^{-\chi\nu_1^2/4}\nu_1^{\chi\nu_1^2/2}$$
$$\cdot \exp\left\{-\chi\int_{\nu_1}^{\hat{\xi}_1} (\tau^2-\nu_1^2)^{1/2}\,d\tau\right\} \quad \text{if } \sigma = 1.$$

On combining these expressions and substituting for A_{11}, $\chi\nu_1^2$ and $\hat{\xi}_1$, one obtains

$$E_{2n+p}^{m(\sigma)}(K) \sim (-1)^{m+\sigma} k^{\sigma-1/2}(1-\alpha^2)^{\sigma/2-1/4}\chi^{\sigma}\mu_1, \tag{10.6a}$$

where

$$\mu_1 = \pi^{1/2}(\nu_1/k\alpha)^{\rho-1/2} 2^{1/2-2m-2\rho}\chi^m(\nu_1^2/e)^{m+\rho/2+1/4}$$
$$\cdot \exp\left\{-\chi\int_{z_0}^{K} (\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi\right\}\Big/\Gamma(m+\rho+\tfrac{1}{2}). \tag{10.6b}$$

Comparison of equations (10.6) and (10.2a) indicates that the sign of the above expression is correct.

Sufficient information concerning the behavior of Lamé polynomials on the real line has been obtained, and it is now possible to investigate the behavior of Lamé polynomials in $[K, K+iK']$, matching the solutions at $z = K$. The Liouville–Green approximation is uniformly valid in this region, the form of a Lamé

polynomial being

$$A_{21}W_{21}(z)+A_{22}W_{22}(z). \tag{10.7}$$

At $z=K$,

$$W_{21}(K)\sim W_{22}(K)\sim k^{-1/2}(1-\alpha^2)^{-1/4}$$

and

$$W'_{21}(K)\sim -W'_{22}(K)\sim \chi k^{1/2}(1-\alpha^2)^{1/4}.$$

The above relations imply that

$$A_{22}\sim(-1)^{\sigma}A_{21} \tag{10.8}$$

in order that the Lamé polynomial may satisfy the boundary condition at $z=K$. Thus (10.7) may be expressed as

$$E^m_{2n+p}(z)=2A_{21}\{k^2(\mathrm{sn}^2\, z-\alpha^2)\}^{-1/4}$$
$$\cdot\left[\cos\left\{-i\chi k\int_K^z(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi\right\}+O(\chi^{-1})\right] \quad \text{if } \sigma=0 \tag{10.9a}$$

or

$$E^m_{2n+p}(z)=2iA_{21}\{k^2(\mathrm{sn}^2\, z-\alpha^2)\}^{-1/4}$$
$$\cdot\left[\sin\left\{-i\chi k\int_K^z(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi\right\}+O(\chi^{-1})\right] \quad \text{if } \sigma=1. \tag{10.9b}$$

Thus on comparing the expressions (10.9a) if $\sigma=0$ or the corresponding expression for the derivative in (10.9b) if $\sigma=1$, evaluated at $z=K$ with (10.6a, b) it follows that

$$A_{21}\sim(-1)^{m+\sigma}\mu_1. \tag{10.10}$$

The computation of the multiplicative constants A_{21}, A_{22} determines the form of the leading term of a Lamé polynomial on $[K, K+iK']$. It is necessary to evaluate the Lamé polynomial and its derivative at $K+iK'$ for matching purposes, and simultaneously verify that the boundary condition is satisfied at this point. The expressions for the leading term of a Lamé polynomial (10.9a, b) may be conveniently combined into one formula as

$$E^m_{2n+p}(z)=2i^{\sigma}A_{21}\{k^2(\mathrm{sn}^2\, z-\alpha^2)\}^{-1/4}$$
$$\cdot\left[\sin\left\{-i\chi k\int_K^z(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi+\tfrac{1}{2}\pi(1-\sigma)\right\}+O(\chi^{-1})\right].$$

At $z=K+iK'$, Lamé polynomials have the following behavior:

$$E^m_{2n+p}(K+iK')=2i^{\sigma}A_{21}(1-k^2\alpha^2)^{-1/4}$$
$$\cdot\left[\sin\left\{-i\chi k\int_K^{K+iK'}(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi+\tfrac{1}{2}\pi(1-\sigma)\right\}+O(\chi^{-1})\right], \tag{10.11a}$$

and

$$E^{m'}_{2n+p}(K+iK') = 2i^{\sigma-1}A_{21}(1-k^2\alpha^2)^{1/4}\chi$$

$$\cdot\left[\cos\left\{-i\chi k\int_K^{K+iK'}(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi+\tfrac{1}{2}\pi(1-\sigma)\right\}+O(\chi^{-1})\right], \tag{10.11b}$$

since the argument of the trigonometric function is increasing as z increases in the positive imaginary direction. Thus for the case $\tau=0$, the boundary condition will be satisfied if

$$\cos\left\{-i\chi k\int_K^{K+iK'}(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi+\tfrac{1}{2}\pi(1-\sigma)\right\}=O(\chi^{-1}) \tag{10.12a}$$

and if $\tau=1$, it will be satisfied if

$$\sin\left\{-i\chi k\int_K^{K+iK'}(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi+\tfrac{1}{2}\pi(1-\sigma)\right\}=O(\chi^{-1}). \tag{10.12b}$$

The conditions (10.12a, b) may be combined into the single condition,

$$\sin\left\{-i\chi k\int_K^{K+iK'}(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi+\tfrac{1}{2}\pi(1-\sigma)+\tfrac{1}{2}\pi(1-\tau)\right\}=O(\chi^{-1}). \tag{10.13}$$

The integral appearing in (10.13) occurred in § 3 and its value is given by (3.8), i.e.,

$$-ik\chi\int_K^{K+iK'}(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi=[(n-m)\pi+\tfrac{1}{2}\pi(\sigma+\tau)]+O\left(\frac{1}{\chi}\right).$$

This immediately verifies that the boundary condition at $z=K+iK'$ is satisfied by the leading term of the Lamé polynomial. The presence of the $(n-m)\pi$ term in the integral of (10.13) indicates that the Lamé polynomial has $(n-m)$ zeros in the interval $(K, K+iK')$. The value of the Lamé polynomial and its derivative at $z=K+iK'$ are therefore given by

$$E^{m}_{2n+p}(K+iK')\sim 2i^{\sigma}A_{21}\{1-k^2\alpha^2\}^{-1/4}(-1)^{n-m} \quad \text{if } \tau=0 \tag{10.14a}$$

and

$$E^{m'}_{2n+p}(K+iK')\sim 2i^{\sigma-1}A_{21}\{1-k^2\alpha^2\}^{-1/4}\chi(-1)^{n-m+1} \quad \text{if } \tau=1. \tag{10.14b}$$

Hence on combining (10.14a, b) and substituting for A_{21} from (10.10), one obtains

$$E^{m(\tau)}_{2n+p}(K+iK')\sim 2i^{\sigma-\tau}(-1)^{n+\sigma+\tau}\mu_1\chi^{\tau}(1-k^2\alpha^2)^{\tau/2-1/4}. \tag{10.15}$$

The argument of the right hand side of (10.15) may be compared with (10.2b) and be seen to have the correct form.

In the region $[iK', K+iK']$, a Lamé polynomial is described by (8.4). As neither of the functions W_{31}, W_{32} may be zero or have a vanishing derivative at $z=K+iK'$, the boundary condition determines a relation between A_{31} and A_{32}. As z moves away from $K+iK'$ towards iK', W_{32} is exponentially increasing whilst W_{31} is exponentially decreasing. Obviously, away from $z=K+iK'$ in this region

the term $A_{32}W_{32}$ will dominate the term $A_{31}W_{31}$. At $z = K + iK'$,

$$W_{31}(K+iK') \sim W_{32}(K+iK') \sim (1-k^2\alpha^2)^{-1/4}, \tag{10.16a}$$

and

$$W'_{31}(K+iK') \sim -W'_{32}(K+iK') \sim \chi(1-\alpha^2k^2)^{1/4}. \tag{10.16b}$$

Thus

$$A_{31} \sim (-1)^{\tau}A_{32},$$

to satisfy the boundary condition at $z = K + iK'$, and

$$E^{m(\tau)}_{2n+p}(K+iK') \sim 2A_{32}\chi^{\tau}(1-k^2\alpha^2)^{\tau/2-1/4}(-1)^{\tau}.$$

Hence on comparing the above expression and (10.15)

$$A_{32} \sim (-1)^n i^{-(\tau+\sigma)}\mu_1 \tag{10.17}$$

and

$$E^m_{2n+p}(z) \sim (-1)^n i^{-(\tau+\sigma)}\mu_1\{k^2(\mathrm{sn}^2\, z - \alpha^2)\}^{-1/4} \cdot \exp\left\{-\chi k \int_{K+iK'}^{z} (\mathrm{sn}^2\, \xi - \alpha^2)^{1/2}\, d\xi\right\}. \tag{10.18}$$

It has been shown in § 8 that the error term in the Liouville–Green approximation is uniformly of $O(1/\chi)$ multiplied by the leading term, so that (10.18) will describe the behavior of a Lamé polynomial in the neighborhood of $z = iK'$. Equation (10.18) shows that the order of magnitude in the neighborhood of the pole and the phase of the Lamé polynomial agree with the value predicted by the form (1.2).

The remaining region on which the Lamé polynomials are to be identified is $[0, iK']$. The general form of the solution on this interval is given by (9.12). The parabolic cylinder functions occurring in this expression have nonnegative first argument so that neither W_{41} or W_{42} may be zero and both functions are monotonic. In order to satisfy the boundary condition at the origin one of the following relations must hold.

$$A_{41} \sim A_{42} \quad \text{if } \rho = 0,$$

or

$$A_{41} \sim A_{42} \quad \text{if } \rho = 1.$$

Hence

$$A_{41} \sim (-1)^{\rho}A_{42}. \tag{10.19}$$

As $\mathrm{Im}\,(z)$ increases along the line $\mathrm{Re}\,(z) = 0$, ζ_4 increases so that W_{41} is exponentially decreasing whilst W_{42} is exponentially increasing. Thus away from the origin the second term of (9.12) will dominate the first term, i.e.

$$E^m_{2n+p}(z) \sim A_{42}\left(\frac{dz}{d\zeta_4}\right)^{1/2} U(\tfrac{1}{2}\chi\nu_1^2, -\zeta_4\sqrt{2\chi}) \tag{10.20}$$

in this region. Replacing the parabolic cylinder function by the leading term of its asymptotic expansion, one finds that

$$E^m_{2n+p}(z) \sim A_{42}\left(\frac{dz}{d\zeta_4}\right)^{1/2}(2\pi)^{1/2}\{\Gamma(\tfrac{1}{2}+\tfrac{1}{2}\chi\nu_1^2)\}^{-1}2^{-\chi\nu_1^2/4-1/4}\chi^{\chi\nu_1^2/4-1/4}$$
$$\cdot e^{\chi\nu_1^2/4}(\zeta^2+\nu_1^2)^{-1/4}\nu_1^{\chi\nu_1^2/2}\exp\left\{\chi\int_0^{\zeta_4}(\tau^2+\nu_1^2)^{1/2}\,d\tau\right\}.$$

Replacing $(dz/d\zeta_4)$ and substituting for $\chi\nu_1^2$ and ζ_4, one obtains

$$\begin{aligned} E^m_{2n+p}(z) &\sim iA_{42}(2\pi)^{1/2}\{\Gamma(2m+\rho+1)\}^{-1} \\ &\cdot 2^{-m-\rho/2-1/2}\chi^{m+\rho/2}e^{-m-\rho/2-1/4}\nu_1^{2m+\rho+1/2} \\ &\cdot\{k^2(\mathrm{sn}^2\, z-\alpha^2)\}^{-1/4}\exp\left\{-\chi k\int_0^z(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi\right\}. \end{aligned} \tag{10.21}$$

On account of the double pole at $z = iK'$, Lamé polynomials are affected by Stokes' phenomenon in the neighborhood of this point. The relationship between solutions on $[0, iK')$ and $(iK', K+iK']$ is given by (10.1a, b), namely

$$E^m_{2n+p}(iK'-i\varepsilon) \sim E^m_{2n+p}(iK'+\varepsilon)\exp\{\tfrac{1}{2}i\pi(2n+p)\},$$

for small ε.

Thus on matching expressions (10.18) and (10.21), using the above relation, one obtains, after some reduction,

$$\begin{aligned} A_{42} \sim \frac{i^{\rho-1}\Gamma(2m+\rho+1)}{2^{m+3\rho/2+1/2}\chi^{\rho/2}\Gamma(m+\rho+\frac{1}{2})}\left(\frac{\nu_1}{k\alpha}\right)^{\rho-1/2}\exp\left\{-\chi k\int_{z_0}^{K}(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi\right. \\ \left.-\chi k\int_{K+iK'}^{iK'+\varepsilon}(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}.d\xi+\chi k\int_0^{iK'-i\varepsilon}(\mathrm{sn}^2\,\xi-\alpha^2)^{1/2}\,d\xi\right\}. \end{aligned} \tag{10.22}$$

Recalling (3.10), we see that the exponential term in the above disappears. Furthermore

$$\Gamma(2z) = \frac{2^{2z-1}}{\pi^{1/2}}\Gamma(z)\Gamma(z+\tfrac{1}{2}),$$

so that

$$\frac{\Gamma(2m+\rho+1)}{\Gamma(m+\rho+\frac{1}{2})} = \frac{\Gamma(m+1)2^{2m+\rho}}{\pi^{1/2}}$$

and

$$A_{42} \sim \frac{2^{m-1/2}e^{i\pi\rho/2-i\pi/4}\Gamma(m+1)}{(2\chi)^{\rho/2}\pi^{1/2}}\left(\frac{\nu_1}{k\alpha}\right)^{\rho-1/2}. \tag{10.23}$$

The behavior of the leading term of the Lamé polynomial for $z \in [0, iK']$ is now completely described by (10.19), (10.21) and (10.23). One may observe from these relations that the leading term of the Lamé polynomial does in fact satisfy

the normalization condition

$$E_{2n+p}^{m(\rho)}(0) = 1$$

which serves as a check on the results presented in this section.

11. Identification of Lamé polynomials for Case (b). The determination of the constants for this case follows the same pattern as in the previous section. The remarks made in that section concerning the arguments of Lamé polynomials and the behavior near the pole will hold also in this case.

For the region $[0, K]$, the behavior of Lamé polynomials is described by the expression (6.17). The parabolic cylinder functions occurring in this expression have nonnegative first argument, so that they are oscillatory in $[0, K)$. The second argument is zero when $z = K$, so in order to satisfy the boundary condition at this point, it is necessary that

$$A_{13} = (-1)^{\sigma} A_{14} l(\tfrac{1}{2}\chi\nu_2^2). \tag{11.1}$$

As $\zeta_2 = 0$, the two components of (6.17) are of the same order of magnitude. In the neighborhood of this point the solution W_{13} is exponentially decreasing whilst the solution W_{14} is exponentially increasing as ζ_2 increases. However as $\nu_2 \to 0$, the measure of the interval on which this type of behavior takes place is tending to zero. Thus for certain small values of ν_2, the behavior of the leading term of Lamé polynomials will be governed solely by the solution W_{14}, whilst for other small values both terms will be significant.

Away from $z = K$, asymptotic expansions may be used to represent the parabolic cylinder functions as the second argument of these functions is large. Thus

$$\begin{aligned} E_{2n+p}^{m}(z) = &\left(\frac{dz}{d\zeta_2}\right)^{1/2} A_{14} \frac{2^{1/4} l^{-1/2}(\tfrac{1}{2}\chi\nu_2^2)}{[\chi(\zeta_2^2 - \nu_2^2)]^{1/4}} \\ &\cdot [\cos y(\zeta_2) l(\tfrac{1}{2}\chi\nu_2^2)(-1)^{\sigma} + \sin y(\zeta_2) + O(\chi^{-1})], \end{aligned} \tag{11.2a}$$

where

$$y(\zeta_2) = \chi \int_{\nu_2}^{\zeta_2} (\tau^2 - \nu_2^2)^{1/2}\, d\tau + \phi(\tfrac{1}{2}\chi\nu_2^2) + \tfrac{1}{4}\chi\nu_2^2 - \tfrac{1}{4}\chi\nu_2^2 \ln \tfrac{1}{2}\chi\nu_2^2. \tag{11.2b}$$

The latter term inside the square brackets will be dominant if $\chi\nu_2^2$ is large, since

$$l(a) \sim \tfrac{1}{2} \exp\{-\pi a\} \quad \text{for large } a.$$

However if $\chi\nu_2^2$ is of $o(1)$ the two terms will be of the same order of magnitude. In order to express (11.2a) more concisely we introduce

$$\eta(a) = \tan^{-1} l(a),$$

such that $\eta(a) \in [0, \tfrac{1}{4}\pi]$, $\forall a$.

Now one may use the asymptotic representation (11.2a) to deduce that

$$E^{m}_{2n+p}(0) = \frac{iA_{14}}{(k\alpha)^{1/2}} \frac{2^{1/4} l^{-1/2}(\frac{1}{2}\chi\nu_2^2)}{\chi^{1/4}\cos\eta(\frac{1}{2}\chi\nu_2^2)} \cdot \sin[y(\hat{\zeta}_2) + (-1)^{\sigma}\eta(\tfrac{1}{2}\chi\nu_2^2) + O(\chi^{-1})] \tag{11.3a}$$

and

$$E^{m'}_{2n+p}(0) = -\frac{iA_{14}(k\alpha)^{1/2}2^{1/4}\chi^{3/4}}{l^{1/2}(\frac{1}{2}\chi\nu_2^2)\cos\eta(\frac{1}{2}\chi\nu_2^2)} \cdot \cos[y(\hat{\zeta}_2) + (-1)^{\sigma}\eta(\tfrac{1}{2}\chi\nu_2^2) + O(\chi^{-1})]. \tag{11.3b}$$

Thus the boundary condition at the origin may be satisfied if

$$\sin[y(\hat{\zeta}_2) + (-1)^{\sigma}\eta(\tfrac{1}{2}\chi\nu_2^2) + \tfrac{1}{2}\pi(1-\rho)] = O(\chi^{-1}). \tag{11.4}$$

This condition enables us to obtain a relation for the eigenvalue h of Lamé's equation, when the turning point is such that $\alpha \geqq \frac{1}{2}$. As the Lamé polynomial has m zeros in $(0, K)$, (11.4) may be written as

$$y(\hat{\zeta}_2) + (-1)^{\sigma}\eta(\tfrac{1}{2}\chi\nu_2^2) + \tfrac{1}{2}(1-\rho) = (m+1)\pi + O(\chi^{-1}).$$

Using (6.12b), one obtains

$$\chi k\int_0^{z_0}(\alpha^2 - \operatorname{sn}^2\xi)^{1/2}\,d\xi + \phi(\tfrac{1}{2}\chi\nu_2^2) + \tfrac{1}{4}\chi\nu_2^2 - \tfrac{1}{4}\chi\nu_2^2\ln\tfrac{1}{2}\chi\nu_2^2 + (-1)^{\sigma}\eta(\tfrac{1}{2}\chi\nu_2^2) = (m + \tfrac{1}{2}\rho + \tfrac{1}{2})\pi + O(\chi^{-1}). \tag{11.5}$$

The relation (11.5) is the eigenvalue condition which must be satisfied by the eigenvalue h of Lamé's equation.

As the argument of the sine function in (11.4) has an error term $O(1/\chi)$, the value of h obtained from (11.5) will have an error term of the same order. For particular cases depending on the size of $\frac{1}{2}\chi\nu_2^2$ it is possible to simplify (11.5) a little.

$$\chi k\int_0^{z_0}(\alpha^2 - \operatorname{sn}^2\xi)^{1/2}\,d\xi = (m + \tfrac{1}{2}\rho + \tfrac{1}{4})\pi + o(1) \quad \text{if } \tfrac{1}{2}\chi\nu_2^2 \gg 1, \tag{11.6}$$

$$\chi k\int_0^{z_0}(\alpha^2 - \operatorname{sn}^2\xi)^{1/2}\,d\xi = [m + \tfrac{1}{2}\rho + \tfrac{1}{4} - \tfrac{1}{8}(-1)\sigma]\pi + o(1) \quad \text{if } \tfrac{1}{2}\chi\nu_2^2 \ll 1. \tag{11.7}$$

Equations (11.6), (11.7) follow either from the asymptotic expansions for ϕ and η or from the limiting values as their arguments tend to zero, i.e.,

$$\phi(0) = \tfrac{1}{4}\pi, \qquad \eta(0) = \tfrac{1}{8}\pi.$$

These relations do not have the same order of accuracy as all other asymptotic formulas for Lamé polynomials. However, they are a little simpler than (11.5) and indicate the value of the eigenvalue h in two particular cases. One may also confirm that (11.6) agrees with (3.12).

The normalization condition implies that

$$A_{14} \sim \frac{(-1)^m l^{1/2}(\frac{1}{2}\chi\nu_2^2)\cos\eta(\frac{1}{2}\chi\nu_2^2)}{i(k\alpha)^{\rho-1/2}2^{1/4}\chi^{\rho-1/4}}. \tag{11.8}$$

The matching coefficient at $z = K$ may be obtained from the value of the parabolic cylinder function and its derivative at the origin and (11.8).

One finds that

$$E^{m(\sigma)}_{2n+p}(K) \sim \frac{(-1)^{m+\sigma}\chi^{\sigma/2}2^{\sigma}\nu_2^{1/2-\sigma}\chi^{\sigma/2+1/4-\rho}}{(k\alpha)^{\rho-1/2}[k(1-\alpha^2)^{1/2}]^{1/2-\sigma}} \cdot l^{1/2}(\tfrac{1}{2}\chi\nu_2^2)\cos\eta(\tfrac{1}{2}\chi\nu_2^2)\left|\frac{\Gamma(\frac{3}{4}+\frac{1}{4}i\chi\nu_2^2)}{\Gamma(\frac{1}{4}+\frac{1}{4}i\chi\nu_2^2)}\right|^{\sigma-1/2}. \tag{11.9}$$

The sign of the formula for this matching coefficient may be seen to be in agreement with the value predicted by (10.2a). The behavior of Lamé polynomials on the line $[0, K]$ has now been determined and we can now attempt to continue the solution into the interval $[K, K+iK']$. In this region expression (7.4) describes the behavior of Lamé polynomials. The parabolic cylinder functions occurring in this expression are nonzero when $\zeta_3 = 0$, so that the boundary condition at $\zeta_3 = 0$ determines a relation between A_{23} and A_{24} as follows:

$$A_{24} \sim (-1)^{\sigma} A_{23}. \tag{11.10}$$

In order to evaluate A_{23}, the value of the matching coefficient at $z = K$ is required. From the value of the parabolic cylinder function at the origin one obtains the following value for the matching coefficient:

$$E^{m(\sigma)}_{2n+p}(K) \sim 2^{\sigma+1/4}\left(\frac{k(1-\alpha^2)^{1/2}}{\nu_2}\right)^{\sigma-1/2} e^{i\pi/4+i\sigma\pi/2}\chi^{\sigma/2}A_{23} \cdot \left|\frac{\Gamma(\frac{3}{4}-\frac{1}{4}i\chi\nu_2^2)}{\Gamma(\frac{1}{4}-\frac{1}{4}i\chi\nu_2^2)}\right|^{\sigma-1/2}. \tag{11.11}$$

The multiplicative constant may be evaluated by matching the two expressions (11.9) and (11.11). This leads to

$$A_{23} \sim 2^{-1/4}(-1)^{m+\sigma} l^{1/2}(\tfrac{1}{2}\chi\nu_2^2)\cos\eta(\tfrac{1}{2}\chi\nu_2^2) \cdot e^{-(i\pi/4+i\sigma\pi/2)}\chi^{1/4-\rho}(k\alpha)^{1/2-\rho}. \tag{11.12}$$

Both of the solutions W_{23} and W_{24} are oscillatory for real ζ_3 and of the same order of magnitude. Hence the leading term of a Lamé polynomial will be dependent on both of the terms in (7.4). The leading term of a Lamé polynomial away from $z = K$ may be expressed as

$$E^{m}_{2n+p}(z) = \frac{2^{1/4}l^{-1/2}(-\frac{1}{2}\chi\nu_2^2)e^{i\pi/4}A_{23}(-1)^{\sigma}}{\chi^{1/4}k^{1/2}(\mathrm{sn}^2 z - \alpha^2)^{1/4}\cos\eta(-\frac{1}{2}\chi\nu_2^2)} \cdot \{\sin[v(\zeta_3)+(-1)^{\sigma}\eta(-\tfrac{1}{2}\chi\nu_2^2)] + O(\chi^{-1})\}, \tag{11.13}$$

where

$$v(\zeta_3)=\chi\int_0^{\zeta_3}(\tau^2+\nu_2^2)^{1/2}\,d\tau+\phi(-\tfrac{1}{2}\chi\nu_2^2)+\tfrac{1}{4}\chi\nu_2^2\ln\tfrac{1}{2}\chi\nu_2^2-\tfrac{1}{4}\chi\nu_2^2.$$

The boundary condition at $z=K+iK'$ will be satisfied if

$$E_{2n+p}^{m(1-\tau)}(K+iK')=0.$$

Thus the form of (11.13) implies that this condition may be satisfied provided that

$$\sin[v(\hat{\zeta}_3)+(-1)^\sigma\eta(-\tfrac{1}{2}\chi\nu_2^2)+\tfrac{1}{2}\pi(1-\tau)]=O(\chi^{-1}). \tag{11.14}$$

As the Lamé polynomial has $(n-m)$ zeros in the interval $(K, K+iK')$, the above condition may be expressed as

$$v(\hat{\zeta}_3)+(-1)^\sigma\eta(-\tfrac{1}{2}\chi\nu_2^2)+\tfrac{1}{2}\pi(1-\tau)=(n-m+1)\pi+O(\chi^{-1}). \tag{11.15}$$

Substituting for $v(\hat{\zeta}_3)$, the above equation becomes

$$\begin{aligned}\chi k\int_K^{K+iK'}(\alpha^2-\operatorname{sn}^2\xi)^{1/2}\,d\xi+\phi(-\tfrac{1}{2}\chi\nu_2^2)+\tfrac{1}{4}\chi\nu_2^2\ln\nu_2-\tfrac{1}{4}\chi\nu_2^2\\ +(-1)^\sigma\eta(-\tfrac{1}{2}\chi\nu_2^2)=(n-m+\tfrac{1}{2}+\tfrac{1}{2}\tau)\pi+O(\chi^{-1}).\end{aligned} \tag{11.16}$$

In order to show that this relation is consistent to our order of approximation with the eigenvalue condition (11.5), one may add the two equations together and recall that (3.11) implies that

$$k\int_0^{z_0}(\alpha^2-\operatorname{sn}^2\xi)^{1/2}\,d\xi+k\int_K^{K+iK'}(\alpha^2-\operatorname{sn}^2\xi)^{1/2}\,d\xi=\tfrac{1}{2}\pi.$$

The sum of (11.5) and (11.16) is thus

$$\begin{aligned}\tfrac{1}{2}\chi\pi+\phi(\tfrac{1}{2}\chi\nu_2^2)+\phi(-\tfrac{1}{2}\chi\nu_2^2)+(-1)^\sigma[\eta(\tfrac{1}{2}\chi\nu_2^2)+\eta(-\tfrac{1}{2}\chi\nu_2^2)]\\ =[n+1+\tfrac{1}{2}(\rho+\tau)]\pi+O(\chi^{-1}).\end{aligned} \tag{11.17}$$

Now

$$\phi(\tfrac{1}{2}\chi\nu_2^2)+\phi(-\tfrac{1}{2}\chi\nu_2^2)=\tfrac{1}{2}\pi,$$

and

$$\tfrac{1}{2}\chi\pi=[n+\tfrac{1}{2}(\rho+\sigma+\tau)+\tfrac{1}{4}]\pi$$

so that (11.17) will be true provided that

$$\begin{aligned}(-1)^\sigma[\eta(\tfrac{1}{2}\chi\nu_2^2)+\eta(-\tfrac{1}{2}\chi\nu_2^2)]&=\tfrac{1}{4}\pi-\tfrac{1}{2}\sigma\pi\\ &=\tfrac{1}{4}(-1)^\sigma\pi\end{aligned}$$

since σ is either zero or unity. Thus the equations will be consistent if

$$\eta(\tfrac{1}{2}\chi\nu_2^2)+\eta(-\tfrac{1}{2}\chi\nu_2^2)=\tfrac{1}{4}\pi.$$

Now for any value of a, one easily finds that

$$\eta(a)+\eta(-a)=\tfrac{1}{4}\pi.$$

Thus, as (11.5) and (11.16) are consistent, the eigenvalue condition imposed on h to enable Lamé's equation to admit a solution satisfying the boundary conditions on the real line is sufficient to ensure that the solution also satisfies the boundary condition at $z = K + iK'$.

In order to continue the asymptotic expansions around the fundamental rectangle the value of the matching coefficient at $z = K + iK'$ is required. From the formulas (11.12), (11.13) and (11.15) one sees that

$$E^{m(\tau)}_{2n+p}(K+iK') \sim \frac{(-1)^n (k\alpha)^{1/2-\rho} i^{\tau-\sigma} \chi^{\tau-\rho} (1-k^2\alpha^2)^{\tau/2-1/4} l^{1/2}(\frac{1}{2}\chi\nu_2^2) \cos \eta(\frac{1}{2}\chi\nu_2^2)}{l^{1/2}(-\frac{1}{2}\chi\nu_2^2) \cos \eta(-\frac{1}{2}\chi\nu_2^2)}. \tag{11.18}$$

The argument of the right hand side of (11.18) agrees with the value predicted by (10.2b).

The form of a Lamé polynomial on $[iK', K+iK']$ is described by (8.5). Neither of the solutions occurring in this expression may vanish or possess a vanishing derivative at $z = K + iK'$ so that the boundary condition will be satisfied if

$$A_{33} \sim (-1)^\tau A_{34}.$$

Hence the value of the matching coefficient is given by

$$E^{m(\tau)}_{2n+p}(K+iK') \sim 2(-1)^\tau A_{34} (1-k^2\alpha^2)^{\tau/2-1/4} \chi^\tau,$$

and comparison of this expression with (11.18) implies that

$$A_{34} \sim \frac{(-1)^n i^{-\tau-\sigma} (k\alpha)^{1/2-\rho} l^{1/2}(\frac{1}{2}\chi\nu_2^2) \cos \eta(\frac{1}{2}\chi\nu_2^2)}{2\chi^\rho l^{1/2}(-\frac{1}{2}\chi\nu_2^2) \cos \eta(-\frac{1}{2}\chi\nu_2^2)}. \tag{11.19}$$

As z moves from $K+iK'$ in the direction of iK' the solution $W_{32}(z)$ is exponentially increasing whilst $W_{31}(z)$ is exponentially decreasing. Consequently for $z \in (iK', K+iK')$, with z away from $K+iK'$,

$$E^m_{2n+p}(z) \sim A_{34} k^{-1/2} (\mathrm{sn}^2 z - \alpha^2)^{-1/4} \exp\left\{-\chi k \int_{K+iK'}^{z} (\mathrm{sn}^2 \xi - \alpha^2)^{1/2}\, d\xi\right\}. \tag{11.20}$$

Equation (11.20) shows that the order of magnitude in the neighborhood of the pole and the argument of the Lamé polynomial agree with the value predicted by the form (1.2).

On the interval $[0, iK']$, Liouville–Green approximations of the form (9.14) are available to describe the behavior of Lamé polynomials. The solutions $W_{43}(z)$ and $W_{44}(z)$ are neither zero nor possess vanishing derivatives at the origin so that the boundary condition determines a relation between A_{43} and A_{44}. This relation is given by

$$A_{43} \sim (-1)^\rho A_{44}. \tag{11.21}$$

As $W_{44}(z)$ is exponentially increasing and $W_{43}(z)$ is exponentially decreasing the general form of the solution on $[0, iK']$ away from the origin is given by

$$E^m_{2n+p}(z) \sim A_{44} k^{-1/2} (\alpha^2 - \mathrm{sn}^2 z)^{-1/4} \exp\left\{-i\chi k \int_0^z (\alpha^2 - \mathrm{sn}^2 \xi)^{1/2}\, d\xi\right\}. \tag{11.22}$$

In order to compute the value of A_{43}, the asymptotic representations on $[0, iK')$ and $(iK', K+iK']$ are matched using the relations (10.1), which result from the discontinuity due to Stokes' phenomenon. Hence from (11.20) and (11.21)

$$A_{44} \sim A_{34} \frac{(\alpha^2 - \operatorname{sn}^2\{iK' - i\varepsilon\})^{1/4}}{(\operatorname{sn}^2\{iK' + \varepsilon\} - \alpha^2)^{1/4}} \exp\left\{ i\chi k \int_0^{iK - i\varepsilon} (\alpha^2 - \operatorname{sn}^2 \xi)^{1/2}\, d\xi \right. \tag{11.23}$$
$$\left. - \chi k \int_{K+iK'}^{iK'+\varepsilon} (\alpha^2 - \operatorname{sn}^2 \xi)^{1/2}\, d\xi \right\} \exp\{in\pi + \tfrac{1}{2}\rho\pi\}.$$

for small values of ε.

Thus (cf. § 10)

$$A_{44} \sim \frac{l^{1/2}(\frac{1}{2}\chi\nu_2^2) \cos \eta(\frac{1}{2}\chi\nu_2^2) i^\rho}{2l^{1/2}(-\frac{1}{2}\chi\nu_2^2) \cos \eta(-\frac{1}{2}\chi\nu_2^2)(k\alpha)^{\rho - 1/2}\chi^\rho} e^{\chi\nu_2^2/4}. \tag{11.24}$$

Now, for all values of a,

$$\frac{l^{1/2}(a) \cos \eta(a)}{l^{1/2}(-a) \cos \eta(-a)} = e^{-\pi a/2}.$$

Thus (11.23) may be simplified to

$$A_{44} \sim \frac{i^\rho}{2(k\alpha)^{\rho - 1/2}\chi^\rho}. \tag{11.25}$$

On substituting this value for A_{44} and the corresponding value of A_{43} determined by (11.21), the value of the matching coefficient $E_{2n+p}^{m(\rho)}(0)$ is equal to unity at the origin. This is in fact the required value on account of the normalization condition and serves as a check on the results of this section.

12. Extension of results. The simplest extension of the results presented in §§ 6 to 11 is the continuation of the asymptotic representations of Lamé polynomials along the lines $\operatorname{Re}(z) = 0, K$ and $\operatorname{Im}(z) = 0, K'$. That is, one may continue the expansions defined on the perimeter of the fundamental rectangle. Periodicity and parity considerations enable one to extend the results along the entire length of these lines, e.g.

$$sE_{2n+1}^m(z) = -sE_{2n+1}^m(-z) \quad \text{for } z \in [-K, 0],$$
$$scE_{2n+2}^m(K + iK' + ix) = scE_{2n+2}^m(K + iK' - ix) \quad \text{for } x \in [0, K'],$$

and

$$uE_{2n}^m(z + 2rK) = uE_{2n}^m(z) \quad \text{for } z \in [0, K]$$

and some integer r.

The results may also be extended to any line of the lattice L defined by

$$L = \{z : \operatorname{Re}(z) \equiv 0 (\operatorname{mod} K)\} \cup \{z : \operatorname{Im}(z) \equiv 0 (\operatorname{mod} K')\}.$$

In practice one is not usually concerned with Lamé polynomials whose argument lies outside the fundamental rectangle. However, values of the argument lying outside the fundamental rectangle occur in the final extension of the

results. We have on several occasions referred to the possibility of deriving results, for the case in which Lamé polynomials have a turning point on the line Re $(z) = K$, by an application of the Jacobi imaginary transformation. Thus we list in full the relation between Lamé polynomials, which is described schematically by (1.5b). Suppose that

$$z = K + iK' - i\xi; \tag{12.1}$$

then

$$\begin{aligned}
uE_{2n}^{m} &= AuE_{2n}^{n-m}(\xi, k'),\\
cE_{2n+1}^{m}(z, k) &= AcE_{2n+1}^{n-m}(\xi, k'),\\
sdE_{2n+2}^{m}(z, k) &= AsdE_{2n+2}^{n-m}(\xi, k'),\\
cdE_{2n+3}^{m}(z, k) &= AscdE_{2n+3}^{n-m}(\xi, k'),\\
sE_{2n+1}^{m}(z, k) &= AdE_{2n+1}^{n-m}(\xi, k'),\\
dE_{2n+1}^{m}(z, k) &= AsE_{2n+1}^{n-m}(\xi, k'),\\
scE_{2n+2}^{m}(z, k) &= AcdE_{2n+2}^{n-m}(\xi, k'),\\
cdE_{2n+2}^{m}(z, k) &= AscE_{2n+2}^{n-m}(\xi, k'),
\end{aligned} \tag{12.2a–h}$$

where A denotes a generic constant.

The following regions correspond under this transformation:

(i) $z \in [0, K]$, $\xi \in [K' - iK, K']$,

(ii) $z \in [K, K + iK']$, $\xi \in [0, K']$,

(iii) $z \in [iK', K + iK']$, $\xi \in [-iK, 0]$,

(iv) $z \in [0, iK]$, $\xi \in [-iK', K' - iK]$.

Thus if the original Lamé polynomial, on the left hand side of equations (12.2), has a turning point in $[K, K + iK']$, the corresponding Lamé polynomial with complementary modulus will have a turning point in the interval $[0, K']$. Each turning point in $[K, K + iK']$ will correspond to one of the Cases (a) or (b) of § 6. For a Lamé polynomial with turning point in $[K, K + iK']$ the corresponding Lamé polynomial on the right of (12.2) will be Case (a) or (b) according as,

(a) if $\frac{1}{2}(1 + k^{-1}) \leqq \alpha \leqq k^{-1}$,

(b) if $1 \leqq \alpha \leqq \frac{1}{2}(1 + k^{-1})$.

The remaining problem for a Lamé polynomial having a turning point in $[K, K + iK']$ is the determination of the constant A. On account of the relation between z and ξ in (12.1), it is clear that

$$A = (-i)^{\tau}/E_{2n+p}^{n-m(\tau)}(K' - iK, k'). \tag{12.3}$$

With this condition each type of polynomial will satisfy the normalization condition

$$E_{2n+p}^{m(\rho)}(z) = 1.$$

The correct form for $E_{2n+p}^{n-m(\tau)}(K'-iK, k')$ will be found by determining which case either (a) or (b) applies and by substituting $n-m$ for m and k' for k in the relevant formula for $E_{2n+p}^{m(\tau)}(K+iK', k)$ and observing that

$$E_{2n+p}^{n-m(\tau)}(K'-iK, k') = (-1)^{\sigma} E_{2n+p}^{n-m(\tau)}(K'+iK, k').$$

13. Special cases. When two turning points of Lamé's equation are very close to the origin, it is possible to simplify some of the expressions occurring in § 10. The turning points must be sufficiently close to the origin to ensure that $\chi\nu_1^2$ is $O(1)$. One may then use the asymptotic formulas of [13] to approximate the parabolic cylinder functions, the error being $O(1/\chi)$. The results which are obtained by this method correspond to those calculated by Ince [7]. The advantage of the approach used here is that the results are uniformly valid on the interval $[0, K]$, whilst those of [7] were based on the Liouville–Green method and consequently were not valid in the neighborhood of the turning point.

Similarly the second special case mentioned in § 2 has also been covered by the results given here. In this case Lamé's equation has a turning point close to $z = K+iK'$, and the corresponding Lamé polynomial may be evaluated from the results of §§ 10 and 12. Once again the asymptotic formulas of [13] may be used and the results reduced to those which are obtained by the Liouville–Green approach.

The other special case mentioned in § 2 is that of a Lamé polynomial with a turning point at z_0 very close to $z = K$. Such Lamé polynomials were discussed in § 11, where an eigenvalue condition and asymptotic solutions were given. In that section a simplified form of the eigenvalue condition was given for the case in which $\chi\nu_2^2 \ll 1$. This condition (11.7) is

$$\chi k \int_0^{z_0} (\alpha^2 - \operatorname{sn}^2 \xi)^{1/2}\, d\xi = [m + \tfrac{1}{2}\rho + \tfrac{1}{4} - \tfrac{1}{8}(-1)^{\sigma}]\pi + o(1). \tag{13.1}$$

For the special case, $m = n/2$, $k^2 = \frac{1}{2}$, $\rho = \tau$ the parameter ν_2 is given by (6.11).

$$\chi\nu_2^2 = \frac{4k\chi}{\pi} \int_{z_0}^{K} \left(\frac{1}{4\chi^2} - \operatorname{cn}^2 \xi\right)^{1/2} d\xi$$

since $\alpha^2 k^2 \chi^2 = h = k^2(\chi^2 - \frac{1}{4})$.

Now z_0 is such that

$$\operatorname{cn}^2 z_0 = \frac{1}{4\chi^2},$$

so that $K - z_0 = O(1/\chi)$.

As the integrand is $O(1/\chi)$ one may deduce that $\chi\nu_2^2$ is in fact $o(1)$ and that (13.1) holds. The integral on the left hand side is for large values of χ given by

$$\begin{aligned} \chi k \int_0^{z_0} (\alpha^2 - \operatorname{sn}^2 \zeta)^{1/2}\, d\zeta &= \chi \int_0^{K} \operatorname{cn} \zeta\, d\zeta + o(1) \\ &\sim \chi \sin^{-1} k \sim \tfrac{1}{4}\chi\pi \\ &\sim \tfrac{1}{2}(n + \tfrac{1}{2}(\rho+\tau) + \tfrac{1}{2}\sigma + \tfrac{1}{4})\pi. \end{aligned}$$

Now as $\rho = \tau$, and $\frac{1}{2}\sigma = \frac{1}{4}(-1)^{\sigma} + \frac{1}{4}$ for $\sigma = 0$ or 1, it follows that

$$\chi k \int_0^{z_0} (\alpha^2 - \mathrm{sn}^2\,\xi)^{1/2}\, d\xi = [m + \tfrac{1}{2}\rho + \tfrac{1}{8}(-1)^{\rho} + \tfrac{1}{4}]\pi,$$

as predicted by the condition (11.7).

For certain values of α, z_0 will actually be located at $z = K$ and the parameter ν_2 will vanish. In this case the parabolic cylinder functions will have first argument equal to zero, which means that the Lamé polynomial may be expressed in terms of Bessel functions of order $\frac{1}{4}$ [13].

One would expect that such Lamé polynomials could be approximated by Bessel functions of order $\frac{1}{4}$ as Lamé's equation has a double turning point at $z = K$ in this case.

There is a difference between the coalescing cases at the origin or $z = K + iK'$ and $z = K$ in that ν_1 can never be equal to zero for Lamé polynomials as

$$\tfrac{1}{2}\chi\nu_1^2 = 4m + 2\rho + 1 + O(\chi^{-1}),$$

whilst ν_2 may well be zero. Thus, although the turning points at z_0 may be arbitrarily close together for large χ they cannot coalesce to form a double turning point, whilst those at z_0, $2K - z_0$ may coalesce to form such a turning point.

14. Conclusions. Asymptotic approximations which take the form of an approximating standard function and an error term, for which realistic bounds have been obtained, have been constructed for the solution of Lamé's equation. On certain occasions the approximating functions were the Liouville–Green functions, i.e., exponential functions or trigonometric functions, whilst on other occasions the approximating functions were parabolic cylinder functions. In the latter case we have achieved a uniform reduction from the three free variables h, n and z to the two variables of the parabolic cylinder function. These asymptotic approximations are uniformly valid on the fundamental rectangle, provided that the moduli of the Jacobian elliptic functions occurring are neither close to zero nor close to unity. If either of these latter conditions were to hold a different approach would be preferable. It would probably be most useful to consider the equation as a perturbation of the appropriate degenerate form of Lamé's equation.

In §§ 10 and 11, it was possible to construct the leading term of the uniform asymptotic approximation of Lamé polynomials based on the results obtained earlier for the solutions of Lamé's equation. In all cases we were able to conclude that the error term was uniformly small, $O(1/\chi)$, with respect to the leading term, except possibly in the neighborhood of the zeros of Lamé polynomials. The asymptotic representations are uniformly valid on the fundamental rectangle and may be extended to the lattice L defined in § 12, based on the elliptic integrals K, K', by parity and periodicity arguments.

In several formulas functions occur which are not doubly periodic. However as the solutions on L are constructed by parity and periodicity arguments, all asymptotic representations on any line of L are periodic on that line. Thus the asymptotic representations of Lamé polynomials, which are matched solutions at the intersections of the lines of L, are doubly periodic.

It may appear that something has been lost from Ince's approach [7], since he was able to eliminate all singly periodic terms. However the functions occurring when $m = O(1)$ could be integrated analytically, whilst here this is not always the case. Thus the periodic and nonperiodic terms cannot be treated in a simpler manner, and we must rely on the interpretation mentioned above to ensure that these functions are doubly periodic. As we have not appealed to the periodic nature of the solutions when constructing Lamé polynomials on the fundamental rectangle, the method used here is applicable to more general problems.

Formulas for the determination of eigenvalues have been established for all cases and a simple numerical algorithm may be used to compute the eigenvalues which are required.

Except for the special cases of § 2, all the results mentioned above are new. Even in the special cases we have been able to improve on existing results.

Several points, which merit further attention, have arisen in the development. Simpler differential equations such as Mathieu's equation and the spheroidal wave equation require similar analysis involving a uniform reduction of free variables. In the case of the former equation it is also a degenerate case of Lamé's equation and the study of the solutions of Lamé's equation as it degenerates into Mathieu's equation would be of interest, particularly if the solutions of Lamé's equation are the eigenfunctions occurring in the delta wing problem [6]. The solutions in this case would correspond to the eigenfunctions for a slender wing. Using similar methods to those used here, it should also be possible to obtain asymptotic representations for the eigenfunctions of this problem.

If only eigenvalues of the problem are required methods based on the modified Prüfer transformation appear to be of interest in the case in which the differential equation possesses turning points. Here we were able to overcome the problem either by restricting attention to the interval which was free from turning points or by obtaining an eigenvalue condition from the solution of the differential equation. However there are sufficiently well developed methods for treating asymptotic integrals with a pole in the integrand, e.g. [3], and these may form a useful approach to this problem.

Acknowledgment. The authors are extremely grateful to Professor F. W. J. Olver for his advice and criticisms during the preparation of this article.

REFERENCES

[1] F. M. ARSCOTT, *Periodic Differential Equations*, Pergamon Press, Oxford, 1964.

[2] G. BIRKHOFF AND G.-C. ROTA, *Ordinary Differential Equations*, Ginn, Boston, 1962.

[3] N. BLEISTEIN, *Uniform asymptotic expansions of integrals with stationary point near algebraic singularity*, Comm. Pure Appl. Math., 19 (1966), pp. 353–370.

[4] A. ERDÉLYI, *On Lamé functions*, Philos. Mag., 31 (1941), pp. 123–130.

[5] B. A. HARGRAVE, *Numerical approximation of eigenvalues of Sturm–Liouville systems*, J. Comput. Phys., 20 (1976), pp. 381–396.

[6] B. A. HARGRAVE AND B. D. SLEEMAN, *The numerical solution of two-parameter eigenvalue problems in ordinary differential equations with an application to the problem of diffraction by a plane angular sector*, J. Inst. Math. Appl., 14 (1974), pp. 9–22.

[7] I. L. INCE, *Periodic Lamé functions*, Proc. Roy. Soc. Edinburgh, 60 (1940), pp. 447–463.

[8] J. C. P. MILLER, *On the choice of standard solutions for a homogeneous linear differential equation of the second order*, Quart. J. Mech. Appl. Math., 3 (1950), pp. 225–235.
[9] ———, *Tables of Weber Parabolic Cylinder Functions*, Her Majesty's Stationery Office, London, 1955.
[10] ———, *Parabolic cylinder functions*, Handbook of Mathematical Functions, M. Abramowitz and I. A. Stegun, eds., National Bureau of Standards Appl. Math. Ser. No. 55, U.S. Government Printing Office, Washington, D.C., pp. 685–720.
[11] F. W. J. OLVER, *Uniform asymptotic expansions for Weber parabolic cylinder functions of large orders*, J. Res. Nat. Bur. Standards, 63B (1959), pp. 131–169.
[12] ———, *Asymptotics and Special Functions*, Academic Press, New York, 1974.
[13] ———, *Second-order linear differential equations with two turning points*, Philos. Trans. Roy. Soc. London Ser. A, 278 (1975), pp. 137–174.

SIAM J. MATH. ANAL.
Vol. 8, No. 5, October 1977

A TRANSVERSELY ISOTROPIC ELLIPTIC EQUATION: AN ARR ANALOGY APPROACH*

LIM CHEE-SENG†

Abstract. A "specially" elliptic equation is posed for a singular point source within a transversely isotropic medium. This occupies an infinite space with an odd number of dimensions equaling or exceeding three. By appropriately transforming the given equation, one can deduce an implicit solution anywhere off a source plane. This is done through an analogy with the time-dependent ARR problem [4] involving an even number of spatial dimensions. Explicit results can be established, as well as eigenfunction behavior near the symmetry axis. Principles are illustrated with an application to an elastodynamic problem for a propagating concentrated force.

1. Introduction. Let position $\mathbf{x}=(x_1, x_2, \cdots, x_n)\in R_n$, the infinite n-dimensional space. Consider the differential equation

$$G(\partial/\partial x_1, \nabla_1^2)\phi = F(\partial/\partial x_1, \nabla_1^2)\delta(\mathbf{x}), \tag{1.1}$$

where $\delta(\mathbf{x})$ denotes the Dirac delta(point source) function in R_n. Being dependent on $\partial/\partial x_1$ and the $(n-1)$-dimensional Laplacian $\nabla_1^2 \equiv \partial^2/\partial x_2^2+\cdots+\partial^2/\partial x_n^2$, both F- and G-(polynomial) operators are said to be transversely isotropic. Suppose they correspond to a real, originally homogeneous medium which is nondispersive either permanently or during an eventual steady state, i.e., they possess real constant coefficients and are homogeneous in $\partial/\partial x_1, \partial/\partial x_2, \cdots, \partial/\partial x_n$. So, generally,

$$F(\partial/\partial x_1, \nabla_1^2) = \sum_{\mu=0}^{(1/2)(m-p)} A_\mu \nabla_1^{2\mu}(\partial/\partial x_1)^{m-l-2\mu}, \tag{1.2}$$

$$G(\partial/\partial x_1, \nabla_1^2) = \sum_{\mu=0}^{(1/2)(m-q)} B_\mu \nabla_1^{2\mu}(\partial/\partial x_1)^{m-2\mu}, \tag{1.3}$$

all A_μ's and B_μ's being real constants and $1\leqq l\leqq m$. Also, assuming that m is even, while $A_{(1/2)(m-p)}\neq 0$ and $B_{(1/2)(m-q)}\neq 0$, the Laplacian indices p and q (necessarily even integers), together with l, must satisfy $l\leqq p\leqq m$, $0\leqq q<m$. The case $q=m$ is trivial. Throughout this paper, we restrict n to be *odd* and $\geqq 3$, and G to be elliptic in a special sense. In particular, (1.1) may have evolved from a matrix differential equation. Actual nondispersive, transversely isotropic systems are encountered in magnetogasdynamics, magnetoelasticity, elasticity of certain crystalline media, uniaxial crystal optics, compressible flow theory, source propagation within an elastic medium.

Physically, (1.1) may be envisaged as the steady state development, either after some large positive time t, or ultimately as $t\to\infty$, of the following radiation problem:

$$Q(\partial/\partial t, \partial/\partial x_1, \nabla_1^2)\phi = P(\partial/\partial t, \partial/\partial x_1, \nabla_1^2)\delta(\mathbf{x})H(t), \tag{1.4}$$

$$\phi = 0 \quad \text{during } t<0, \tag{1.5}$$

* Received by the editors September 14, 1975, and in revised form December 17, 1975.

† Institute of Geophysics and Planetary Physics, University of California, Los Angeles, California. On leave from Department of Mathematics, University of Malaya. Now at Department of Mathematics, University of Singapore, Singapore. This appears as Publication no. 1547, Institute of Geophysics and Planetary Physics, UCLA.

where $H(t)$ denotes the Heaviside unit function, whilst P and Q are homogeneous in $\partial/\partial t, \partial/\partial x_1, \cdots, \partial/\partial x_n$, and possess real constant coefficients. Postulating Q as being "hyperbolic cum elliptic" (i.e., originally hyperbolic, but eventually elliptic during the steady state), the present author has studied such a problem for n odd and $\geqq 3$ [3], as well as n even and $\geqq 2$ [4]; but solutions hold only for axial radiation reception (ARR), i.e., with observer's position $\mathbf{x}$ constrained along the x_1-axis, which is one of symmetry.

However, the ARR technique for the even n radiation problem comprising (1.4) and (1.5) can be adapted to the present odd n configuration governed by (1.1) and moreover for any $\mathbf{x}$ position noncoincident with the delta source point at $\mathbf{0}$. This is largely because (1.1) does not involve $\partial/\partial t$. An obvious relationship between F, G, P and Q is

$$F(\partial/\partial x_1, \nabla_1^2) = P(0, \partial/\partial x_1, \nabla_1^2), \qquad G(\partial/\partial x_1, \nabla_1^2) = Q(0, \partial/\partial x_1, \nabla_1^2). \tag{1.6}$$

2. Inversion formalities. Suppose, within R_{n-1}, the typical position $\mathbf{r} = (x_2, \cdots, x_n)$. Then Fourier inversion of (1.1) yields

$$\phi = \frac{i^{-l}}{(2\pi)^n} \int_{R_{n-1}} \exp(i\boldsymbol{\kappa} \cdot \mathbf{r})\, d\boldsymbol{\kappa} \int_{-\infty}^{\infty} \frac{F(\alpha, \kappa^2)}{G(\alpha, \kappa^2)} \exp(i\alpha x_1)\, d\alpha, \tag{2.1}$$

where $\boldsymbol{\kappa} \in R_{n-1}$ and $\kappa = |\boldsymbol{\kappa}|$. We have incorporated the homogeneity effects of (1.2) and (1.3), viz., that for any real or complex β:

$$F(\beta\alpha, \beta^2\kappa^2) \equiv \beta^{m-l} F(\alpha, \kappa^2), \qquad G(\beta\alpha, \beta^2\kappa^2) \equiv \beta^m G(\alpha, \kappa^2). \tag{2.2}$$

Ellipticity of the G-operator normally requires the nonvanishing of (see, e.g., [6])

$$G(\xi_1, \xi_2^2 + \cdots + \xi_n^2) \equiv G(\xi_1, 1 - \xi_1^2)$$

$\forall\, \boldsymbol{\xi} = (\xi_1, \xi_2, \cdots, \xi_n) \in \Omega_n$, the unit sphere (circle if $n = 2$) in R_n. (Note: we may take $\xi_1 = \alpha(\alpha^2 + \kappa^2)^{(-1/2)}$.) However, it is easily seen from (1.3) that, unless $q = 0$, vanishing does occur along every $\boldsymbol{\xi}$-direction orthogonal to $(1, 0, \cdots, 0)$, i.e., when $\xi_1 = 0$. Instead, we impose a *special ellipticity* on G, viz., (cf. [3], [4])

$$\xi_1^{-q} G(\xi_1, 1 - \xi_1^2) \neq 0 \quad \forall\, \boldsymbol{\xi} \in \Omega_n \quad (\text{i.e., } \forall\, \xi_1 \in [-1, 1]), \tag{2.3}$$

which obviously holds at $\xi_1 = 0$ since then the left side takes the nonzero value $B_{(1/2)(m-q)}$. At $\xi_1 = \pm 1$, (2.3) implies $B_0 \neq 0$.

In the ARR problem, the additional partial zero condition (1.5) plus the hyperbolic-cum-elliptic condition on the Q-operator are accommodated by contour integration during inversion of the Fourier transform with respect to time. The corresponding situation never arises with (2.1), for which no contour integration need be performed.

Regarding the integrand in (2.1), its factor

$$F(\alpha, \kappa^2)/G(\alpha, \kappa^2) = \alpha^{p-q-l}[\alpha^{l-p} F(\alpha, \kappa^2)]/[\alpha^{-q} G(\alpha, \kappa^2)].$$

From (1.2), $\alpha^{l-p} F(\alpha, \kappa^2)$ is a polynomial of degree $m - p$ in κ and of degree less than or equal to $m - p$ in α. From (1.3),

$$\alpha^{-q} G(\alpha, \kappa^2) \equiv (\alpha^2 + \kappa^2)^{(1/2)(m-q)} \xi_1^{-q} G(\xi_1, 1 - \xi_1^2),$$

a polynomial of degree $m-q$ in both α and κ which clearly never vanishes whenever $\alpha \in (-\infty, \infty)$, $\kappa \in (0, \infty)$, $\boldsymbol{\xi} \in \Omega_n$. Then for

$$p \geqq q + l, \tag{2.4}$$

the infinite integral in (2.1) is nonsingular and convergent. In contrast, Chee-Seng's ([3], [4]) Fourier integral representations of the dependent ϕ to (1.4) are, in a sense, singular but convergent.

Concerning (2.1), the $\boldsymbol{\kappa}$-integral over R_{n-1} can be expressed as the combination of a line integral over $0 < \kappa < \infty$ and a spherical integral over Ω_{n-1}. The latter can be reduced by the method of spherical means [6]. The former can then be converted into a line integral over $(-\infty, \infty)$; whereupon, we arrive at

$$\begin{aligned} \phi = {} & \frac{i^{-l}\omega_{n-2}}{2(2\pi)^n} \int_{-1}^{1} (1-\xi^2)^{(1/2)n-2}\, d\xi \\ & \cdot \int_{-\infty}^{\infty} (\operatorname{sgn} \kappa) \exp(i\kappa\xi r)\kappa^{n-2}\, d\kappa \int_{-\infty}^{\infty} \frac{F(\alpha, \kappa^2)}{G(\alpha, \kappa^2)} \exp(i\alpha x_1)\, d\alpha, \end{aligned} \tag{2.5}$$

where $\omega_n = 2\pi^{(1/2)n}/\Gamma(\frac{1}{2}n)$, the surface area of Ω_n. Observe an axisymmetry about the x_1-axis. Hereafter, we confine our attention off the plane $x_1 = 0$. By substituting the α-integration variable and appealing to (2.2), we find

$$\int_{-\infty}^{\infty} \cdots d\alpha = (-1)^l |x_1|^{-1} \kappa^{1-l} (\operatorname{sgn} \kappa) \int_{-\infty}^{\infty} \frac{F(\alpha x_1^{-1}, 1)}{G(\alpha x_1^{-1}, 1)} \exp(-i\kappa\alpha)\, d\alpha.$$

Consequently, noting that by the one-dimensional Fourier transformation-inversion

$$\int_{-\infty}^{\infty} \exp(i\kappa\xi r)\, d\kappa \int_{-\infty}^{\infty} \frac{F(\alpha x_1^{-1}, 1)}{G(\alpha x_1^{-1}, 1)} \exp(-i\kappa\alpha)\, d\alpha = 2\pi \frac{F(\xi\chi, 1)}{G(\xi\chi, 1)}$$

with $\chi = r x_1^{-1}$, we obtain

$$\phi = \frac{(-1)^{l+1}\omega_{n-2}}{2(2\pi)^{n-2}} \left(\frac{\partial}{\partial r}\right)^{n-l-1} \frac{K(\chi)}{|x_1|} \quad \text{if } n \geqq l+1, \tag{2.6}$$

$$\left(\frac{\partial}{\partial r}\right)^{l+1-n} \phi = \frac{(-1)^{l+1}\omega_{n-2}}{2(2\pi)^{n-2}} \frac{K(\chi)}{|x_1|} \quad \text{if } n \leqq l+1, \tag{2.7}$$

where

$$K(\chi) = \frac{(-1)^{(1/2)(n-3)}}{2\pi} \int_{-1}^{1} \frac{(1-\xi^2)^{(1/2)n-2} F(\xi\chi, 1)}{\xi^{n-l-1} G(\xi\chi, 1)}\, d\xi. \tag{2.8}$$

3. Application of the analogy. Following Chee-Seng [4], we transform the ξ-integration variable in (2.8):

$$\zeta = \xi^{-1}(1-\xi^2)^{1/2}. \tag{3.1}$$

Then via (2.2), (2.8) becomes

$$K(\chi) = \frac{(-1)^{(1/2)(n-3)}}{2\pi} \int_{-\infty}^{\infty} \zeta^{n-3} \frac{F(\chi, 1+\zeta^2)}{G(\chi, 1+\zeta^2)}\, d\zeta. \tag{3.2}$$

Now, the even n solution to (1.4) and (1.5) is first related to a similar integral, viz., ([4], (3.4))

$$\frac{(-1)^{(1/2)n-1}}{2\pi}\int_{-\infty}^{\infty}\zeta^{n-2}\frac{P(-x_1/t,1,\zeta^2)}{Q(-x_1/t,1,\zeta^2)}d\zeta,$$

precisely, an extension of the form (3.2) with n replacing $n-1$ and $-x_1/t$ replacing χ. Its derivation and subsequent evaluation develop from certain concepts put forward by Weitzner [8], Burridge [2], Bazer and Yen [1] and Payton [7]. It is convergent under a certain inequality relating n to the Laplacian indices of both P- and Q-operators in (1.4), and is evaluated in the sense of a principal value (see [4], (3.14), (3.15), (3.19)).

The similarity provides an analogy with the ARR result. By virtue of this analogy, we can now deduce under the convergence inequality

$$p\geqq q+n-1, \tag{3.3}$$

that

$$K(\chi)=\sum_{0<\arg\zeta_\nu<(1/2)\pi}\Phi_\nu(\chi)+\sum_{\arg\zeta_\nu=(1/2)\pi}\Psi_\nu(\chi), \tag{3.4}$$

where $\zeta_\nu=\zeta_\nu(\chi)$ denotes a typical ζ-root to

$$G(\chi,1+\zeta^2)=0, \tag{3.5}$$

$\sum_{0<\arg\zeta_\nu<(1/2)\pi}$ and $\sum_{\arg\zeta_\nu=(1/2)\pi}$ range over, respectively, every complex ζ-root in the first quadrant $0<\arg\zeta<\frac{1}{2}\pi$ and every purely imaginary ζ-root in the upper half-plane $\operatorname{Im}\zeta>0$. Furthermore, corresponding to a first quadrant root ζ_ν of order m_ν we have

$$\Phi_\nu(\chi)=\frac{2(-1)^{(1/2)(n-1)}}{(m_\nu-1)!}\operatorname{Im}\left\{\lim_{\zeta\to\zeta_\nu(\chi)}\left(\frac{\partial}{\partial\zeta}\right)^{m_\nu-1}\cdot\left[(\zeta-\zeta_\nu(\chi))^{m_\nu}\zeta^{n-3}\frac{F(\chi,1+\zeta^2)}{G(\chi,1+\zeta^2)}\right]\right\}, \tag{3.6}$$

whilst corresponding to a purely imaginary root ζ_ν of order m_ν in $\operatorname{Im}\zeta>0$, we have

$$\Psi_\nu(\chi)=\frac{-1}{(m_\nu-1)!}\lim_{\zeta\to|\zeta_\nu(\chi)|}\left(\frac{\partial}{\partial\zeta}\right)^{m_\nu-1}\left[(\zeta-|\zeta_\nu(\chi)|)^{m_\nu}\zeta^{n-3}\frac{F(\chi,1-\zeta^2)}{G(\chi,1-\zeta^2)}\right]. \tag{3.7}$$

Our solution is now complete.

Remarks. On the ARR theory, (3.5) should not possess any real repeated ζ-root; furthermore, in case simple real roots occur, their contributions cancel out. We shall later see that such roots actually never arise.

If $n\geqq l+1$, as for (2.6), then postulation of (3.3) automatically implies the satisfaction of (2.4); alternatively, if $n\leqq l+1$, as for (2.7), then we need only impose (2.4), from which (3.3) automatically follows.

4. Explicit dependences. The forms (3.6) and (3.7) are implicitly dependent on χ. Before any explicit dependence can be established, that of $\zeta_\nu(\chi)$ must first be found. Now, defining

$$g(\zeta^2) \equiv G(1, \zeta^2), \tag{4.1}$$

then via (2.2), we find

$$G(\xi, 1-\xi^2) \equiv \xi^m g(\zeta^2) \tag{4.2}$$

under the mapping of (3.1). The infinite real ζ-range $(-\infty, \infty)$ is the image of the combined ξ-interval of $[-1, 0) \cup (0, 1]$, over which $G(\xi, 1-\xi^2)$ never vanishes by virtue of the ellipticity condition (2.3). Hence $g(\zeta^2)$ has no real zeros. Also, from (1.3), it is a polynomial of degree $\frac{1}{2}(m-q)$ in ζ^2 and can thereby be represented by

$$g(\zeta^2) \equiv B_{(1/2)(m-q)} \prod_{\nu=1}^{(1/2)(m-q)} \zeta^2 - \lambda_\nu^2, \tag{4.3}$$

λ_ν, $-\lambda_\nu$ $(\nu = 1, \cdots, \frac{1}{2}m - \frac{1}{2}q)$ being its $\frac{1}{2}(m-q)$ pairs of symmetrical zeros, all essentially lying off the Re ζ axis. The λ_ν's from more than one pair may coincide. From (2.2) and (4.3), we have

$$G(\chi, 1+\zeta^2) \equiv \chi^m g\left(\frac{1+\zeta^2}{\chi^2}\right) \equiv \chi^m B_{(1/2)(m-q)} \prod_{\nu=1}^{(1/2)(m-q)} \frac{\zeta^2 - \chi^2\lambda_\nu^2 + 1}{\chi^2}; \tag{4.4}$$

also,

$$F(\chi, 1+\zeta^2) \equiv \chi^{m-l} f\left(\frac{1+\zeta^2}{\chi^2}\right), \tag{4.5}$$

wherein

$$f(\zeta^2) \equiv F(1, \zeta^2). \tag{4.6}$$

Evidently, the ζ-roots to (3.5) can now be explicitly identified as satisfying

$$\zeta_\nu^2 = \chi^2\lambda_\nu^2 - 1, \qquad \nu = 1, \cdots, \tfrac{1}{2}(m-q). \tag{4.7}$$

Suppose λ_ν is a purely imaginary zero, with order m_ν, of $g(\zeta^2)$. Whence, expression (4.3), and likewise (4.4), contains exactly m_ν identical factors involving this specific λ_ν. Consequently, (3.5) has a pair of symmetrical, purely imaginary ζ-roots, each of order m_ν, at

$$i|\zeta_\nu(\chi)|, \ -i|\zeta_\nu(\chi)| \quad \text{with } |\zeta_\nu(\chi)| \equiv (\chi^2|\lambda_\nu|^2 + 1)^{1/2}. \tag{4.8}$$

Only the upper root $i|\zeta_\nu(\chi)|$ contributes a $\Psi_\nu(\chi)$ term represented by (3.7).

Since the coefficients of the polynomial $g(\zeta^2)$ are all real, then if λ_ν is a complex ζ-zero of order m_ν, so is its complex conjugate at $\overline{\lambda_\nu}$ (this being just $-\lambda_\nu$ if λ_ν is purely imaginary). Correspondingly, if Re $\lambda_\nu \neq 0$, there is a quartet of four complex ζ-roots, each of order m_ν, to (3.5) at

$$\zeta_\nu(\chi), \ -\zeta_\nu(\chi), \ \overline{\zeta_\nu(\chi)}, \ -\overline{\zeta_\nu(\chi)}, \tag{4.9}$$

with $\zeta_\nu(\chi) \equiv (\chi^2\lambda_\nu^2 - 1)^{1/2}$. Suppose $0 < \arg \zeta_\nu(\chi) < \frac{1}{2}\pi$. Then among this quartet, only $\zeta_\nu(\chi)$ imparts a $\Phi_\nu(\chi)$ contribution of the type (3.6).

Clearly, there are no real ζ-roots to (3.5). So, unlike the more general ARR theory itself, it is, primarily, unnecessary to resort to a principal value interpretation of (3.2).

To formulate $\Phi_\nu(\chi)$ explicitly, we expand $F(\chi, 1+\zeta^2)$ and $G(\chi, 1+\zeta^2)$ about $\chi^{-2}(1+\zeta^2)=\lambda_\nu^2$ ($\operatorname{Re}\lambda_\nu \neq 0$), to provide us with two expansions about the point $\zeta=\zeta_\nu \equiv (\chi^2\lambda_\nu^2-1)^{1/2}$ within $0<\arg\zeta<\frac{1}{2}\pi$. Thus from (4.4) and (4.5), we have

$$F(\chi, 1+\zeta^2) \equiv \chi^{m-l}[f(\lambda_\nu^2)+\chi^{-2}(\zeta^2-\zeta_\nu^2)f'(\lambda_\nu^2)+\cdots], \tag{4.10}$$

$$\begin{aligned} G(\chi, 1+\zeta^2) \equiv \chi^m[&\chi^{-2}(\zeta^2-\zeta_\nu^2)g'(\lambda_\nu^2)+\tfrac{1}{2}\chi^{-4}(\zeta^2-\zeta_\nu^2)^2 g''(\lambda_\nu^2) \\ &+\tfrac{1}{6}\chi^{-6}(\zeta^2-\zeta_\nu^2)^3 g'''(\lambda_\nu^2)+\cdots], \end{aligned} \tag{4.11}$$

$g'(\lambda_\nu^2)$, $g''(\lambda_\nu^2)$, $g'''(\lambda_\nu^2)$ denoting first, second and third ζ^2-derivatives taken at $\zeta^2=\lambda_\nu^2$.

Case (i) λ_ν is of order $m_\nu=1$. By (4.3),

$$g'(\lambda_\nu^2) \equiv B_{(1/2)(m-q)} \prod_{\mu=1:\mu\neq\nu}^{(1/2)(m-q)} \lambda_\nu^2-\lambda_\mu^2 \neq 0, \tag{4.12}$$

with the product ranging over $\frac{1}{2}(m-q-2)$ μ-factors, avoiding $\mu=\nu$. It then follows from (3.6), (4.10) and (4.11) that

$$\Phi_\nu(\chi)=(-1)^{(1/2)(n-1)}\chi^{2-l}\operatorname{Im}\{(\chi^2\lambda_\nu^2-1)^{(1/2)n-2}f(\lambda_\nu^2)/g'(\lambda_\nu^2)\}. \tag{4.13}$$

Case (ii) λ_ν is of order $m_\nu=2$. Here $g'(\lambda_\nu^2)=0$, while

$$g''(\lambda_\nu^2) \equiv 2B_{(1/2)(m-q)} \prod_{\mu=1:\mu\neq\nu}^{(1/2)(m-q)} \lambda_\nu^2-\lambda_\mu^2 \neq 0; \tag{4.14}$$

the product ranges over $(1/2)(m-q-4)\mu$-factors. We can now go on to show that

$$\begin{aligned} \Phi_\nu(\chi)=&(-1)^{(1/2)(n-1)}(n-4)\chi^{4-l}\operatorname{Im}\{(\chi^2\lambda_\nu^2-1)^{(1/2)n-3}f(\lambda_\nu^2)/g''(\lambda_\nu^2)\} \\ &+2(-1)^{(1/2)(n-1)}\chi^{2-l} \\ &\cdot\operatorname{Im}\left\{(\chi^2\lambda_\nu^2-1)^{(1/2)n-2}\frac{f'(\lambda_\nu^2)g''(\lambda_\nu^2)-\frac{1}{3}f(\lambda_\nu^2)g'''(\lambda_\nu^2)}{[g''(\lambda_\nu^2)]^2}\right\}. \end{aligned} \tag{4.15}$$

One can thus proceed to generate more elaborate Φ_ν-forms for higher m_ν-orders. All such forms, including (4.13) and (4.15), are explicit in χ.

If λ_ν is substituted by $i|\lambda_\nu|$ so that $\zeta_\nu(\chi) \equiv (\chi^2\lambda_\nu^2-1)^{1/2}$ becomes $i|\zeta_\nu(\chi)| \equiv i(\chi^2|\lambda_\nu|^2+1)^{1/2}$, then after a transformation of the dummy ζ-variable in (3.6) to $i\zeta$, it is seen by comparison with (3.7) that $\Phi_\nu(\chi)$ becomes $2\Psi_\nu(\chi)$ with the m_ν-order preserved. We can therefore deduce from (4.13) and (4.15), the following Ψ_ν-forms for a purely imaginary λ_ν.

Case (iii) λ_ν is of order $m_\nu=1$.

$$\Psi_\nu(\chi)=\tfrac{1}{2}\chi^{2-l}(\chi^2|\lambda_\nu|^2+1)^{(1/2)n-2}f(-|\lambda_\nu|^2)/g'(-|\lambda_\nu|^2). \tag{4.16}$$

Case (iv) λ_ν is of order $m_\nu = 2$.

$$\Psi_\nu(\chi) = \tfrac{1}{2}(4-n)\chi^{4-l}(\chi^2|\lambda_\nu|^2+1)^{(1/2)n-3} f(-|\lambda_\nu|^2)/g''(-|\lambda_\nu|^2)$$
$$+\chi^{2-l}(\chi^2|\lambda_\nu|^2+1)^{(1/2)n-2} \tag{4.17}$$
$$\cdot \frac{f'(-|\lambda_\nu|^2)g''(-|\lambda_\nu|^2) - \tfrac{1}{3}f(-|\lambda_\nu|^2)g'''(-|\lambda_\nu|^2)}{[g''(-|\lambda_\nu|^2)]^2}.$$

5. Near-axial behaviors. Near the axis of symmetry, r, hence χ, is small:

$$|\chi| < \min_{\nu=1,\cdots,(1/2)m-(1/2)q} |\lambda_\nu|^{-1}, \tag{5.1}$$

say. We shall show that corresponding approximations for $\Phi_\nu(\chi)$ and $\Psi_\nu(\chi)$ are significant within the context of residue calculus.

By taking the binomial expansion of $(\chi^2\lambda_\nu^2 - 1)^{(1/2)n-2}$, and remembering that (2.4) and (3.3) now hold, we derive from (4.13) in the case $m_\nu = 1$,

$$\Phi_\nu(\chi) = \sum_{\mu=0}^{(1/2)(p-q-2)} 2c_{2\mu}\chi^{2-l+2\mu} \operatorname{Re}\left\{\frac{\lambda_\nu^{2\mu} f(\lambda_\nu^2)}{2g'(\lambda_\nu^2)}\right\} + O(\chi^{p-q-l+2}), \tag{5.2}$$

where

$$c_{2\mu} = (-1)^\mu \Gamma(\tfrac{1}{2}n-1)/\mu!\,\Gamma(\tfrac{1}{2}n-1-\mu). \tag{5.3}$$

Similarly, from (4.15), it can be shown that for $m_\nu = 2$,

$$\Phi_\nu(\chi) = \sum_{\mu=0}^{(1/2)(p-q-2)} 2c_{2\mu}\chi^{2-l+2\mu}$$
$$\cdot \operatorname{Re}\left\{\lambda_\nu^{2\mu}\frac{f'(\lambda_\nu^2)g''(\lambda_\nu^2) - \tfrac{1}{3}f(\lambda_\nu^2)g'''(\lambda_\nu^2)}{[g''(\lambda_\nu^2)]^2} + \mu\lambda_\nu^{2\mu-2}\frac{f(\lambda_\nu^2)}{g''(\lambda_\nu^2)}\right\} \tag{5.4}$$
$$+ O(\chi^{p-q-l+2}).$$

Each trailing $O(\chi^{p-q-l+2})$ quantity, being small, will henceforth be discarded. Regarding each curly bracketed factor, an even function of λ_ν:

$$\left\{\ \right\} \text{ in (5.2)} = \left[\frac{\zeta^{2\mu+1}f(\zeta^2)}{dg(\zeta^2)/d\zeta}\right]_{\zeta=\pm\lambda_\nu}, \tag{5.5}$$

$$\left\{\ \right\} \text{ in (5.4)} = 2\left\{\frac{d\zeta^{2\mu+1}f(\zeta^2)/d\zeta}{d^2g(\zeta^2)/d\zeta^2} - \frac{\zeta^{2\mu+1}f(\zeta^2)d^3g(\zeta^2)/d\zeta^3}{3[d^2g(\zeta^2)/d\zeta^2]^2}\right\}_{\zeta=\pm\lambda_\nu}. \tag{5.6}$$

The right sides in (5.5) and (5.6) are residues of the meromorphic function

$$\zeta^{2\mu+1}f(\zeta^2)/g(\zeta^2) \tag{5.7}$$

at the pole $\zeta = \pm\lambda_\nu$ of respective orders one and two. From arguments in § 4, all complex poles over $\operatorname{Re}\zeta \neq 0$ must group into ν-quartets of the type (cf. (4.9)) $\{\nu\} = \{\lambda_\nu, -\lambda_\nu, \bar{\lambda}_\nu, -\bar{\lambda}_\nu\}$, wherein every member has the same order m_ν and is

allocated to a separate quadrant of the ζ-plane. Accordingly then, regardless of whether $m_\nu = 1$ or 2,

$$\Phi_\nu(\chi) \approx \sum_{\mu=0}^{(1/2)(p-q-2)} \tfrac{1}{2} c_{2\mu} \chi^{2-l+2\mu} \sum_{\zeta=\{\nu\}} \text{residue}\,[\zeta^{2\mu+1} f(\zeta^2)/g(\zeta^2)] \tag{5.8}$$

with $\sum_{\zeta=\{\nu\}}$ ranging over the associated ν-quartet.

Corresponding approximations for (4.16) and (4.17) may be deduced from (5.2) and (5.4), respectively, by substituting the upper purely imaginary pole $i|\lambda_\nu|$ for λ_ν, and halving each result. Now, all purely imaginary poles arise in symmetric ν-pairs of the type $[\nu] = [i|\lambda_\nu|, -i|\lambda_\nu|]$. We can go on to show that, for an order $m_\nu = 1$ or 2 of the poles $i|\lambda_\nu|, -i|\lambda_\nu|$,

$$\Psi_\nu(\chi) \approx \sum_{\mu=0}^{(1/2)(p-q-2)} \tfrac{1}{2} c_{2\mu} \chi^{2-l+2\mu} \sum_{\zeta=[\nu]} \text{residue}\,[\zeta^{2\mu+1} f(\zeta^2)/g(\zeta^2)]. \tag{5.9}$$

Unless $l = 1$ or 2, the expressions (5.8) and (5.9) involve terms that are singular at $\chi = 0$. As we shall see next, when the $\Phi_\nu(\chi)$'s and $\Psi_\nu(\chi)$'s combine near $\chi = 0$ in the sense of (3.4), their near-singular terms can mutually cancel out.

Suppose the zeros (all nonreal) of $g(\zeta^2)$ are, at most, of order two. Then (3.4), (5.8) and (5.9) imply, via Cauchy's residue theorem, the following corollary:

$$K(\chi) \approx \sum_{\mu=0}^{(1/2)(p-q-2)} (4\pi i)^{-1} c_{2\mu} \chi^{2-l+2\mu} \oint \zeta^{2\mu+1} f(\zeta^2)/g(\zeta^2) d\zeta \tag{5.10}$$

with $\oint$ performed over the closed circle $\zeta = R \exp(i\theta)$ enclosing all such zeros and of arbitrarily large radius $R > \max_{\nu=1,\cdots,(1/2)(m-q)} |\lambda_\nu|$. Now, via (1.2) and (1.3), if $p < m$,

$$\begin{aligned}
&\left| \oint \zeta^{2\mu+1} f(\zeta^2)/g(\zeta^2)\, d\zeta - \oint \zeta^{2\mu+1+q-p} A_{(1/2)(m-p)}/B_{(1/2)(m-q)}\, d\zeta \right| \\
&= \left| \oint \zeta^{2\mu+1+q-p} \frac{A_{(1/2)(m-p)}}{B_{(1/2)(m-q)}} \left[1 - \frac{1 + \cdots + \zeta^{p-m} A_0/A_{(1/2)(m-p)}}{1 + \cdots + \zeta^{q-m} B_0/B_{(1/2)(m-q)}} \right] d\zeta \right| \\
&\leqq \frac{2\pi \{ |A_{(1/2)(m-p)}| (|B_{(1/2)(m-q-2)}| R^{-2} + \cdots + |B_0| R^{q-m}) + |B_{(1/2)(m-q)}| (|A_{(1/2)(m-p-2)}| R^{-2} + \cdots + |A_0| R^{p-m}) \}}{R^{p-q-2-2\mu} |B_{(1/2)(m-q)}| \{ |B_{(1/2)(m-q)}| - (|B_{(1/2)(m-q-2)}| R^{-2} + \cdots + |B_0| R^{q-m}) \}}
\end{aligned} \tag{5.11}$$

for a sufficiently large R; but if $p = m$,

$$\begin{aligned}
&\left| \oint \zeta^{2\mu+1} f(\zeta^2)/g(\zeta^2)\, d\zeta - \oint \zeta^{2\mu+1+q-p} A_0/B_{(1/2)(m-q)}\, d\zeta \right| \\
&\leqq \frac{2\pi |A_0| (|B_{(1/2)(m-q-2)}| R^{-2} + \cdots + |B_0| R^{q-m})}{R^{p-q-2-2\mu} |B_{(1/2)(m-q)}| \{ |B_{(1/2)(m-q)}| - (|B_{(1/2)(m-q-2)}| R^{-2} + \cdots + |B_0| R^{q-m}) \}}.
\end{aligned} \tag{5.12}$$

When $\mu = 0, \cdots, \frac{1}{2}(p-q-2)$, both expressions (5.11) and (5.12) approach zero

as $R \to \infty$, in which event,

$$\lim_{R\to\infty} \oint \zeta^{2\mu+1} f(\zeta^2)/g(\zeta^2)\, d\zeta$$

$$= \lim_{R\to\infty} (A_{(1/2)(m-p)}/B_{(1/2)(m-q)}) \oint \zeta^{2\mu+1+q-p}\, d\zeta$$

$$= \lim_{R\to\infty} iR^{2\mu-(p-q-2)}(A_{(1/2)(m-p)}/B_{(1/2)(m-q)}) \int_0^{2\pi} \exp[-i(p-q-2-2\mu)\theta]\, d\theta$$

$$= \begin{cases} 0 & \text{if } 0 \leqq \mu < \tfrac{1}{2}(p-q-2), \\ 2\pi i A_{(1/2)(m-p)}/B_{(1/2)(m-q)} & \text{if } \mu = \tfrac{1}{2}(p-q-2), \end{cases}$$

valid for $p \leqq m$. So (5.10) reduces to

$$K(\chi) \approx \tfrac{1}{2}\chi^{p-q-l} c_{p-q-2} A_{(1/2)(m-p)}/B_{(1/2)(m-q)}, \tag{5.13}$$

which is, in view of (2.4), nonsingular at $\chi = 0$. Obviously, possible singularly inclined terms encountered in the $\Phi_\nu(\chi)$'s and $\Psi_\nu(\chi)$'s have combined, via the residue theorem, into zero converging contour integrals along an indefinitely expanding circuit.

6. Moving concentrated force in an elastic medium. We shall next illustrate a simple application to the Eason, Fulton and Sneddon [5] elastodynamic problem concerning a force $\mathbf{X}$ (per unit mass), moving with uniform speed v along an x_1-direction and within an infinite medium. The three authors cited also started with a Fourier transformation; immediately after, however, their subsequent interpretation followed a different route.

With reference to a moving $\mathbf{x} = (x_1, x_2, x_3)$ R_3-frame, fixed relative to the force, the equation of motion governing the elastic displacement $\mathbf{u}$ is, in the steady state,

$$(v_1^2 - v_2^2)\nabla(\nabla \cdot \mathbf{u}) + v_2^2 \nabla^2 \mathbf{u} + \mathbf{X} = v^2 \partial^2 \mathbf{u}/\partial x_1^2, \tag{6.1}$$

v_1 and v_2 being, respectively, the speeds of compressional and shear waves, while ∇ denotes the gradient operator. Now, (6.1) can be vectorially manipulated into

$$G(\partial/\partial x_1, \nabla_1^2)\mathbf{u} = (v_1^2 - v_2^2)\nabla(\nabla \cdot \mathbf{X}) + (v^2 \partial^2/\partial x_1^2 - v_1^2 \nabla^2)\mathbf{X}, \tag{6.2}$$

where

$$G(\partial/\partial x_1, \nabla_1^2) \equiv [(v_1^2 - v^2)\partial^2/\partial x_1^2 + v_1^2 \nabla_1^2][(v_2^2 - v^2)\partial^2/\partial x_1^2 + v_2^2 \nabla_1^2], \tag{6.3}$$

and is precisely of the form (1.3) with $m = 4$ and $q = 0$.

Suppose the force is point concentrated and acts only along its path direction, viz., $\mathbf{X} = (1, 0, 0)\delta(\mathbf{x})$. Let u_1, u_r and u_θ denote, respectively, the axial, radial and azimuthal components of $\mathbf{u}$ relative to a (x_1, r, θ) cylindrical polar frame: $x_2 = r\cos\theta$, $x_3 = r\sin\theta$. Thus it follows from (6.2) that $u_\theta \equiv 0$; but

$$G(\partial/\partial x_1, \nabla_1^2)u_1 = [(v^2 - v_2^2)\partial^2/\partial x_1^2 - v_1^2 \nabla_1^2]\delta(\mathbf{x}); \tag{6.4}$$

also if

$$G(\partial/\partial x_1, \nabla_1^2)\psi = (v_1^2 - v_2^2)(\partial/\partial x_1)\delta(\mathbf{x}), \tag{6.5}$$

then

$$u_r = \partial\psi/\partial r. \tag{6.6}$$

Applying (4.1) to (6.3), we have

$$g(\zeta^2) \equiv G(1, \zeta^2) \equiv v_1^2 v_2^2 (1-\mu_1^2+\zeta^2)(1-\mu_2^2+\zeta^2), \tag{6.7}$$

with $\mu_1 = vv_1^{-1}$, $\mu_2 = vv_2^{-1}$. Normally, $v_1 > v_2$. To achieve ellipticity of the G-operator, we assume a slow passage of the force: $1 > \mu_2\ (> \mu_1)$. Whence, $g(\zeta^2)$ has four purely imaginary ζ-zeros, each of order one, at

$$\zeta = i\sqrt{1-\mu_1^2},\quad -i\sqrt{1-\mu_1^2},\quad i\sqrt{1-\mu_2^2},\quad -i\sqrt{1-\mu_2^2}, \tag{6.8}$$

of which $i\sqrt{1-\mu_1^2}$, $i\sqrt{1-\mu_2^2}$ impart two Ψ_ν-contributions, each representable by (4.16); no Φ_ν-term participates.

Equation (6.4) comes under (1.1) if

$$F(\partial/\partial x_1, \nabla_1^2) \equiv (v^2 - v_2^2)\partial^2/\partial x_1^2 - v_1^2\nabla_1^2, \tag{6.9}$$

which is of the type (1.2) with $l = 2 = p$. Since $q = 0$, the convergence criteria (2.4) and (3.3) are satisfied. So, via (2.6), (3.4) and (4.16), we eventually arrive at

$$u_1 = -(2\pi|x_1|)^{-1}K(\chi), \tag{6.10}$$

$$= (4\pi v^2)^{-1}[(\mathbf{x}^2 - \mu_1^2 r^2)^{-1/2} - (1-\mu_2^2)(\mathbf{x}^2 - \mu_2^2 r^2)^{-1/2}]. \tag{6.11}$$

Equation (6.5) is also covered by (1.1), with the same G-operator of (6.3), but involving a different

$$F(\partial/\partial x_1, \nabla_1^2) \equiv (v_1^2 - v_2^2)\partial/\partial x_1. \tag{6.12}$$

This excludes ∇_1^2 and represents a highly elementary subform of (1.2) with $l = 3$ and $p = 4$. Again (2.4) and (3.3) hold. Hence, as above, but starting from (2.7), and incorporating (6.12) instead, via (4.6), into (4.16), we finally establish

$$u_r = \partial\psi/\partial r = (2\pi|x_1|)^{-1}K(\chi) \tag{6.13}$$

$$= (4\pi r v^2)^{-1} x_1[(\mathbf{x}^2 - \mu_2^2 r^2)^{-1/2} - (\mathbf{x}^2 - \mu_1^2 r^2)^{-1/2}]. \tag{6.14}$$

Our solution for $\mathbf{u}$ is now complete; it agrees with that found in [5].

Suppose $v \to 0$. Then it can be deduced from (6.11) and (6.14) that

$$u_1 = (8\pi v_1^2 v_2^2)^{-1}[r^2|\mathbf{x}|^{-3}(v_2^2 - v_1^2) + 2v_1^2|\mathbf{x}|^{-1}] + O(v^2), \tag{6.15}$$

$$u_r = (8\pi v_1^2 v_2^2)^{-1}(v_1^2 - v_2^2) r x_1 |\mathbf{x}|^{-3} + O(v^2). \tag{6.16}$$

Consider the limit when $v = 0$. Expression (6.3) reveals that

$$(v_1 v_2)^{-2} G(\partial/\partial x_1, \nabla_1^2) \equiv \nabla^4, \tag{6.17}$$

the biharmonic operator. Correspondingly, both upper contributing zeros in the set of (6.8) now coincide at $\zeta = i$, a purely imaginary ζ-zero of order two. Thereupon, it can be verified that if the above procedures for deriving u_1 and u_r

are repeated with (4.17) instead of (4.16), the respective results obtained are the two dominant (or limit) terms in (6.15) and (6.16).

REFERENCES

[1] J. BAZER AND D. H. Y. YEN, *Lacunas of the Riemann matrix of symmetric hyperbolic systems in two space variables*, Comm. Pure Appl. Math., 22 (1969), pp. 279–333.
[2] R. BURRIDGE, *Lacunas in two-dimensional wave propagation*, Proc. Cambridge Philos. Soc., 63 (1967), pp. 819–825.
[3] L. CHEE-SENG, *Axial radiation reception*, Ibid., 74 (1973), pp. 369–395.
[4] ———, *Axial radiation reception. II*, this Journal, 8 (1977), pp. 24–51.
[5] G. EASON, J. FULTON AND I. N. SNEDDON, *Generation of waves in an infinite elastic solid by variable body forces*, Philos. Trans. Roy. Soc. London Ser. A, 248 (1956), pp. 575–607.
[6] F. JOHN, *Plane Waves and Spherical Means*, Interscience, New York, 1955.
[7] R. G. PAYTON, *Two-dimensional anisotropic elastic waves emanating from a point source*, Proc. Cambridge Philos. Soc., 70 (1971), pp. 191–210.
[8] H. WEITZNER, *Green's function for two-dimensional magnetohydrodynamic waves*, Phys. Fluids, 4 (1961), pp. 1238–1245.

SIAM J. MATH. ANAL.
Vol. 8, No. 5, October 1977

ON THE DISTRIBUTION OF THE EIGENVALUES OF A TWO-PARAMETER SYSTEM OF ORDINARY DIFFERENTIAL EQUATIONS OF THE SECOND ORDER*

M. FAIERMAN†

Abstract. In this paper two simultaneous Sturm–Liouville systems are considered, the first defined for the interval $0 \leqq x_1 \leqq 1$, the second for the interval $0 \leqq x_2 \leqq 1$, and each containing the parameters λ and μ. Denoting the eigenvalues and eigenfunctions of the simultaneous systems by $(\lambda_{j,k}, \mu_{j,k})$ and $\psi_{j,k}(x_1, x_2)$, respectively, $j, k = 0,1, \cdots$, the principle of the argument and asymptotic methods are employed to derive asymptotic formulas for these expressions, as $j^2+k^2 \to \infty$, when (j, k) is restricted to lie in each of several sectors of the (x, y)-plane. These results partially resolve a problem posed by Atkinson concerning the behavior of the eigenvalues and eigenfunctions of multiparameter Sturm–Liouville systems and constitute a further stage in the development of the theory related to these questions.

1. Introduction. The multiparameter analogue of the classical Sturm–Liouville problem arises naturally when we seek to solve boundary value problems associated with the potential or wave equation by use of the method of separation of variables. For, depending upon the coordinate system concerned, it may not always be possible to separate the spectral parameters when going from the partial differential equation to a set of ordinary differential equations (we refer to Sleeman [14] for a detailed discussion). In spite of the importance of such multiparameter problems in mathematical physics, it has been pointed out by Atkinson [2, Introd.], [3, § 4] that this field has been relatively neglected in recent years, and in particular he states that as opposed to the single parameter case, the detailed behavior of the eigenvalues and eigenfunctions of multiparameter Sturm–Liouville systems is still far from clear. Since the appearance of Atkinsons' papers, the author [7], [8] has obtained some results pertaining to the behavior of the eigenvalues of a two-parameter system, and in this paper we shall continue with our investigation of the two-parameter case and further results concerning the behavior of the eigenvalues, as well as results concerning the eigenfunctions, will be established.

We shall be concerned here with the simultaneous two-parameter systems

$$y_1'' + (\lambda A_1(x_1) - \mu B_1(x_1) + q_1(x_1))y_1 = 0, \qquad 0 \leqq x_1 \leqq 1, \quad ' = d/dx_1, \tag{1.1}$$

$$\begin{aligned} y_1(0)\cos\alpha_1 - y_1'(0)\sin\alpha_1 &= 0, & 0 &\leqq \alpha_1 < \pi, \\ y_1(1)\cos\beta_1 - y_1'(1)\sin\beta_1 &= 0, & 0 &< \beta_1 \leqq \pi, \end{aligned} \tag{1.2}$$

and

$$y_2'' + (-\lambda A_2(x_2) + \mu B_2(x_2) + q_2(x_2))y_2 = 0, \qquad 0 \leqq x_2 \leqq 1, \quad ' = d/dx_2, \tag{1.3}$$

$$\begin{aligned} y_2(0)\cos\alpha_2 - y_2'(0)\sin\alpha_2 &= 0, & 0 &\leqq \alpha_2 < \pi, \\ y_2(1)\cos\beta_2 - y_2'(1)\sin\beta_2 &= 0, & 0 &< \beta_2 \leqq \pi, \end{aligned} \tag{1.4}$$

* Received by the editors April 23, 1975, and in final revised form September 13, 1976.
† Department of Mathematics, Ben Gurion University of the Negev, Beer Sheva, Israel.

where it will be supposed that, for $r=1, 2$, A_r, B_r, and q_r are real-valued, continuous functions in $0 \leqq x_r \leqq 1$, with both A_r and B_r having absolutely continuous first derivatives in this interval, and that $\Delta = (A_1B_2 - A_2B_1) \neq 0$ in I^2 (the product of the intervals $0 \leqq x_1 \leqq 1$, $0 \leqq x_2 \leqq 1$). Furthermore, there is no loss of generality in assuming henceforth that the A_r, B_r, and Δ are all positive for all values of x_1 and x_2 in their respective intervals, since this can always be achieved, if necessary, by introducing a nonsingular transformation in the parameters λ and μ (see Appendix A). In § 2 we collect some known facts concerning the above systems and these are used in § 3 to establish results concerning the asymptotic developments of the eigenvalues and eigenfunctions (see Theorem 3.4). In § 4 we extend the foregoing results for the case where A_2/B_2 is constant in $0 \leqq x_2 \leqq 1$.

2. Preliminary results.

2.0. Introduction. We shall now collect some known facts concerning the systems (1.1)–(1.2) and (1.3)–(1.4) which we require for later use. Firstly we need the following definitions. Let b_1 and b_2 be the infimum and supremum, respectively, of $A_1(x_1)/B_1(x_1)$ in $0 \leqq x_1 \leqq 1$ and a_1 and a_2 the infimum and supremum, respectively, of $A_2(x_2)/B_2(x_2)$ in $0 \leqq x_2 \leqq 1$ (hence $0 < a_1 \leqq a_2 < b_1 \leqq b_2$). For $r=1, 2$, let ϕ_r denote the solution of the differential equation $(1.2r-1)$ satisfying $\phi_r(0, \lambda, \mu) = \sin \alpha_r$, $\phi_r'(0, \lambda, \mu) = \cos \alpha_r$. We note that ϕ_r and ϕ_r' are entire functions of λ and μ for each fixed x_r, $r=1, 2$.

2.1. The system (1.1)–(1.2). We know from the Sturm theory that for each real λ, the totality of the values of μ for which (1.1) has a nontrivial solution satisfying (1.2) forms a countably infinite set of real numbers which we shall denote by $\mu_n(\lambda)$, $n \geqq 0$, where $\mu_0(\lambda) > \mu_1(\lambda) > \cdots$, $\mu_n(\lambda) \to -\infty$ as $n \to \infty$, and the solution corresponding to $\mu_n(\lambda)$ has precisely n zeros in $0 < x_1 < 1$. From [8, § 2.1], [13, § 2] we also know that for each n, $\mu_n(\lambda)$ is analytic in $-\infty < \lambda < \infty$ and at each point of this interval, $b_1 \leqq d\mu_n(\lambda)/d\lambda \leqq b_2$. Finally, we note for later use that if (1.1), with $\lambda = \lambda^*$ and $\mu = \mu^*$, has a nontrivial solution satisfying (1.2), then (λ^*, μ^*) is a zero of $W_1(\lambda, \mu) = [\phi_1(1, \lambda, \mu) \cos \beta_1 - \phi_1'(1, \lambda, \mu) \sin \beta_1]$, and conversely. From [9, Lem. 4.1, p. 206] we also know that if $W_1(\lambda^*, \mu^*) = 0$, then

$$\partial W_1(\lambda^*, \mu^*)/\partial \mu = -K \int_0^1 B_1(x_1) \phi_1^2(x_1, \lambda^*, \mu^*)\, dx_1, \tag{2.1}$$

where K equals $(\sin \beta_1)/\phi_1(1, \lambda^*, \mu^*)$ if $\beta_1 \neq \pi$ and equals $-1/\phi_1'(1, \lambda^*, \mu^*)$ otherwise.

2.2. The system (1.3)–(1.4). Definitions and results (with obvious modifications) completely analogous to the system (1.1)–(1.2) hold for the system (1.3)–(1.4). Here the analogue of $\mu_n(\lambda)$ will be denoted by $\mu_n^*(\lambda)$, $n \geqq 0$, and that of W_1 by W_2. We might note that now $a_1 \leqq d\mu_n^*(\lambda)/d\lambda \leqq a_2$ for $-\infty < \lambda < \infty$ and $n \geqq 0$.

2.3. The system (1.1)–(1.4). By an eigenvalue of the system (1.1)–(1.4) we mean a pair of numbers, (λ^*, μ^*), such that for $\lambda = \lambda^*$ and $\mu = \mu^*$, (1.1) and (1.3) have nontrivial solutions satisfying (1.2) and (1.4), respectively. If $y_1(x_1, \lambda^*, \mu^*)$ and $y_2(x_2, \lambda^*, \mu^*)$ denote these solutions, respectively, then the product,

$\prod_{r=1}^{2} y_r(x_r, \lambda^*, \mu^*)$, is called an eigenfunction of the system (1.1)–(1.4) corresponding to (λ^*, μ^*).

It is clear that the eigenvalues of the system (1.1)–(1.4) are precisely the zeros of the simultaneous equations $W_1 = 0$, $W_2 = 0$. Now let $J(\lambda, \mu)$ denote the Jacobian of W_1 and W_2 at the point (λ, μ). Then from [4, Prob. 16, p. 551, and pp. 160–168], [9, Lem. 4.1, p. 206], and [12, pp. 248–251] we know that the eigenvalues of the system (1.1)–(1.4) (and hence the zeros of the simultaneous equations $W_1 = 0$, $W_2 = 0$) form a countably infinite subset of E^2 (real Euclidean 2-space) such that at each eigenvalue $J \neq 0$ and with eigenfunctions corresponding to distinct eigenvalues being orthogonal in L_Δ^2 (the Hilbert space constructed from those functions which are absolutely square-integrable in I^2 and with inner product $(g, h) = \iint_{I^2} \Delta g \bar{h}\, dx_1\, dx_2$). Furthermore, if p_1 and p_2 are any nonnegative integers, then there is precisely one eigenvalue of the system (1.1)–(1.4), say (λ^*, μ^*), such that $\phi_r(x_r, \lambda^*, \mu^*)$ has exactly p_r zeros in $0 < x_r < 1$, $r = 1, 2$. From these results it follows that the eigenvalues of the system (1.1)–(1.4) may be denoted by $\{(\lambda_{j,k}, \mu_{j,k})\}_{j,k=0}^{\infty}$, where $\phi_1(x_1, \lambda_{j,k}, \mu_{j,k})$ has precisely j zeros in $0 < x_1 < 1$ and $\phi_2(x_2, \lambda_{j,k}, \mu_{j,k})$ has precisely k zeros in $0 < x_2 < 1$; and it is this notation for the eigenvalues of the system (1.1)–(1.4) which will be used in the sequel. Henceforth we shall also put $\psi_{j,k}(x_1, x_2) = \psi_{j,k}^*(x_1, x_2)/\|\psi_{j,j}^*\|$, where $\psi_{j,k}^*(x_1, x_2) = \prod_{r=1}^{2} \phi_r(x_r, \lambda_{j,k}, \mu_{j,k})$ and $\|\cdot\|$ denotes the norm in L_Δ^2 constructed from the scalar product $(\cdot, \cdot)$.

Finally, to see the connection between the results of this subsection and those of § 2.1 and § 2.2, we note that as λ runs from $-\infty$ to ∞, the $\mu_n(\lambda)$ determine a countably infinite number of disjoint analytic curves in E^2, which we shall denote by C_n. A similar result also holds for the $\mu_n^*(\lambda)$ and we shall denote the curves which they determine by C_n^*. Then from the foregoing results, it follows that the eigenvalues of the system (1.1)–(1.4) and the points of intersection of the curves C_n with C_n^* are identical. Indeed, if j and k are any nonnegative integers, then C_j intersects C_k^* in precisely one point, namely at the eigenvalue of the system (1.1)–(1.4), $(\lambda_{j,k}, \mu_{j,k})$.

3. Main results.

3.0. Introduction. In this section we shall use the results of § 2 to derive asymptotic formulas for $\lambda_{j,k}$, $\mu_{j,k}$, and $\psi_{j,k}$ as $j \to \infty$ when (j, k) belongs to a certain sector Ω of the (x, y)-plane which will be defined below.

Notation. For the remainder of this paper we let:

1. $P_1(x_1, \lambda, \mu) = \lambda A_1(x_1) - \mu B_1(x_1)$ and $P_2(x_2, \lambda, \mu) = -\lambda A_2(x_2) + \mu B_2(x_2)$;
2. $h_1(t) = \int_0^1 (A_1(x_1) - tB_1(x_1))^{1/2}\, dx_1$ for $-\infty < t \leqq b_1$ and $h_2(t) = \int_0^1 (-A_2(x_2) + tB_2(x_2))^{1/2}\, dx_2$ for $a_2 \leqq t < \infty$ (here positive square roots are taken);
3. $g(t) = h_2(t)/h_1(t)$ for $a_2 \leqq t < b_1$, $g(b_1) = h_2(b_1)/h_1(b_1)$ if $h_1(b_1) \neq 0$, and $g(b_1) = \infty$ otherwise (observe that $g(t)$ is nonnegative, continuous, strictly increasing in $[a_2, b_1)$, and tends to $g(b_1)$ as $t \to b_1$);
4. $\theta_1^* = \tan^{-1} g(a_2)$, $\theta_2^* = \tan^{-1} g(b_1)$ if $g(b_1) \neq \infty$, $\theta_2^* = \pi/2$ otherwise, where the principal branch of the inverse tangent is taken;
5. θ_1, θ_2 be any two numbers satisfying $\theta_1^* < \theta_1 < \theta_2 < \theta_2^*$;

6. t_j denote the solution of the equation $g(t) = \tan \theta_j$ for $j = 1, 2$ (observe that $a_2 < t_1 < t_2 < b_1$);

7. $\delta = \min\{(t_1 - a_2), (b_1 - t_2)\}$;

8. Ω denote the sector in the (x, y)-plane defined by the inequalities $\theta_1 \leqq \theta \leqq \theta_2$, where θ denotes the angle which a ray emanating from the origin makes with the positive x-axis;

9. $B^\dagger = \max\{B_1^\dagger, B_2^\dagger\}$, $B^* = \min\{B_1^*, B_2^*\}$, where $B_j^\dagger$ and B_j^* denote the supremum and infimum, respectively, of $B_j(x_j)$ in $0 \leqq x_j \leqq 1$ for $j = 1, 2$.

3.1. Some estimates. We are now going to derive some estimates for $\lambda_{j,k}$ and $\mu_{j,k}$ for $(j, k) \in \Omega$. Our main result here, given in Theorem 3.3, will then be used in subsection 3.2 to establish asymptotic formulae for these expressions as $j \to \infty$.

Notation. 1. Let p_1, p_2 be any positive integers satisfying $\tan \theta_1 \leqq p_2/p_1 \leqq \tan \theta_2$ such that for $r = 1, 2$, (i) p_r is odd if $\alpha_r \neq 0$ and $\beta_r \neq \pi$ or if $\alpha_r = 0$ and $\beta_r = \pi$, (ii) p_r is even if $\alpha_r \neq 0$ and $\beta_r = \pi$ or if $\alpha_r = 0$ and $\beta_r \neq \pi$ (see (1.2) and (1.4) for the definitions of the α_r and β_r).

2. Let t^* denote the solution of the equation $g(t) = p_2/p_1$ (observe that $t_1 \leqq t^* \leqq t_2$).

3. Let p_3 be the smallest odd integer greater than σ $(= p_1 h_1(0)/h_1(t^*))$ if $\alpha_1 \neq 0$ and $\beta_1 \neq \pi$ or if $\alpha_1 = 0$ and $\beta_1 = \pi$, and the smallest even integer greater than σ if $\alpha_1 \neq 0$ and $\beta_1 = \pi$ or if $\alpha_1 = 0$ and $\beta_1 \neq \pi$.

4. Let $t^\dagger$ denote the solution of the equation $h_1(t) = p_3 h_1(t^*)/p_1$. Observe that $t^\dagger < 0$, and a simple calculation involving the definition of $h_1(t)$ shows that if $t = -c$ is the solution of the equation $h_1(t) = h_1(0) + 2(B^\dagger b_2)^{1/2}$, then $|t^\dagger| < c$. Observe also that $p_3 > p_1$ and

$$h_1(t^*)/p_1 = h_2(t^*)/p_2 = h_1(t^\dagger)/p_3. \tag{3.1}$$

5. Let n_0 be the smallest integer exceeding $[3 + (2d^2/\pi)(B^\dagger b_2)^{1/2}]$, where d denotes a number exceeding 4 chosen large enough so that $(4/d)\log d < \min\{1, B^*\delta/B^\dagger\langle 1 + [B^\dagger(b_2 + c)]^{1/4}\rangle\}$.

6. Let n be an odd integer exceeding n_0.

7. Let $\alpha_3 = \alpha_1$, $\beta_3 = \beta_1$, and $N_j = (np_j - r_j)/2$ for $j = 1, 2, 3$, where $r_j = 1$ if $\alpha_j \neq 0$ and $\beta_j \neq \pi$, $r_j = 2$ if $\alpha_j \neq 0$ and $\beta_j = \pi$ or if $\alpha_j = 0$ and $\beta_j \neq \pi$, and $r_j = 3$ if $\alpha_j = 0$ and $\beta_j = \pi$.

8. Let $\lambda_n = (p_1 n\pi/2h_1(t^*))^2$.

Let Γ denote the closed curve lying in the plane of the complex variable μ defined by

$$\begin{aligned}
\mu = \mu(s) &= [t^* + is]\lambda_n, && 0 \leqq s \leqq 1,\\
&= [\langle (2-s)t^* + (s-1)t^\dagger\rangle + i]\lambda_n, && 1 < s \leqq 2,\\
&= [t^\dagger + i(5-2s)]\lambda_n, && 2 < s \leqq 3,\\
&= [\langle (4-s)t^\dagger + (s-3)t^*\rangle - i]\lambda_n, && 3 < s \leqq 4,\\
&= [t^* - i(5-s)]\lambda_n, && 4 < s \leqq 5.
\end{aligned} \tag{3.2}$$

Observing that $t^\dagger\lambda_n < 0 < a_1\lambda_n \leqq a_2\lambda_n < t^*\lambda_n < b_1\lambda_n \leqq b_2\lambda_n$, it follows from the definition of the P_j above that $P_1(x_1, \lambda_n, \mu)$ cannot vanish within and on Γ for any

x_1 in [0, 1], while if x_2 is any point of [0, 1], then $P_2(x_2, \lambda_n, \mu)$ has precisely one (simple) zero, which lies within Γ, and in particular, in the interval $a_1\lambda_n \leqq \operatorname{Re}\mu \leqq a_2\lambda_n$, $\operatorname{Im}\mu = 0$. Turning next to the zeros of the $W_j(\lambda_n, \mu)$ within and on Γ and referring to the definitions given in § 2, we have

THEOREM 3.1. *There exists the integer n_1 ($> n_0$), depending on the θ_r, but not on the p_r, $r = 1, 2$, such that $W_j(\lambda_n, \mu)$ cannot vanish on Γ for $n > n_1$, $j = 1, 2$. Furthermore, if $n > n_1$, then $W_1(\lambda_n, \mu)$ and $W_2(\lambda_n, \mu)$ have exactly $(N_3 - N_1)$ and $(N_2 + 1)$ zeros, respectively, within Γ (a zero of order p being counted p times); these zeros are all simple, with the zeros of $W_1(\lambda_n, \mu)$ occurring at precisely the points $\mu_j(\lambda_n)$, $j = (N_1 + 1), \cdots, N_3$, and those of $W_2(\lambda_n, \mu)$ at precisely the points $\mu_k^*(\lambda_n)$, $k = 0, 1, \cdots, N_2$.*

In order to prove the theorem, we shall need to adopt a convention for dealing with powers of the P_j. Hence let (λ, μ) be any tuple of complex numbers for which $P_j(x_j, \lambda, \mu) \neq 0$ in $0 \leqq x_j \leqq 1$, and for convenience of notation write $P_j(x_j)$ for $P_j(x_j, \lambda, \mu)$. Then we shall agree *for the remainder of this paper*, unless otherwise stated, to assign to $\arg P_j(0)$ its principal value, determine $\arg P_j(x_j)$ in $0 \leqq x_j \leqq 1$ by continuity, and in this interval interpret the expression $P_j^{\nu}(x_j)$ (ν real) in accordance with the rule

$$P_j^{\nu}(x_j) = |P_j(x_j)|^{\nu} \exp\{i\nu \arg P_j(x_j)\}. \tag{3.3}$$

We now observe from (3.1) and the definition of λ_n that

$$\begin{aligned} &\int_0^1 P_j^{1/2}(x_j, \lambda_n, t^*\lambda_n)\, dx_j = p_j n\pi/2, && j = 1, 2, \\ &\int_0^1 P_1^{1/2}(x_1, \lambda_n, t^{\dagger}\lambda_n)\, dx_1 = p_3 n\pi/2. \end{aligned} \tag{3.4}$$

Proof of Theorem 3.1. We shall only prove the theorem for W_2 and only under the assumption that $\alpha_2 = 0$, $\beta_2 = \pi$; the other cases can be treated in a similar manner (we remark that in proving the theorem for W_1, we argue with each of the closed curves bounding the rectangles $t^*\lambda_n \leqq \operatorname{Re}\mu \leqq (a_1 + b_2)\lambda_n$, $|\operatorname{Im}\mu| \leqq \lambda_n$, and $t^{\dagger}\lambda_n \leqq \operatorname{Re}\mu \leqq (a_1 + b_2)\lambda_n$, $|\operatorname{Im}\mu| \leqq \lambda_n$, respectively, instead of with Γ). *Then with μ a point of* Γ, and writing $P(x_2)$ for $P_2(x_2, \lambda_n, \mu)$, $\phi(x_2)$ for $\phi_2(x_2, \lambda_n, \mu)$, we see from the variation-of-constants formula that

$$\phi^*(x_2) = \sin\Theta(x_2, 0) - \int_0^{x_2} \sin\Theta(x_2, \tau) P^{-1/2}(\tau) Q(\tau) \phi^*(\tau)\, d\tau, \tag{3.5}$$

where $\phi^*(\tau) = P^{1/4}(0) P^{1/4}(\tau) \phi(\tau)$, $\Theta(\tau_2, \tau_1) = \int_{\tau_1}^{\tau_2} P^{1/2}(\tau)\, d\tau$, and $Q(\tau) = \{q_2(\tau) + P^{1/4}(\tau)[d^2 P^{-1/4}(\tau)/d\tau^2]\}$. Arguing with the Gronwall lemma, it then follows that

$$\phi(x_2) = P^{-1/4}(0) P^{-1/4}(x_2)[\sin\Theta(x_2, 0) + O(n^{-1} \exp\Theta^*(x_2, 0))] \tag{3.6}$$

for $0 \leqq x_2 \leqq 1$, where $\Theta^*(\tau_2, \tau_1) = \int_{\tau_1}^{\tau_2} |\operatorname{Im} P^{1/2}(\tau)|\, d\tau$. It is important to observe here that the constant implied in the O symbol *does not depend upon x_2, μ, n, p_1, or p_2, but depends upon δ, and hence upon θ_1 and θ_2*. By noting that $\operatorname{Im} P^{1/2}(x_2)$ is of one sign in $0 \leqq x_2 \leqq 1$ and by arguing with the results of [1, p. 3] and the definition of n_0, we may now show that: (i) when $\operatorname{Re}\mu = t^*\lambda_n$ and $|\operatorname{Im}\mu| \leqq \lambda_n^{1/2} \log\lambda_n$, then

$|\mathrm{Re}\, P^{1/2}(x_2) - P_2^{1/2}(x_2, \lambda_n, t^*\lambda_n)| < \pi/4$ in $0 \leqq x_2 \leqq 1$, and hence it follows from (3.4) that $|\sin \Theta(1, 0)| > 8^{-1/2} \exp \Theta^*(1, 0)$; (ii) when $\mathrm{Re}\, \mu = t^*\lambda_n$ and $|\mathrm{Im}\, \mu| > \lambda_n^{1/2} \log \lambda_n$, or when $|\mathrm{Im}\, \mu| = \lambda_n$, then $|\mathrm{Im}\, P^{1/2}(x_2)| > d^* \log \lambda_n > d^*$ in $0 \leqq x_2 \leqq 1$ (here $d^* = B^*/2^{5/4}[B^\dagger(1 + c + 2b_2)]^{1/2}$, where c is defined in the notation above), and hence $|\sin \Theta(1, 0)| \geqq \sinh \Theta^*(1, 0) > [(1 - e^{-2d^*})/2] \exp \Theta^*(1, 0)$; and (iii) when $\mathrm{Re}\, \mu = t^\dagger \lambda_n$, $|\mathrm{Im}\, P^{1/2}(x_2)| > d^\dagger \lambda_n^{1/2} > d^\dagger$ in $0 \leqq x_2 \leqq 1$ ($d^\dagger = (B^* a_1)^{1/2}$), and hence $|\sin \Theta(1, 0)| > [(1 - e^{-2d^\dagger})/2] \exp \Theta^*(1, 0)$. Denoting by C the constant implied in the O symbol in (3.6), choose n^* ($> n_0$) large enough so that $8^{1/2}C/n^*$, $2C/(1 - e^{-2d^*})n^*$, and $2C/(1 - e^{-2d^\dagger})n^*$ are all less than $1/2$, and sufficiently large so as to ensure that analogous results also hold for W_1. Then it follows from (3.6) than when $n > n^*$,

$$W_2(\lambda_n, \mu) = -\phi(1) = -P^{-1/4}(0)P^{-1/4}(1) \sin \Theta(1, 0)\{1 + O(1/n)\} \neq 0. \tag{3.7}$$

We wish next to examine the behavior of $W_2(\lambda_n, \mu)$ in the interval X: $\mathrm{Re}\, \mu \leqq t^\dagger \lambda_n$, $\mathrm{Im}\, \mu = 0$. Indeed, if we employ the same notation as above *with* μ *now being a point of* X, then it is not difficult to verify that $\phi(x_2)$ may be expressed in the form (3.6), with the constant implied in the O symbol, which we shall denote by C^*, being *independent* of x_2, μ, n, and of the p_j and θ_j for $j = 1, 2$. Since $\mathrm{Im}\, P^{1/2}(x_2) > d^\dagger$ in $0 \leqq x_2 \leqq 1$ (where $d^\dagger$ is defined in the above paragraph), it follows that $|\sin \Theta(1, 0)| > [(1 - e^{-2d^\dagger})/2] \exp \Theta^*(1, 0)$. Now choose $n^\dagger$ ($> n_0$) large enough so that $2C^*/(1 - e^{-2d^\dagger})n^\dagger < 1/2$ and sufficiently large so as to ensure that an analogous result holds for $W_1(\lambda_n, \mu)$ in the interval $\mathrm{Re}\, \mu \geqq (a_1 + b_2)\lambda_n$, $\mathrm{Im}\, \mu = 0$. Then for $n > n^\dagger$, $W_2(\lambda_n, \mu)$ ($= -\phi(1)$) cannot vanish in X.

Putting $n_1 = \max\{n^*, n^\dagger\}$, *we shall henceforth assume that* $n > n_1$. We note from (3.7) and the above discussion that $|1 - H(\mu)| < 1/2$ for $\mu \in \Gamma$, where

$$H = W/P^{-1/4}(0)P^{-1/4}(1) \sin \Theta(1, 0) \tag{3.8}$$

and $W = W(\mu) = -W_2(\lambda_n, \mu)$. Then to complete the proof of our theorem we wish to apply the *principle of the argument* to the function W with respect to the closed curve Γ. To this end we shall abandon for the remainder of this proof the convention adopted above to define $P^\nu(x_2)$ and instead redefine this expression in the following way. Describing Γ by means of the parameter s (see (3.2)), putting

$$P(x_2, s) = P_2(x_2, \lambda_n, \mu(s)) \tag{3.9}$$

for $(x, s) \in D$, where D is the rectangle defined by the inequalities $0 \leqq x_2 \leqq 1$, $0 \leqq s \leqq 5$, and observing that $P(0, 0) > 0$, we shall then agree to take $\arg P(0, 0) = 0$. This convention, together with the continuity of the function concerned, determines $\arg P(x_2, s)$ in D, and the expression $P^\nu(x_2, s)$ (ν real) is unambiguous provided that it is interpreted according to (3.3). It then follows that $P^{-1/4}(0, s)$, $P^{-1/4}(1, s)$ are continuous in $0 \leqq s \leqq 5$ and both their arguments are equal to 0 when $s = 0$ and to $-\pi/2$ when $s = 5$. Moreover, if we put

$$z(s) = \int_0^1 P^{1/2}(x_2, s)\, dx_2, \tag{3.10}$$

then we can show that as s runs from 0 to 5, $z(s)$ traces out a piecewise smooth simple arc in the z-plane, which we shall denote by Γ_1, such that $z(0) = -z(5) =$

$p_2 n\pi/2$ and $\operatorname{Im} z(s) > 0$ in $0 < s < 5$. It is also clear from (3.8) that

$$H^*(s) = H(\mu(s)) = W^*(s)/P^{-1/4}(0, s)P^{-1/4}(1, s) \sin z(s) \tag{3.11}$$

for $0 \leqq s \leqq 5$, where $W^*(s) = W(\mu(s))$. Hence from this fact and the foregoing results, we are led to assign to both $\arg\{\sin z(0)\}$ and $\arg W^*(0)$ the value 0 or π according to whether $(p_2 n - 1)/2$ is even or odd, and then determining $\arg\{\sin z(s)\}$ and $\arg W^*(s)$ for $0 \leqq s \leqq 5$ by continuity, to define $\arg H^*(s)$ in this interval by means of the equation

$$\arg W^*(s) = \arg P^{-1/4}(0, s) + \arg P^{-1/4}(1, s) + \arg\{\sin z(s)\} + \arg H^*(s). \tag{3.12}$$

In light of our above results it is easy to see from (3.11) that $H^*(5)$ is real, and hence we conclude from (3.12) that $\arg H^*(0) = \arg H^*(5) = 0$ and the variation of $\arg W$ around the contour Γ is equal to $[-\pi + \arg\{\sin z(5)\} - \arg\{\sin z(0)\}]$, which clearly is just $[-\pi - i \int_{\Gamma_1} \cot z \, dz]$.

Now let Γ_2 denote the path traced out by z as μ traces out Γ twice in succession. That is to say, if we extend the definition of $\mu(s)$ given in (3.2) to the interval $0 \leqq s \leqq 10$ by putting $\mu(s) = \mu(s-5)$ for $5 < s \leqq 10$, if in the rectangle $0 \leqq x_2 \leqq 1$, $0 \leqq s \leqq 10$, we define $P(x_2, s)$ by means of (3.9), and with $\arg P(0, 0) = 0$, determine $\arg P(x_2, s)$ by continuity and define $P^\nu(x_2, s)$ according to (3.3), and if we define $z(s)$ by means of (3.10), then we denote by Γ_2 the path traced out by $z(s)$ as s runs from 0 to 10. It is easy to show that Γ_2 is a piecewise smooth Jordan curve and that $\int_{\Gamma_2} \cot z \, dz = 2 \int_{\Gamma_1} \cot z \, dz$. Hence it follows that the variation of $\arg W$ around Γ is just $(p_2 n - 1)\pi$. The proof of the theorem (for W_2) is then completed by applying the principle of the argument, by utilizing the fact that $t^\dagger \lambda_n < \mu_0^*(\lambda_n)$, and by appealing to the analogue of (2.1) for the system (1.3)–(1.4).

Recalling again the terminology of § 2, we have next

THEOREM 3.2. *If* $n > (n_1 + 2)$, *then* $\lambda_{n-2} < \lambda_{j,k} < \lambda_{n+2}$, $t^* \lambda_{n-2} < \mu_{j,k} < t^* \lambda_{n+2}$, *and* $|t^* - \mu_{j,k}/\lambda_{j,k}| < 60 b_2/n$ *for* $(N_1 - p_1) < j \leqq (N_1 + p_1)$, $(N_2 - p_2) < k \leqq (N_2 + p_2)$.

Proof. From Theorem 3.1 we have

$$\mu_0^*(\lambda_{n+2}) < \cdots < \mu^*_{N_2+p_2}(\lambda_{n+2}) < t^* \lambda_{n+2} < \mu^*_{N_2+p_2+1}(\lambda_{n+2}) < \cdots,$$

$$\mu_0(\lambda_{n+2}) > \cdots > \mu_{N_1+p_1}(\lambda_{n+2}) > t^* \lambda_{n+2} > \mu_{N_1+p_1+1}(\lambda_{n+2}) > \cdots,$$

$$\mu_0^*(\lambda_{n-2}) < \cdots < \mu^*_{N_2-p_2}(\lambda_{n-2}) < t^* \lambda_{n-2} < \mu^*_{N_2-p_2+1}(\lambda_{n-2}) < \cdots,$$

$$\mu_0(\lambda_{n-2}) > \cdots > \mu_{N_1-p_1}(\lambda_{n-2}) > t^* \lambda_{n-2} > \mu_{N_1-p_1+1}(\lambda_{n-2}) > \cdots.$$

Hence in light of the results of § 2 and the definition of λ_n, our theorem follows immediately.

We shall henceforth assume that n is fixed at a value exceeding $\max\{(n_1 + 2), 120 B^\dagger b_2 (1 + \tan\theta_2)/B^* \delta\}$.

Referring to Theorem 3.2, we now note that $60 b_2/n < \delta/2$. Hence if we choose the positive integer p to have the same parity as p_1 and to satisfy

$$2n/p < \min\{(1 + \tan\theta_2)^{-1} \tan\theta_1, (\tan\theta_2 - \tan\theta_1)\},$$

and if we again recall the definitions of the r_i (see statement 7 of the notation given at the beginning of this subsection), then we have

THEOREM 3.3. *Let j, k be any positive integers for which $(j, k) \in \Omega$ and $j \geqq (np - r_1)/2$. Then $(\pi j/2)^2/B^{\dagger}b_2 < \lambda_{j,k} < (3\pi j)^2/B^{*}\delta$ and $a_2 + \delta/2 < \mu_{j,k}/\lambda_{j,k} < b_1 - \delta/2$.*

Proof. Let i be a positive integer and put $p_i^{*} = p + 2(i-1)$. For each positive integer m let $p_m^{\dagger} = p^{\dagger} + 2(m-1)$, where $p^{\dagger}$ is 1 or 2 according to whether p_2 is odd or even, and denote by $(m_i - 1)$ the largest value of m for which $p_m^{\dagger}/p_i^{*} < \tan\theta_1$ and by $(M_i + 1)$ the smallest value of m for which $p_m^{\dagger}/p_i^{*} > \tan\theta_2$. In light of the definitions of n and p above, it is clear that $9 \leqq m_i < M_i \leqq [1 + (p_i^{*}\tan\theta_2 - 1)/2]$, $2n < \min\{p_i^{*}, p_i^{*}\tan\theta_1\}$, and $p_i^{*}\tan\theta_1 \leqq p_{m_i}^{\dagger} < \cdots < p_{M_i}^{\dagger} \leqq p_i^{*}\tan\theta_2$. Putting $N_1(i) = (np_i^{*} - r_1)/2$, $N_2(i, m) = (np_m^{\dagger} - r_2)/2$, $m = m_i, \cdots, M_i$, denote by $G_m(i)$, $m_i \leqq m \leqq M_i$, the set of tuples $(N_1(i) + s_1, N_2(i, m) + s_2)$, where $s_1 = 0, \cdots, (n-1)$, and (a) $s_2 = -2n, (-2n+1), \cdots, (n-1)$ if $m = m_i$, (b) $s_2 = 0, \cdots, (n-1)$ if $m_i < m < M_i$, and (c) $s_2 = 0, \cdots, p_{M_i}^{\dagger}$ if $m = M_i$. It is important to observe that the $G_m(i)$ are all disjoint and

$$[N_2(i, m_i) - 2n] < N_1(i)\tan\theta_1 < [N_1(i) + n]\tan\theta_2 < [N_2(i, M_i) + p_{M_i}^{\dagger}]$$

(these inequalities follow easily from the definitions of n and p). Finally, we shall denote by $G(i)$ the union of the $G_m(i)$, $m = m_i, \cdots, M_i$, and by G the union of the disjoint sets $G(i)$, with i running from 1 to ∞.

It is clear from the assumptions made in our theorem that $(j, k) \in G$. Hence $(j, k) \in G_m(i)$ for some i and m. The assertion of the theorem now follows easily from Theorem 3.2 if we take $p_1 = p_i^{*}$, $p_2 = p_m^{\dagger}$, $N_1 = N_1(i)$, and $N_2 = N_2(i, m)$.

3.2. Asymptotic formulas. *Throughout this subsection we shall suppose that j, k are any positive integers for which $(j, k) \in \Omega$ and $j \geqq (np - r_1)/2$* (we refer to the statements preceding Theorem 3.3 for definitions and assumptions). Then as a consequence of Theorem 3.3 we are now in a position to establish asymptotic formulas for $\lambda_{j,k}$, $\mu_{j,k}$, and $\psi_{j,k}$ as $j \to \infty$. Accordingly, let us return again to the function $g(t)$ defined in § 3.0 and observe that when $a_2 < t < b_1$,

$$2h_1^2(t)g'(t) = \iint_{I^2} [B_1(x_1)f(t, x_1, x_2) + B_2(x_2)/f(t, x_1, x_2)]\, dx_1\, dx_2,$$

where $' = d/dt$ and $f(t, x_1, x_2) = P_2^{1/2}(x_2, 1, t)/P_1^{1/2}(x_1, 1, t)$. Thus $g'(t) \geqq B^{*}h_1^{-2}(t) > B^{*}/B^{\dagger}b_2$ for $a_2 < t < b_1$. By utilizing this fact and taking into account the assumptions concerning n and p made in the statements preceding Theorem 3.3, it is not difficult to verify that if we put

$$\begin{aligned} \nu_r &= 0 \quad \text{if } \alpha_r \neq 0 \text{ and } \beta_r \neq \pi, \\ &= \tfrac{1}{2} \quad \text{if } \alpha_r \neq 0 \text{ and } \beta_r = \pi \text{ or if } \alpha_r = 0 \text{ and } \beta_r \neq \pi, \\ &= 1 \quad \text{if } \alpha_r = 0 \text{ and } \beta_r = \pi, \end{aligned} \tag{3.13}$$

for $r = 1, 2$ (see (1.2) and (1.4)), then there is precisely one value of t in $[a_2, b_1]$ for which

$$g(t) = (k + \nu_2)/(j + \nu_1), \tag{3.14}$$

and if we henceforth agree to denote this value of t by $t_{j,k}$, then

$$t_1-\delta/4<t_{j,k}<t_2+\delta/4$$

(we refer to § 3.0 for terminology).

Notation. For $r=1, 2$, $0\leqq x_r\leqq 1$, $0<\gamma_r<\pi$, and $a_2<t<b_1$, let

$$X_r(x_r; t)=\int_0^{x_r} P_r^{1/2}(\tau_r, 1, t)\,d\tau_r, \qquad X_r^*(x_r, t)=\int_0^{x_r} B_r(\tau_r)P_r^{-1/2}(\tau_r, 1, t)\,d\tau_r,$$

$$X_r^\dagger(x_r, \gamma_r, t)=[\,P_r^{1/2}(x_r, 1, t)\cot\gamma_r+P_r'(x_r, 1, t)/4P_r^{3/2}(x_r, 1, t)], \qquad '=d/dx_r,$$

$$X_r^{\#}(x_r, t)=-2^{-1}\int_0^{x_r} Q_r(\tau_r, t)P_r^{-1/2}(\tau_r, 1, t)\,d\tau_r,$$

where $Q_r(\tau_r, t)=q_r(\tau_r)+P_r^{1/4}(\tau_r, 1, t)[d^2P_r^{-1/4}(\tau_r, 1, t)/d\tau_r^2]$,

$$\begin{aligned}
c_r^*(t)&=[X_r^\dagger(0, \alpha_r, t)-X_r^\dagger(1, \beta_r, t)+X_r^{\#}(1, t)] && \text{if } \alpha_r\neq 0 \text{ and } \beta_r\neq\pi,\\
&=[X_r^\dagger(0, \alpha_r, t)+X_r^{\#}(1, t)] && \text{if } \alpha_r\neq 0 \text{ and } \beta_2=\pi,\\
&=[-X_r^\dagger(1, \beta_r, t)+X_r^{\#}(1, t)] && \text{if } \alpha_r=0 \text{ and } \beta_r\neq\pi,\\
&=X_r^{\#}(1, t) && \text{if } \alpha_r=0 \text{ and } \beta_r=\pi,
\end{aligned}$$

$$c_r(t)=c_r^*(t)/h_1^2(t)g'(t) \quad (\,'=d/dt), \qquad d_r(t)=X_r^*(1, t),$$

$$d_r^*(t)=td_r(t)+(-1)^{r-1}h_r(t), \quad D(t)=\iint_{I^2}\left\{\Delta(x_1, x_2)\Big/\prod_{r=1}^{2} P_r^{1/2}(x_r, 1, t)\right\}dx_1\,dx_2,$$

$$\begin{aligned}
Y_r^*(x_r, t)&=[c_1(t)d_2(t)+c_2(t)d_1(t)]X_r(x_r, t)\\
&\quad+(-1)^{r-1}[c_1(t)h_2(t)-c_2(t)h_1(t)]X_r^*(x_r, t),\\
Y_r^\dagger(x_r, t)&=[-2^{-1}Y_r^*(x_r, t)+X_r^\dagger(0, \alpha_r, t)+X_r^{\#}(x_r, t)] \quad \text{if } \alpha_r\neq 0,\\
&=[2^{-1}Y_r^*(x_r, t)-X_r^{\#}(x_r, t)] \quad \text{if } \alpha_r=0,
\end{aligned}$$

$$Y_r(x_r, j, k)=N_{j,k}^{-1}Y_r^\dagger(x_r, t_{j,k}), \qquad Z_r(x_r, j, k)=N_{j,k}X_r(x_r, t_{j,k}),$$

where $N_{j,k}=[(j+\nu_1)\pi/h_1(t_{j,k})]$,

and finally, writing E_m for $E_m(x_1, x_2, j, k)$, $m=1, \cdots, 4$, let: (a) $E_1=1$, $E_2=Y_2(x_2, j, k)$, $E_3=Y_1(x_1, j, k)$, and $E_4=0$ if $\alpha_1\neq 0$ and $\alpha_2\neq 0$, (b) $E_1=Y_2(x_2, j, k)$, $E_2=1$, $E_3=0$, and $E_4=Y_1(x_1, j, k)$ if $\alpha_1\neq 0$ and $\alpha_2=0$, (c) $E_1=Y_1(x_1, j, k)$, $E_2=0$, $E_3=1$, and $E_4=Y_2(x_2, j, k)$ if $\alpha_1=0$ and $\alpha_2\neq 0$, (d) $E_1=0$, $E_2=Y_1(x_1, j, k)$, $E_3=Y_2(x_2, j, k)$, and $E_4=1$ if $\alpha_1=\alpha_2=0$.

THEOREM 3.4. *It is the case that*

$$\lambda_{j,k}=[(j+\nu_1)\pi/h_1(t_{j,k})]^2[1+(h_1(t_{j,k})/(j+\nu_1)\pi)^2\langle c_1(t_{j,k})d_2(t_{j,k})+c_2(t_{j,k})d_1(t_{j,k})\rangle+o(1/j^2)], \tag{3.15}$$

$$\mu_{j,k}=[(j+\nu_1)\pi/h_1(t_{j,k})]^2[t_{j,k}+(h_1(t_{j,k})/(j+\nu_1)\pi)^2\langle c_1(t_{j,k})d_2^*(t_{j,k})+c_2(t_{j,k})d_1^*(t_{j,k})\rangle+o(1/j^2)], \tag{3.16}$$

$$\psi_{j,k}(x_1, x_2) = 2D^{-1/2}(t_{j,k}) \prod_{r=1}^{2} P_r^{-1/4}(x_r, 1, t_{j,k})$$
$$\cdot \Big[\prod_{r=1}^{2} \cos Z_r(x_r, j, k)\{E_1(x_1, x_2, j, k) + o(1/j)\}$$
$$+ \cos Z_1(x_1, j, k) \sin Z_2(x_2, j, k)\{E_2(x_1, x_2, j, k) + o(1/j)\} \tag{3.17}$$
$$+ \sin Z_1(x_1, j, k) \cos Z_2(x_2, j, k)\{E_3(x_1, x_2, j, k) + o(1/j)\}$$
$$+ \prod_{r=1}^{2} \sin Z_r(x_r, j, k)\{E_4(x_1, x_2, j, k) + o(1/j)\} \Big]$$

as $j \to \infty$, $(j, k) \in \Omega$. *This last result holds uniformly for* $0 \leqq x_r \leqq 1$, $r = 1, 2$.

Proof. We shall only prove the theorem for the case $\alpha_r = 0$, $\beta_r = \pi$ for $r = 1, 2$; the other cases can be similarly treated. Then putting

$$P_{j,k}(x_2) = P_2(x_2, \lambda_{j,k}, \mu_{j,k}) \quad \text{and} \quad \phi_{j,k}(x_2) = \phi_2(x_2, \lambda_{j,k}, \mu_{j,k}),$$

we may argue with (3.5), the Gronwall lemma, and the results of Theorem 3.3 to verify that for $0 \leqq x_2 \leqq 1$,

$$\phi_{j,k}(x_2) = P_{j,k}^{-1/4}(0) P_{j,k}^{-1/4}(x_2)[\sin \Theta_{j,k}(x_2) + O(1/k)],$$

where

$$\Theta_{j,k}(x_2) = \int_0^{x_2} P_{j,k}^{1/2}(\tau)\, d\tau,$$

and similarly we can also verify that

$$\phi'_{j,k}(x_2) = P_{j,k}^{-1/4}(0) P_{j,k}^{1/4}(x_2)[\cos \Theta_{j,k}(x_2) + O(1/k)],$$

$(' = d/dx_2)$, with these results holding uniformly in $0 \leqq x_2 \leqq 1$. A standard argument now shows that as $k \to \infty$,

$$\int_0^1 P_2^{1/2}(x_2, \lambda_{j,k}, \mu_{j,k})\, dx_2 = (k+1)\pi + O(1/k), \tag{3.18}$$

and by repeating the same argument for ϕ_1, we may also show that

$$\int_0^1 P_1^{1/2}(x_1, \lambda_{j,k}, \mu_{j,k})\, dx_1 = (j+1)\pi + O(1/j) \tag{3.19}$$

and

$$g(\mu_{j,k}/\lambda_{j,k}) = ((k+1)/(j+1)) + O(1/j^2) \tag{3.20}$$

as $j \to \infty$. By appealing to the inverse function theorem, it is not difficult to deduce from (3.20) that

$$\mu_{j,k}/\lambda_{j,k} = t_{j,k} + O(1/j^2) \tag{3.21}$$

as $j \to \infty$, and hence it follows from (3.19) that

$$\lambda_{j,k} = N_{j,k}^2(1 + O(1/j^2)) \tag{3.22}$$

as $j \to \infty$, where

$$N_{j,k} = [(j+1)\pi/h_1(t_{j,k})] = [(k+1)\pi/h_2(t_{j,k})].$$

Turning again to (3.5), we may now appeal to the method of Horn [11] (see also [12, p. 272]), to arguments similar to those used in the proof of the Riemann–Lebesgue lemma, and to (3.21), (3.22) to establish that

$$\begin{aligned}\phi_{j,k}(x_2) = P_{j,k}^{-1/4}(0)P_{j,k}^{-1/4}(x_2)[&\sin\Theta_{j,k}(x_2)\{1+o(1/k)\}\\ &-\cos\Theta_{j,k}(x_2)\{N_{j,k}^{-1}X_2^{\#}(x_2, t_{j,k})+o(1/k)\}]\end{aligned} \tag{3.23}$$

as $k\to\infty$, uniformly in $0\leqq x_2\leqq 1$, where we refer to the statements immediately preceding this theorem for terminology. Hence if we now argue in a manner similar to the way in which we argued in arriving at (3.18), then we can also show that as $k\to\infty$,

$$\int_0^1 P_2^{1/2}(x_2, \lambda_{j,k}, \mu_{j,k})\,dx_2 = (k+1)\pi + N_{j,k}^{-1}c_2^*(t_{j,k})+o(1/k);$$

and similarly, we can also show that

$$\int_0^1 P_1^{1/2}(x_1, \lambda_{j,k}, \mu_{j,k})\,dx_1 = (j+1)\pi + N_{j,k}^{-1}c_1^*(t_{j,k})+o(1/j) \tag{3.24}$$

as $j\to\infty$. From these results is now follows that as $j\to\infty$,

$$\begin{aligned}g(\mu_{j,k}/\lambda_{j,k}) = &((k+1)/(j+1))\\ &-[1/(j+1)\pi]^2\langle c_1^*(t_{j,k})h_2(t_{j,k})-c_2^*(t_{j,k})h_1(t_{j,k})\rangle+o(1/j^2),\end{aligned}$$

and hence we conclude that

$$\mu_{j,k}/\lambda_{j,k} = t_{j,k} - N_{j,k}^{-2}\langle c_1(t_{j,k})h_2(t_{j,k})-c_2(t_{j,k})h_1(t_{j,k})\rangle+o(1/j^2) \tag{3.25}$$

as $j\to\infty$. The assertion of our theorem concerning $\lambda_{j,k}$ and $\mu_{j,k}$ now follows immediately from (3.24), (3.25), and the relation $2h_1^2(t)g'(t) = (d_1(t)h_2(t)+d_2(t)h_1(t))$.

To prove the assertion concerning $\psi_{j,k}$, we first observe from (3.21) and (3.22) that (3.23) remains valid if the expression $P_{j,k}^{-1/4}(0)P_{j,k}^{-1/4}(x_2)$ in the right-hand side of this equation is replaced by

$$N_{j,k}^{-1}P_2^{-1/4}(0, 1, t_{j,k})P_2^{-1/4}(x_2, 1, t_{j,k}).$$

Furthermore, from (3.25) and the estimate for $\lambda_{j,k}$ which we have just established it is easy to see that

$$\Theta_{j,k}(x_2) = Z_2(x_2, j, k)+(1/(2N_{j,k}))Y_2^*(x_2, t_{j,k})+o(1/k)$$

as $k\to\infty$, uniformly in $0\leqq x_2\leqq 1$. Hence it follows from (3.23) that

$$\begin{aligned}\phi_2(x_2, \lambda_{j,k}, \mu_{j,k}) = &N_{j,k}^{-1}P_2^{-1/4}(0, 1, t_{j,k})P_2^{-1/4}(x_2, 1, t_{j,k})\\ &\cdot[\sin Z_2(x_2, j, k)\{1+o(1/k)\}\\ &+\cos Z_2(x_2, j, k)\{E_3(x_1, x_2, j, k)+o(1/k)\}]\end{aligned}$$

as $k\to\infty$, uniformly in $0\leqq x_2\leqq 1$. Arguing in the same way with ϕ_1, we may easily establish that (see § 2.3 for terminology)

$$\psi_{j,k}^*(x_1, x_2) = N_{j,k}^{-2}\Big(\prod_{r=1}^{2}P_r^{-1/4}(0, 1, t_{j,k})P_r^{-1/4}(x_r, 1, t_{j,k})\Big)F(x_1, x_2, j, k)$$

as $j \to \infty$, uniformly in $0 \leqq x_r \leqq 1$, $r = 1, 2$, where $F(x_1, x_2, j, k)$ denotes the expression within the square brackets on the right-hand side of (3.17). From this result it is not difficult to verify that

$$\|\psi_{j,k}^*\|^{-1} = 2D^{-1/2}(t_{j,k})N_{j,k}^2 \prod_{r=1}^{2} P_r^{1/4}(0, 1, t_{j,k})\{1 + o(1/j)\}$$

as $j \to \infty$, and hence the assertion of our theorem concerning $\psi_{j,k}$ follows immediately.

As a consequence of Theorem 3.4 and the definition of Ω, we have

THEOREM 3.5. *There exists the positive number K, depending upon θ_1 and θ_2, such that $|\psi_{j,k}(x_1, x_2)| \leqq K$ for $(x_1, x_2) \in I^2$ and $(j, k) \in \Omega$.*

Remark. We note that the formulas given in Theorem 3.4 have been obtained to an accuracy determined by the conditions which we have imposed upon the coefficients of the differential equations (1.1) and (1.3). It is clear that by proceeding as above and using the method of Horn, we may further develop these formulas if the coefficients in our differential equations are suitably defined.

4. Some extensions.

4.0. Introduction. *Throughout this section we shall again employ the notation given in* § 3.0 *and denote by Ω_1 and Ω_2 the sectors in the (x, y)-plane defined by the inequalities $0 \leqq \theta < \theta_1$ and $\theta_2 < \theta \leqq \pi/2$, respectively.* Then the methods of the foregoing section do not generally suffice in establishing the analogue of Theorem 3.4 for each of the cases: (i) $(j, k) \in \Omega_1$, $j \to \infty$, (ii) $(j, k) \in \Omega_2$, $k \to \infty$. Indeed, the treatment of these cases may become somewhat intricate (see [7], [8]) and will, with certain exceptions that will be discussed below, be left for later papers. However, we might remark that as a consequence of the results of § 2, as well as from standard arguments, we do have some information at our disposal. For we know that C_j, $j \geqq 0$, has a continuously turning tangent whose slope is a mean among the values of A_1/B_1 and that C_j passes through the point $(0, \mu_j(0))$, where $\infty > \mu_0(0) > \mu_1(0) > \cdots$, and

$$\mu_j(0) = -\left[(j + \nu_1)\pi \Big/ \int_0^1 B_1^{1/2}\, dx_1\right]^2 (1 + O(1/j^2)) \quad \text{as } j \to \infty$$

(we refer to (3.13) for the definitions of the ν_r). We also know that C_k^*, $k \geqq 0$, has a continuously turning tangent whose slope is a mean among the values of A_2/B_2 and that C_k^* passes through the point $(0, \mu_k^*(0))$, where $-\infty < \mu_0^*(0) < \mu_1^*(0) < \cdots$, and

$$\mu_k^*(0) = \left[(k + \nu_2)\pi \Big/ \int_0^1 B_2^{1/2}\, dx_2\right]^2 (1 + O(1/k^2)) \quad \text{as } k \to \infty. \tag{4.1}$$

Hence a simple calculation now shows that when $(j, k) \in \Omega_1$ and j is sufficiently large,

$$\begin{gathered} \frac{1}{4B^\dagger(b_2 - a_1)} < \frac{\lambda_{j,k}}{(j\pi)^2} < \frac{8B^\dagger \sec^2 \theta_1}{(b_1 - a_2)(B^*)^2}, \\ a_1 - \frac{4B^\dagger(b_2 - a_1)|\mu_0^*(0)|}{(j\pi)^2} < \frac{\mu_{j,k}}{\lambda_{j,k}} < a_2 + \frac{8B^\dagger(b_2 - a_1)\tan^2 \theta_1}{B^*}, \end{gathered} \tag{4.2}$$

with similar results also holding in Ω_2 for all large k.

Assumptions. 1. We shall henceforth suppose that $A_2(x_2)/B_2(x_2)$ is constant in $0 \leqq x_2 \leqq 1$ and put $a = a_1 = a_2$ (see § 2.0). We now note that the C_k^* are precisely the straight lines

$$-\lambda a + \mu = \mu_k^*(0), \qquad k \geqq 0. \tag{4.3}$$

We also note from the definitions given in § 3.0 that now $\theta_1^* = 0$.

2. We shall henceforth suppose that

$$0 < \theta_1 < \tan^{-1}\left\{\frac{\gamma B^*(b_1 - a)}{2B^\dagger(b_2 - a)}\right\},$$

where $\gamma = \min\{1, (b_2 - a)^{1/2}\}$ and the principal branch of the inverse tangent is taken.

3. We shall henceforth suppose that $(j, k) \in \Omega_1$.

It now follows from (4.2) that

$$\mu_{j,k}/\lambda_{j,k} < (a + b_1)/2 \tag{4.4}$$

for all large j.

Notation. Throughout this section we let:

1. Ω^* and $\Omega^\dagger$ denote the subsets of Ω_1 composed of points (x, y) satisfying $y^4 \geqq x \tan\theta_1$ and $y^4 < x \tan\theta_1$, respectively;

2. ν_r, $r = 1, 2$, denote those expressions defined in (3.13).

In light of the above results we are now in a position to derive asymptotic formulas for $\lambda_{j,k}$ and $\mu_{j,k}$ as $j \to \infty$; and to this end we shall investigate separately the behavior of these expressions in each of the sets Ω^* and $\Omega^\dagger$.

4.1. Asymptotic formulas in Ω^*. *In this subsection we shall suppose that* $(j, k) \in \Omega^*$ *and put* $\sigma_{j,k} = (k + \nu_2)/(j + \nu_1)$. Then in order to obtain estimates for $\lambda_{j,k}$ and $\mu_{j,k}$ as $j \to \infty$, we may utilize (4.2)–(4.4) and arguments similar to those used in the proof of Theorem 3.4 to show that

$$g(\mu_{j,k}/\lambda_{j,k}) = \sigma_{j,k} + O(1/jk) \tag{4.5}$$

as $j \to \infty$. To deal with (4.5) we must turn to the inverse function theorem. Accordingly, let us firstly extend $h_1(t)$ to an analytic function (of the complex variable t) in the half-plane $\operatorname{Re} t < b_1$ by putting here $h_1(t) = \int_0^1 P_1^{1/2}(x_1, 1, t)\, dx_1$ (recall from (3.3) the convention for dealing with fractional powers of the P_r). Then it is clear that in $\operatorname{Re} t < b_1$ the function $G(t) = g^2(t)$ is analytic and may be expanded about the point $t = a$ in the form $G(t) = \sum_{m=1}^{\infty} b_m^*(t - a)^m$, $|t - a| < (b_1 - a)$, where $b_1^* = [(\int_0^1 B_2^{1/2}\, dx_2)/h_1(a)]^2$ and $b_m^* > 0$ for $m \geqq 1$. Putting $M = B^\dagger/B^*$ and $R = (b_1 - a)/2$, we may argue with the inverse function theorem to show that when $|w| < R_1 = (Rb_1^*)^2/(6M)$, the equation $G(t) = w$ has precisely one solution in $|t - a| < R^* = 3R_1/(2b_1^*)$ $(\leqq R/8)$ and this solution may be expressed in the form

$$t - a = H(w) = \sum_{m=1}^{\infty} b_m^\dagger w^m, \tag{4.6}$$

$|w| < R_1$, where the $b_m^\dagger$ are real, $b_1^\dagger = 1/b_1^*$, and $|b_m^\dagger| \leqq R^*/R_1^m$ for $m \geqq 1$. Observing that $0 < \sigma_{j,k} < 2 \tan\theta_1$ and $4 \tan^2\theta_1 < R_1/4$, we therefore conclude that there is precisely one value of t in the disc $|t - a| < R^*$ satisfying $G(t) = \sigma_{j,k}^2$. It is not

difficult to verify that this value of t is also a point of the interval $[a, b_1]$, and hence it follows that (3.14) has precisely one solution in this interval.

We shall henceforth denote the solution of (3.14) *(for* $a \leqq t \leqq b_1$*) by* $t_{j,k}$

We observe from (4.6) that

$$t_{j,k} - a = H(\sigma_{j,k}^2), \tag{4.7}$$

while if we put

$$H(w) = wH^*(w), \tag{4.8}$$

then a simple calculation shows that

$$b_1^\dagger/2 < H^*(\sigma_{j,k}^2) < 3b_1^\dagger/2, \tag{4.9}$$

and hence it follows that

$$\frac{B^*(\tan\theta_1)^{1/2}}{8B^\dagger j^{3/2}} < \frac{t_{j,k}-a}{b_1-a} < \frac{1}{64}. \tag{4.10}$$

In light of the foregoing results, it is now a simple matter to argue with (4.5) and the series $H((w_1+w_2)^2)$ to show that

$$\mu_{j,k}/\lambda_{j,k} = t_{j,k} + O(1/j^2) \quad \text{as } j \to \infty.$$

Arguments similar to those used in the proof of Theorem 3.4 suffice at this point to establish our main results which will be given in Theorem 4.1 below.

Notation. 1. For $|w| \leqq 2\tan\theta_1$, let $H_1^*(w) = \sum_{m=1}^{\infty} 2mb_m^\dagger w^{2m-2}$. It is not difficult to verify that

$$2b_1^\dagger/9 < |H_1^*(w)| < 4b_1^\dagger. \tag{4.11}$$

2. For $0 \leqq x_2 \leqq 1$, $0 < \gamma_2 < \pi$, and $a < t < g^{-1}(2\tan\theta_1)$, where g^{-1} denotes the inverse of g, let:

$$X_2(x_2, t) = \int_0^{x_2} B_2^{1/2}(\tau)\, d\tau, \qquad h_2^* = X_2(1, t),$$

(here t is introduced in order to conform with the notation of Theorem 3.4),

$$X_2^\dagger(x_2, \gamma_2, t) = [B_2^{-1/2}(x_2)\cot\gamma_2 + B_2'(x_2)/4B_2^{3/2}(x_2)], \qquad ' = d/dx_2,$$

$$X_2^{\#}(x_2, t) = -2^{-1}\int_0^{x_2} Q_2(\tau)B_2^{-1/2}(\tau)\, d\tau,$$

$$\text{where } Q_2(\tau) = q_2(\tau) + B_2^{1/4}(\tau)[d^2B_2^{-1/4}(\tau)/d\tau^2],$$

$X_2^*(x_2, t)$, $d_2(t)$, $d_2^*(t)$, and $c_2^*(t)$ be defined as in the notation immediately preceding Theorem 3.4, except that in defining this latter expression we are now to use our new definitions of $X_2^\dagger$ and $X_2^{\#}$,

$$c_2(t) = \frac{h_2^* H_1^*(g(t))c_2^*(t)}{h_1^3(t)}, \qquad D(t) = \iint_{I^2} \left\{\frac{\Delta(x_1, x_2)}{P_1^{1/2}(x_1, 1, t)B_2^{1/2}(x_2)}\right\} dx_1\, dx_2,$$

$$Y_2^*(x_2, t) = [2c_2^*(t)/h_2^*]X_2(x_2, t),$$

$Y_2^{\dagger}(x_2, t)$ be defined as in the notation immediately preceding Theorem 3.4 (using now our new definitions of Y_2^*, $X_2^{\dagger}$, and $X_2^{\#}$),

$$Y_2(x_2, j, k) = N_k^{-1} Y_2^{\dagger}(x_2, t_{j,k}), \qquad Z_2(x_2, j, k) = N_k X_2(x_2, t_{j,k}),$$

where $N_k = [(k+\nu_2)\pi/h_2^*]$,

and finally, writing e_m for $e_m(j, k)$, $m = 1, \cdots, 4$, let (i) $e_1 = e_2 = 1/k$, $e_3 = e_4 = 1/j$ if $\alpha_1 \neq 0$, (ii) $e_1 = e_2 = 1/j$, $e_3 = e_4 = 1/k$ if $\alpha_1 = 0$.

3. Let c_1, d_1, d_1^*, Z_1, and E_m, $m = 1, \cdots, 4$, be defined as in the notation immediately preceding Theorem 3.4, except now we are to use our new definition of Y_2 in defining the E_m (note also that in defining Y_1^* and hence Y_1, we are to use our new definition of c_2).

THEOREM 4.1. *If $A_2(x_2)/B_2(x_2)$ is constant in $0 \leqq x_2 \leqq 1$, then $\lambda_{j,k}$ and $\mu_{j,k}$ are given by the right-hand sides of* (3.15) *and* (3.16), *respectively, and*

$$\begin{aligned}\psi_{j,k}(x_1, x_2) = 2D^{-1/2}(t_{j,k})P_1^{-1/4}(x_1, 1, t_{j,k})B_2^{-1/4}(x_2)\\ \cdot\Big[\prod_{r=1}^{2}\cos Z_r(x_r, j, k)\{E_1(x_1, x_2, j, k)+o(e_1(j, k))\}\\ +\cos Z_1(x_1, j, k)\sin Z_2(x_2, j, k)\{E_2(x_1, x_2, j, k)+o(e_2(j, k))\}\\ +\sin Z_1(x_1, j, k)\cos Z_2(x_2, j, k)\{E_3(x_1, x_2, j, k)+o(e_3(j, k))\}\\ +\prod_{r=1}^{2}\sin Z_r(x_r, j, k)\{E_4(x_1, x_2, j, k)+o(e_4(j, k))\}\Big]\end{aligned}$$

as $j \to \infty$, $(j, k) \in \Omega^*$. This last result holds uniformly for $0 \leqq x_r \leqq 1$, $r = 1, 2$.

Remark. Writing $s_{j,k}$ for $c_1(t_{j,k})d_2(t_{j,k})$ and $s_{j,k}^*$ for $c_1(t_{j,k})d_2^*(t_{j,k})$, we observe that each of the formulas given in Theorem 4.1 contains the expression $s_{j,k}$ or $s_{j,k}^*$. And although $d_2(t_{j,k})$ and $d_2^*(t_{j,k})$ contain the factor $(t_{j,k}-a)^{-1/2}$, it is important to observe that $c_1(t_{j,k})$ contains the factor $1/g'(t_{j,k})$ which is easily shown to be just $\sigma_{j,k}H_1^*(\sigma_{j,k})$. Hence it follows from (4.7)–(4.8) and the definitions of the terms involved (see the statements preceding Theorem 3.4) that $s_{j,k} = L_{j,k}c_1^*(t_{j,k})$, $s_{j,k}^* = L_{j,k}^*c_1^*(t_{j,k})$, where

$$L_{j,k} = \frac{h_2^*H_1^*(\sigma_{j,k})}{h_1^2(t_{j,k})[H^*(\sigma_{j,k}^2)]^{1/2}}$$

and $L_{j,k}^* = aL_{j,k}$. In light of (4.9)–(4.11) we therefore conclude that the absolute values of $s_{j,k}$ and $s_{j,k}^*$ remain less than some bound independent of j and k.

4.2. Asymptotic formulas in $\Omega^{\dagger}$. *In this subsection we suppose that $(j, k) \in \Omega^{\dagger}$.* Then to deal with this case, we shall have to modify the definitions given in the statements preceding Theorem 3.4. Hence referring to these definitions, we observe that the expressions c_1^*, d_1, and d_1^* have only been defined for the interval $a < t < b_1$. We now extend the definitions of these expressions to the interval $[a, b_1)$ by continuity.

Let $P_{j,k}(x_1) = P_1(x_1, \lambda_{j,k}, \mu_{j,k})$ and $N_j = (j+\nu_1)\pi/h_1(a)$. Then from (4.1)–(4.4) and arguments similar to those used in the proof of Theorem 3.4, it is not difficult to verify that $\mu_{j,k}/\lambda_{j,k} = a + O(1/j^{3/2})$, $\int_0^1 P_{j,k}^{1/2}(x_1)\,dx_1 =$

$(j+\nu_1)\pi+O(1/j)$, and hence $\lambda_{j,k}=N_j^2[1+O(1/j^{3/2})]$ as $j\to\infty$. From this last result and an argument involving the method of Horn, we may next verify that

$$\mu_{j,k}/\lambda_{j,k}=a+\rho_{j,k}/N_j^{3/2}+O(1/j^3),$$

$$\int_0^1 P_{j,k}^{1/2}(x_1)\,dx_1=(j+\nu_1)\pi+c_1^*(a)/N_j+o(1/j)$$

as $j\to\infty$, where $\rho_{j,k}=\mu_k^*(0)/N_j^{1/2}$ (observe from (4.1) that $\rho_{j,k}=O(1)$). Hence putting $\rho_{j,k}^\dagger=\rho_{j,k}d_1(a)/h_1(a)$, $\rho_{j,k}^*=\rho_{j,k}d_1^*(a)/h_1(a)$, and $c_1^\#(a)=2c_1^*(a)/h_1(a)$, it follows that

THEOREM 4.2. *If $A_2(x_2)/B_2(x_2)$ is constant in $0\leqq x_2\leqq 1$, then*

$$\lambda_{j,k}=\left[\frac{(j+\nu_1)\pi}{h_1(a)}\right]^2\left[1+\left(\frac{h_1(a)}{(j+\nu_1)\pi}\right)^{3/2}\rho_{j,k}^\dagger+\left(\frac{h_1(a)}{(j+\nu_1)\pi}\right)^2 c_1^\#(a)+o\left(\frac{1}{j^2}\right)\right],$$

$$\mu_{j,k}=\left[\frac{(j+\nu_1)\pi}{h_1(a)}\right]^2\left[a+\left(\frac{h_1(a)}{(j+\nu_1)\pi}\right)^{3/2}\rho_{j,k}^*+\left(\frac{h_1(a)}{(j+\nu_1)\pi}\right)^2 ac_1^\#(a)+o\left(\frac{1}{j^2}\right)\right]$$

as $j\to\infty$, $(j,k)\in\Omega^\dagger$.

4.3. The eigenfunctions. It is clear from (4.3) that we are unable to obtain estimates for the $\psi_{j,k}$, $(j,k)\in\Omega^\dagger$, in the form given in Theorems 3.4 and 4.1. However we do have

THEOREM 4.3. *If $A_2(x_2)/B_2(x_2)$ is constant in $0\leqq x_2\leqq 1$, then there exists the positive number K such that $|\psi_{j,k}(x_1,x_2)|\leqq K$ for $(x_1,x_2)\in I^2$ and $(j,k)\in\Omega_1$.*

Proof. Let $u_{j,k}(x_1)=\phi_1(x_1,\lambda_{j,k},\mu_{j,k})/J_1(j,k)$, $v_{j,k}(x_2)=\phi_2(x_2,\lambda_{j,k}\mu_{j,k})/J_2(j,k)$, where $J_i(j,k)=(\int_0^1\phi_i^2(x_i,\lambda_{j,k},\mu_{j,k})\,dx_i)^{1/2}$ for $i=1,2$. Then by appealing to (4.2), (4.4), [10, Thm. 3.1], and arguing in a manner similar to that in [6, pp. 334–335], it is not difficult to verify that the absolute values of the $u_{j,k}(x_1)$ for $0\leqq x_1\leqq 1$ remain less than some bound independent of x_1, j, and k. In light of (4.3), it is also clear that a similar result holds for $v_{j,k}(x_2)$. The proof of the theorem now follows from these results.

4.4. Final remarks. To conclude our discussion for Ω_1 we wish to state that the formulas given in Theorems 4.1 and 4.2 may be further developed for suitable coefficients in the differential equations (1.1) and (1.3) (we refer to the remark made at the end of § 3 for further details). Also it is not difficult to verify that analogous results hold in Ω_2 if we assume that A_1/B_1 is constant in $0\leqq x_1\leqq 1$.

Appendix A. In the introduction to this paper we made the hypothesis that the coefficients of the differential equations (1.1) and (1.3) satisfied the condition $\Delta=(A_1B_2-A_2B_1)\neq 0$ in I^2. We then asserted that under this hypothesis it was possible to arrange matters, by means of a nonsingular transformation in the parameters λ and μ, so that the A_r, B_r, and Δ are positive for all values of x_1 and x_2 in I^2. We now prove this assertion.

Let S^1 denote the unit circle in E^2 with center at the origin and $u=(u_1,u_2)$ the points of S^1. Recalling the definition of the $P_r(x_r,\lambda,\mu)$ given in the introduction to § 3, let us define the mapping $f(u)=(f_1(u),f_2(u))$ of S^1 into E^2 in the following way: for each u in S^1 and for $r=1,2$, (a) let $f_r(u)$ denote the infimum of $P_r(x_r,u_1,u_2)$ in $0\leqq x_r\leqq 1$ if $P_r(x_r,u_1,u_2)>0$ at every point of this interval, (b) let

$f_r(u) = 0$ if $P_r(x_r, u_1, u_2)$ has at least one zero in $0 \leqq x_r \leqq 1$, and (c) let $f_r(u)$ denote the supremum of $P_r(x_r, u_1, u_2)$ in $0 \leqq x_r \leqq 1$ if $P_r(x_r, u_1, u_2) < 0$ at every point of this interval. It is not difficult to verify that $f(u)$ is a continuous mapping of S^1 into E^2, $f(u) \neq 0$ (since $\Delta \neq 0$), and $f(-u) = -f(u)$. From a result of Borsuk [5] it follows that the image of S^1 under f meets every ray emanating from the origin in E^2. Thus there exists the point $u^* = (u_1^*, u_2^*)$ of S^1 and the positive number d such that $u_2^* \neq u_1^*$, $f_1(u^*) > d$, and $f_2(u^*) < -d$. Let $v_r = -u_r^* + \varepsilon$ for $r = 1, 2$, where ε $(\neq 0)$ denotes a number chosen so that: (a) the supremum of $(|A_r(x_r)| + |B_r(x_r)|)$ in $0 \leqq x_r \leqq 1$ does not exceed $d/2|\varepsilon|$ for $r = 1, 2$, and (b) $\varepsilon(u_1^* - u_2^*)$ has the same sign as Δ. Hence if in (1.1) and (1.3) we put $\lambda = u_1^* \lambda^{\#} + v_1 \mu^{\#}$ and $\mu = u_2^* \lambda^{\#} + v_2 \mu^{\#}$, then the expressions $\lambda A_1(x_1) - \mu B_1(x_1)$ and $-\lambda A_2(x_2) + \mu B_2(x_2)$ become $\lambda^{\#} A_1^{\#}(x_1) - \mu^{\#} B_1^{\#}(x_1)$ and $-\lambda^{\#} A_2^{\#}(x_2) + \mu^{\#} B_2^{\#}(x_2)$, respectively, where the $A_r^{\#}$, $B_r^{\#}$, and $\Delta^{\#} = (A_1^{\#} B_2^{\#} - A_2^{\#} B_1^{\#})$ satisfy the conditions asserted.

REFERENCES

[1] L. V. AHLFORS, *Complex Analysis*, 2nd ed., McGraw-Hill, New York, 1966.

[2] F. V. ATKINSON, *Multiparameter Eigenvalue Problems*, vol. 1, Academic Press, New York, 1972.

[3] ———, *Multiparameter spectral theory*, Bull. Amer. Math. Soc., 74 (1968), pp. 1–27.

[4] ———, *Discrete and Continuous Boundary Problems*, Academic Press, New York, 1964.

[5] K. BORSUK, *Drei sätze über die n-dimensionale euklidische sphäre*, Fund. Math., 20 (1933), pp. 177–190.

[6] R. COURANT AND D. HILBERT, *Methods of Mathematical Physics*, vol. 1, Interscience, New York, 1953.

[7] M. FAIERMAN, *Asymptotic formulae for the eigenvalues of a two-parameter ordinary differential equation of the second order*, Trans. Amer. Math. Soc., 168 (1972), pp. 1–52.

[8] ———, *Asymptotic formulae for the eigenvalues of a two-parameter system of ordinary differential equation of the second order*, Canad. Math. Bull., 17 (1975), pp. 657–665.

[9] ———, *The completeness and expansion theorems associated with the multiparameter eigenvalue problem in ordinary differential equations*, J. Differential Equations, 5 (1969), pp. 197–213.

[10] ———, *An oscillation theorem for a one-parameter ordinary differential equations of the second order*, Ibid., 11 (1972), pp. 10–37.

[11] J. HORN, *Ueber eine lineare differentialgleichung zweiter ordnung mit einem willkürlichen parameter*, Math. Ann., 52 (1899), pp. 271–292.

[12] E. L. INCE, *Ordinary Differential Equations*, Dover, New York, 1956.

[13] R. G. D. RICHARDSON, *Theorems of oscillation for two linear differential equations of the second order*, Trans. Amer. Math. Soc., 13 (1912), pp. 22–34.

[14] B. D. SLEEMAN, *Multi-parameter eigenvalue problems in ordinary differential equations*, Bul. Inst. Politehn. Iaşi., 17 (1971), pp. 51–60.

SIAM J. MATH. ANAL.
Vol. 8, No. 5, October 1977

MAXIMUM PRINCIPLES AND BOUNDS IN SOME INHOMOGENEOUS ELLIPTIC BOUNDARY VALUE PROBLEMS*

P. W. SCHAEFER† AND R. P. SPERB‡

Abstract. A class of inhomogeneous nonlinear elliptic boundary value problems is considered. The Hopf maximum principles are used to deduce that certain functionals defined for solutions of the problem attain a maximum at a critical point of the solution. These maximum principles are then used to determine bounds for quantities of interest in some physical problems.

1. Introduction. In a recent paper, C. Bandle [1] established isoperimetric bounds for the solution of the Poisson problem. These results extend some earlier work of Pólya and Szegö [10]. In addition, C. Bandle developed estimates for the critical value λ_{cr}, the largest value of λ for which nonlinear boundary value problems for equations of the form $\Delta u + \lambda\rho(x)f(u) = 0$ have a positive solution. We will call the problem homogeneous if ρ is constant and inhomogeneous otherwise. We will consider similar problems as C. Bandle but utilize the Hopf maximum principles [11] to establish maximum principles for certain functionals which are defined on solutions of the particular problem. These results then lead to bounds on the solution and/or the gradient of the solution in various linear and nonlinear problems. Furthermore, for the Poisson problem our bounds are applicable in some cases where C. Bandle's technique is inapplicable.

Several authors [6], [7], [8], [9], [12] have recently used the Hopf maximum principles to establish bounds in certain homogeneous boundary value problems. In § 2 we extend this technique to inhomogeneous equations. The resulting principles will then be used to deduce bounds under mixed or Dirichlet boundary conditions in § 3.

Specifically, we let D be a domain in the plane bounded by a sufficiently smooth curve ∂D and assume u is a solution of

$$\Delta u + \lambda\rho(x)f(u) = 0 \quad \text{in } D, \tag{1.1}$$

where Δ is the Laplace operator, λ is a positive parameter, $\rho(x)$ is a positive C^2 density function in D, and $f(u)$ is a positive C^1 function for $u \geqq 0$. We shall be primarily interested in positive solutions u which satisfy the mixed conditions

$$u = 0 \quad \text{on } \Gamma_1, \quad \frac{\partial u}{\partial n} = 0 \quad \text{on } \Gamma_2, \quad \Gamma_1 \cup \Gamma_2 = \partial D, \quad \Gamma_1 \neq \varnothing, \tag{1.2}$$

where $\partial u/\partial n$ denotes the outward normal derivative. In fact, Γ_2 may be empty. We then determine bounds for the maximum value of the solution in the Poisson problem of an inhomogeneous nuclear reactor operating at critical conditions. In the case of Dirichlet boundary conditions, we also indicate how our results can be extended to the case of more than two dimensions.

* Received by the editors January 28, 1976, and in revised form June 5, 1976.
† Department of Mathematics, University of Tennessee, Knoxville, Tennessee 37916.
‡ Department of Mathematics, Vanderbilt University, Nashville, Tennessee 37235.

2. Maximum principles. We consider the functional

$$\Phi = \frac{|\nabla u|^2}{\rho} g(u) + h(u), \tag{2.1}$$

where u is a solution of (1.1) and g and h are arbitrary functions to be determined. We shall choose g and h so that Φ satisfies an elliptic differential equation.

We use the comma notation for partial differentiation and the summation convention on repeated indices in our computations. Thus, we have

$$\Phi_{,k} = \frac{2gu_{,i}u_{,ik}}{\rho} + \frac{|\nabla u|^2 g' u_{,k}}{\rho} - \frac{|\nabla u|^2 g\rho_{,k}}{\rho^2} + h' u_{,k}, \tag{2.2}$$

$$\begin{aligned}\Phi_{,kk} = {} & \frac{2gu_{,ik}u_{,ik}}{\rho} + \frac{2gu_{,i}u_{,ikk}}{\rho} + \frac{4u_{,i}u_{,ik}g'u_{,k}}{\rho} - \frac{4u_{,i}u_{,ik}g\rho_{,k}}{\rho^2} \\ & + \frac{|\nabla u|^4 g''}{\rho} + \frac{|\nabla u|^2 g' u_{,kk}}{\rho} - \frac{2|\nabla u|^2 g' u_{,k}\rho_{,k}}{\rho^2} - \frac{|\nabla u|^2 g\rho_{,kk}}{\rho^2} \\ & + \frac{2|\nabla u|^2 g|\nabla\rho|^2}{\rho^3} + h''|\nabla u|^2 + h' u_{,kk}.\end{aligned} \tag{2.3}$$

Here the prime denotes differentiation with respect to the argument which has been suppressed.

From (1.1) and (2.2) we determine $u_{,xx}$, $u_{,yy}$, and $u_{,xy}$ so that by the identity

$$u_{,ik}u_{,ik} = (\Delta u)^2 + 2(u_{,xy}^2 - u_{,xx}u_{,yy}) \tag{2.4}$$

we can substitute for the first term in (2.3). We then use (2.2) in the third and fourth terms of (2.3). After a lengthy computation, we obtain

$$\begin{aligned}\Phi_{,kk} + \frac{L_k\Phi_{,k}}{|\nabla u|^2} = {} & \frac{|\nabla u|^4}{\rho}\left[g'' - \frac{g'^2}{g}\right] + |\nabla u|^2\left[(h' - 2\lambda fg)' - \lambda fg' - \frac{g\Delta(\ln\rho)}{\rho}\right] \\ & + \frac{\rho}{g}(h' - \lambda fg)(h' - 2\lambda fg),\end{aligned} \tag{2.5}$$

where

$$L_k = -\frac{\rho}{g}\{\Phi_{,k} + 2u_{,k}[\lambda fg - h']\}.$$

More simply, we write

$$\Delta\Phi + \frac{L_k\Phi_{,k}}{|\nabla u|^2} = c_2|\nabla u|^4 + c_1|\nabla u|^2 + c_0, \tag{2.6}$$

where c_0, c_1, c_2 have the obvious interpretation. Consequently, we have

THEOREM 1. *Let u be a solution of* (1.1). *If ρ, g, and h are C^2 functions with $\rho > 0$ and $g > 0$ such that the quadratic form in* (2.6) *is nonnegative, then Φ attains a maximum either on ∂D or at a critical point of u.*

For simplicity here and later in the applications, we take

$$g(u)\equiv 1, \quad h(u)=2\lambda\int_0^u f(\eta)\,d\eta, \quad \Delta(\ln\rho)\leqq 0 \quad \text{in } D. \tag{2.7}$$

Clearly, one could choose other functions g and h as indicated in [12].

We now consider $\partial\Phi/\partial n$ at an arbitrary point $P\in\partial D$. If $P\in\Gamma_1$, then $|\nabla u|=|\partial u/\partial n|$ so that

$$\frac{\partial\Phi}{\partial n}=\frac{2u_nu_{nn}}{\rho}-\frac{|\nabla u|^2\rho_n}{\rho^2}+2\lambda f u_n, \tag{2.8}$$

where we use the subscript notation for the normal derivative. On the boundary we have

$$\Delta u = u_{nn}+\kappa u_n = -\lambda\rho f. \tag{2.9}$$

Here κ denotes the curvature of the boundary. Consequently,

$$\frac{\partial\Phi}{\partial n}=-|\nabla u|^2\left\{\frac{2\kappa}{\rho}+\frac{1}{\rho^2}\frac{\partial\rho}{\partial n}\right\}. \tag{2.10}$$

On the other hand, if $P\in\Gamma_2$ we again obtain (2.10) since on the boundary

$$u_{sn}=u_{ns}-\kappa u_s, \tag{2.11}$$

where u_s denotes the tangential derivative.

Let

$$\kappa_g=\rho^{-1/2}\left\{\kappa+\frac{1}{2}\frac{\partial}{\partial n}(\ln\rho)\right\} \tag{2.12}$$

represent the geodesic curvature of ∂D considered as a curve on the Riemann surface with line element $ds^2=\rho(d\xi^2+d\eta^2)$. Then if $\kappa_g\geqq 0$ on ∂D, we conclude that $\partial\Phi/\partial n\leqq 0$. Thus by Hopf's maximum principle [2] we deduce

THEOREM 2. *If u is a solution of* (1.1), (1.2) *where the geodesic curvature κ_g of ∂D is nonnegative and if g, h, ρ satisfy* (2.7), *then Φ takes its maximum at a critical point of u.*

Remark 1. Theorems 1 and 2 can be extended to the case of $n>2$ dimensions as follows. We define Φ by (2.1) but use the Schwarz inequality on the first term in (2.3) (see [7]) as no identity of the form (2.4) is available. Proceeding as in the derivation of (2.5), we obtain

$$\begin{aligned}\Phi_{,kk}+\frac{H_k\Phi_{,k}}{|\nabla u|^2}\geqq{}&\frac{|\nabla u|^4}{\rho}\left[g''-\frac{3g'^2}{2g}\right]+\frac{|\nabla u|^2 g|\nabla\rho|^2}{2\rho^3}+\frac{|\nabla u|^2 g' u_{,k}\rho_{,k}}{\rho^2}\\&+\frac{u_{,k}\rho_{,k}}{\rho}[h'-2\lambda fg]+|\nabla u|^2\left[(h'-2\lambda fg)'+\lambda fg'-\frac{g'h'}{g}-\frac{g\Delta\rho}{\rho^2}\right]\\&+\frac{\rho h'}{2g}[h'-2\lambda fg]\end{aligned} \tag{2.13}$$

where

$$H_k = \frac{\rho}{g}\left\{h' u_{,k} - \frac{1}{2}\Phi_{,k}\right\} + |\nabla u|^2\left\{\frac{\rho_{,k}}{\rho} - \frac{g' u_{,k}}{g}\right\}.$$

By means of the arithmetic mean-geometric mean inequality on the third term, we can write

$$\Delta\Phi + \frac{H_k \Phi_{,k}}{|\nabla u|^2} \geqq \frac{|\nabla u|^4}{\rho}\left[g'' - \frac{2g'^2}{g}\right] + \frac{u_{,k}\rho_{,k}}{\rho}[h' - 2\lambda f g]$$
$$+ |\nabla u|^2\left[(h' - 2\lambda f g)' + \lambda f g' - \frac{g' h'}{g} - \frac{g\Delta\rho}{\rho^2}\right] + \frac{\rho h'}{2g}[h' - 2\lambda f g].$$

Thus we see that if ρ, g, and h satisfy

$$(g^{-1})'' \leqq 0, \quad h \equiv 2\lambda \int_0^u f(\eta) g(\eta)\, d\eta, \quad \Delta\rho \leqq 0 \quad \text{in } D, \tag{2.14}$$

we obtain the conclusion of Theorem 1. Again for simplicity one could ask that $g \equiv 1$, in which case $\Delta\rho^{1/2} \leqq 0$ in D is sufficient. Furthermore, under Dirichlet boundary conditions for u, if one asks that

$$\rho^{-1/2}\left\{(n-1)\kappa + \frac{1}{2}\frac{\partial}{\partial n}(\ln \rho)\right\} \geqq 0 \tag{2.15}$$

on ∂D, then the conclusion of Theorem 2 holds.

Remark 2. In the event that $\kappa_g \geqq 0$ on ∂D is not satisfied or that ∂D is nonconvex, we add αu to the functional Φ in (2.1), where α is a positive unspecified constant. The positivity of α ensures that Φ will satisfy the ellipticity requirement. We then choose α so that

$$2\kappa + \frac{\partial}{\partial n}(\ln \rho) + \frac{\alpha\rho}{|\nabla u|} \geqq 0 \quad \text{on } \partial D. \tag{2.16}$$

Thus in the case of Dirichlet boundary conditions, we are again led to the conclusion of Theorem 2.

We now mention another result for solutions of the Poisson equation

$$\Delta u = -\rho \quad \text{in } D. \tag{2.17}$$

The following is an extension of a result of Miranda [4] which was used by Payne in [6] in the case $\rho = 2$.

THEOREM 3. *If u is any solution of* (2.17), *where $\rho > 0$ and $\Delta\rho \leqq 0$ in D, then the function*

$$\Psi = \frac{|\nabla u|^2}{\rho} + u$$

takes its maximum on ∂D.

Proof. We have

$$\Psi_{,k} = \frac{2u_{,i}u_{,ik}}{\rho} - \frac{|\nabla u|^2 \rho_{,k}}{\rho^2} + u_{,k}, \tag{2.18}$$

$$\Delta\Psi=\frac{2u_{,ik}u_{,ik}}{\rho}+\frac{2u_{,i}u_{,ikk}}{\rho}-\frac{4u_{,i}u_{,ik}\rho_{,k}}{\rho^2}+\frac{2|\nabla u|^2|\nabla\rho|^2}{\rho^3}-\frac{|\nabla u|^2\Delta\rho}{\rho^2}+\Delta u. \tag{2.19}$$

In view of (2.17) and (2.18), we find

$$\Delta\Psi+\frac{2}{\rho}\rho_{,k}\Psi_{,k}=\frac{1}{\rho}[2u_{,ik}u_{,ik}-(\Delta u)^2]-\frac{|\nabla u|^2\Delta\rho}{\rho^2}. \tag{2.20}$$

From (2.4) it follows that

$$2u_{,ik}u_{,ik}-(\Delta u)^2=(u_{,xx}-u_{,yy})^2+4u_{,xy}^2,$$

which is clearly nonnegative. Thus our assumptions on ρ lead to the conclusion of the theorem. We note that Ψ could be a constant as in the case when ρ is constant, D is a disk, and u vanishes on the boundary.

Finally, it is interesting to note that if u is the solution of (2.17) which satisfies the boundary conditions (1.2), then from Theorem 2 it follows that

$$\Phi=\frac{|\nabla u|^2}{\rho}+2u$$

takes its maximum at a critical point of u, provided $\kappa_g\geqq 0$ on ∂D and $\Delta(\ln\rho)\leqq 0$ in D.

3. Applications. Our first application is to problem (2.17), (1.2) where $\rho>0$ and $\Delta p\leqq 0$. Since Ψ takes its maximum at some point P on ∂D, we have

$$\frac{\partial\Psi}{\partial n}(P)>0. \tag{3.1}$$

If $P\in\Gamma_1$, then by a calculation similar to that in [6], we find

$$\sigma\equiv\max_{\partial D}\frac{|\nabla u|}{\rho^{1/2}}\leqq\frac{1}{2\kappa_0},\qquad \kappa_g\geqq\kappa_0>0\quad\text{on }\Gamma_1. \tag{3.2}$$

We note that if $\kappa_g\geqq 0$ on Γ_2, then $\Psi_{\max}$ cannot occur there since on Γ_2

$$\frac{\partial\Psi}{\partial n}=-2\rho^{-1/2}\kappa_g|\nabla u|^2\leqq 0,$$

which would contradict the strong maximum principle. Further, from Theorem 3 and (3.2), we conclude

$$\frac{|\nabla u|^2}{\rho}+u\leqq\frac{1}{4\kappa_0^2}\quad\text{in } D, \tag{3.3}$$

and hence

$$u_{\max}\leqq\frac{1}{4\kappa_0^2}. \tag{3.4}$$

As observed by Payne [6], the equality sign holds in (3.2), (3.3), and (3.4) when ρ is constant and D is a circle.

Although one could use (3.3) to deduce other inequalities, we will make use of Theorem 2 instead. As an inequality it states that

$$|\nabla u|^2 + 2\rho u \leqq 2\rho u_{\max}. \tag{3.5}$$

Let

$$M = \int_D \rho \, dx \quad \text{and} \quad L = \oint_{\partial D} \rho^{1/2} \, ds.$$

By integration over ∂D and some manipulation, we obtain ($\Gamma_2 = \varnothing$)

$$u_{\max} \geqq \frac{M^2}{2L^2}. \tag{3.6}$$

The equality sign holds in (3.5) if D is an infinite strip and ρ is constant.

Clearly, other bounds for $u_{\max}$ follow from (3.5). We note that by a calculation analogous to [6], we can extend Payne's result and obtain

$$u_{\max} \leqq \frac{1}{2} d^2 \tag{3.7}$$

where d is given as follows: if P denotes the (unknown) point where u attains its maximum and Q is an arbitrary boundary point, then

$$d = \max_{P \in D} \left(\min_{Q \in \partial D} \int_P^Q \rho^{1/2}(r) \, dr \right) \tag{3.8}$$

where r measures the distance along the ray joining P and Q. Actually, one could take any curve from P to Q. When ρ is a constant, (3.7) gives the result of Payne [6]. Finally, if ρ and D have two axes of symmetry, one could show that P could be taken at the center and Q as the nearest boundary point in the ordinary sense.

As a second application, we consider the inhomogeneous fixed membrane problem

$$\begin{aligned} \Delta\phi + \lambda\rho\phi &= 0 \quad \text{in } D, \\ \phi &= 0 \quad \text{on } \partial D. \end{aligned} \tag{3.9}$$

In a nuclear reactor context (see [7]), the "efficiency ratio" E defined by

$$E = \frac{\int_D \rho\phi \, dx}{\phi_{\max} \cdot \int_D \rho \, dx} = \frac{\int_D \rho\phi \, dx}{\phi_{\max} M} \tag{3.10}$$

plays an important role. Here ϕ is the first eigenfunction of (3.9) with associated eigenvalue λ. From Theorem 2, we have

$$|\nabla\phi|^2 + \lambda\rho\phi^2 \leqq \lambda\rho\phi_{\max}^2. \tag{3.11}$$

An easy calculation then results in

$$E \leqq \frac{L}{M\lambda^{1/2}}, \qquad \left(L = \oint_{\partial D} \rho^{1/2} \, ds \right). \tag{3.12}$$

Thus (3.12) is an extension of the corresponding result of Payne and Stakgold [7].

We note that if one were to consider an efficiency ratio $\tilde{E}$ defined by

$$\tilde{E} = \frac{\int_D \rho\phi^2\, dx}{\phi_{\max}^2 \cdot M}, \tag{3.13}$$

then the integration of (3.11) and use of (3.9) results in

$$\tilde{E} \leqq \frac{1}{2}, \tag{3.14}$$

with equality if D shrinks to an interval.

Finally, we consider the following nonlinear eigenvalue problem ($p > 0$)

$$\begin{aligned} \Delta u + \lambda u^p &= 0 \quad \text{in } D, \\ u &= 0 \quad \text{on } \partial D. \end{aligned} \tag{3.15}$$

We have chosen $\rho \equiv 1$ for simplicity; the extension to nonconstant ρ is immediate. It follows from the results of Levinson [3] that this problem has a positive solution for any $\lambda > 0$.

By Theorem 2, we get

$$|\nabla u|^2 + \frac{2\lambda}{p+1} u^{p+1} \leqq \frac{2\lambda}{p+1} u_{\max}^{p+1}. \tag{3.16}$$

Let P be a point in D where u takes its maximum and Q be the point on ∂D nearest to P. Let r denote the distance along the ray from P to Q. Then from (3.16) we have

$$-\frac{du}{dr} \leqq |\nabla u| \leqq \left[\frac{2\lambda}{p+1}(u_{\max}^{p+1} - u^{p+1})\right]^{1/2}. \tag{3.17}$$

Now integration of (3.17) from P to Q and a short calculation results in

$$(u_{\max})^{(1-p)/2} N(p) \leqq d\sqrt{\frac{2\lambda}{p+1}}. \tag{3.18}$$

Here d is the radius of the largest inscribed circle in D with center at P and

$$N(p) = \sqrt{\pi}\,\Gamma\left(\frac{1}{p+1}\right)\left[(p+1)\Gamma\left(\frac{p+3}{2(p+1)}\right)\right]^{-1}.$$

We note that (3.18) gives an upper bound for $u_{\max}$ if $0 < p < 1$ and a lower bound if $p > 1$. When $p = 1$, we get

$$\lambda \geqq \frac{\pi^2}{4d^2}, \tag{3.19}$$

which is a lower bound for the first eigenvalue of the fixed membrane spanned over D.

4. Concluding remarks. In the previous section we indicated how our results may be used to obtain bounds for various important quantities in some physical problems of interest. Other techniques are presented in [9].

We note that extensions to problems with boundary conditions of the third kind in $n \geqq 2$ dimensions for the inhomogeneous problem and $n > 2$ dimensions for the homogeneous problem remain to be done. Extensions to uniformly elliptic operators will appear in a forthcoming paper.

REFERENCES

[1] C. BANDLE, *Bounds for the solutions of Poisson problems and applications to nonlinear eigenvalue problems*, this Journal, 6 (1975), pp. 146–152.

[2] E. HOPF, *A remark on linear elliptic differential equations of the second order*, Proc. Amer. Math. Soc., 3 (1952), pp. 791–793.

[3] N. LEVINSON, *Positive eigenfunctions for* $\Delta u + \lambda f(u) = 0$, Arch. Rational Mech. Anal., 11 (1962), pp. 258–272.

[4] C. MIRANDA, *Formule di maggiorazione e teorema di esistenza per le funzioni biharmoniche di du variabili*, Gioro. Mat. Battaglini, 78 (1948), pp. 97–118.

[5] L. E. PAYNE, *Isoperimetric inequalities and their applications*, SIAM Rev., 9 (1967), pp. 453–488.

[6] ———, *Bounds for the maximum stress in the Saint Venant torsion problem*, Indian J. Mech. Math., Special Issue (1968), pp. 51–59.

[7] 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.

[8] ———, *Nonlinear problems in nuclear reactor analysis*, Proc. Conference on Nonlinear Problems in Physical Sciences and Biology, Springer Lecture Notes in Math., No. 322, Springer-Verlag, New York, 1972, pp. 298–307.

[9] L. E. PAYNE, R. P. SPERB AND I. STAKGOLD, *On Hopf type maximum principles for convex domains*, to appear.

[10] G. PÓLYA AND G. SZEGÖ, *Isoperimetric inequalities in mathematical physics*, Princeton University Press, Princeton, NJ, 1951.

[11] M. H. PROTTER AND H. F. WEINBERGER, *Maximum principles in differential equations*, Prentice-Hall, Englewood Cliffs, NJ, 1967.

[12] 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.

SIAM J. MATH. ANAL.
Vol. 8, No. 5, October 1977

APPLICATIONS OF RIEMANN–STIELTJES INTEGRALS OF ORDER k IN FUNCTIONAL ANALYSIS*

A. M. RUSSELL†

Abstract. In this paper we show that any bounded linear functional on $C[a, b]$ has a representation as a kth order Riemann–Stieltjes integral. Conversely, a kth order Riemann–Stieltjes integral formed with a fixed function of bounded kth variation defines a bounded linear functional on $C[a, b]$. Finally, an interpretation of our integrals is made in the context of distribution theory.

Introduction. The important initial contribution in this field was made by F. Riesz [2, § 50] when he showed that a bounded linear functional defined on $C[a, b]$ could be represented as a Riemann–Stieltjes integral. Conversely, the Riemann–Stieltjes integral formed with a fixed function g of bounded variation on $[a, b]$ defines a bounded linear functional on $C[a, b]$, and its norm is equal to the total variation of g on $[a, b]$. We establish corresponding results for the kth order Riemann–Stieltjes integral, $\int_a^b f(x)[d^k g(x)/dx^{k-1}]$, the definition of which is given in [5].

The theory of distributions has many important applications in mathematics and physics. Several different approaches to the theory are given in the literature. Distributions as defined by L. Schwartz [7] are linear functionals defined on test functions having derivatives of all orders. Our kth order integrals will be interpreted as linear functionals possessing, in a completely rigorous way, the important properties of Dirac's delta function, but a knowledge of topological concepts will not be required.

The simplest distribution is a Riemann–Stieltjes integral $\int_a^b f(x)\, dg(x)$, and physically this can be used to represent a mass distribution on a line. Higher order integrals of the form $\int_a^b f(x)[d^k g(x)/dx^{k-1}]$ are also necessary, the cases $k \geqq 2$ representing dipole and multipole distributions.

The analysis given is constructive. We assume that k is an integer greater than or equal to three; the case $k = 2$ appears in unpublished form in [3], and can be obtained from the general case by easy and obvious modifications.

Notation and preliminaries. Unless otherwise stated undefined terms and notation can be found in [5] and [6].

DEFINITION 1. Let k be any positive integer greater than 1. Then we will denote by $\Gamma(x_{-k+1}, \cdots, x_{n+k-1})$ a subdivision of the closed interval $[a, b]$ such that

$$a' \leqq x_{-k+1} < \cdots < x_0 = a < x_1 < \cdots < x_n = b < \cdots < x_{n+k-1} \leqq b'.$$

In this paper it will always be understood that $a' < a < b < b'$. In addition we point out at this stage that a Γ subdivision , as opposed to a π subdivision of $[a, b]$ (see [5]), requires a fixed number $2k-2$ of points outside $[a, b]$.

* Received by the editors September 12, 1975 and in revised form September 27, 1976.
† Department of Mathematics, University of Melbourne, Parkville, Victoria 3052, Australia.

Throughout the paper we will make extensive use of the following function:

DEFINITION 2. Define

$$p_x(t) = 0, \quad t \leqq x,$$
$$= 1, \quad t > x.$$

DEFINITION 3. If L is a bounded linear functional on $C[a, b]$, then we define a function g by

$$g(x) = -(k-1)!L[w_x(t)], \qquad a' \leqq x \leqq b', \tag{1}$$

where

$$w_x(t) = \int_{a'}^{t} \frac{(t-u)^{k-2}}{(k-2)!} p_x(u)\, du.$$

We now discuss the function

$$Z(k, t, x_i, \cdots, x_{i+k-1}) \equiv Z(t)$$
$$= (k-1)!Q_{k-1}(w_x(t); x_i, \cdots, x_{i+k-1}) \tag{2}$$
$$= (k-1)! \sum_{s=i}^{i+k-1} \frac{w_{x_s}(t)}{\prod_{j=i}^{i+k-1'} (x_s - x_j)},$$

where $\prod'$ indicates that zero factors are omitted. (See Definition 1(a) of [5].)

When $t \leqq x_i$,

$$Z(t) = 0.$$

When $x_i \leqq t \leqq x_{i+1}$,

$$Z(t) = \frac{(k-1)!w_{x_i}(t)}{(x_i - x_{i+1}) \cdots (x_i - x_{i+k-1})} = \frac{(t-x_i)^{k-1}}{(x_i - x_{i+1}) \cdots (x_i - x_{i+k-1})}.$$

Generally, when $x_{i+j} \leqq t \leqq x_{i+j+1}$, $j = 0, 1, \cdots, k-2$,

$$Z(t) = \sum_{m=0}^{j} \frac{(t - x_{i+m})^{k-1}}{\prod_{s=i}^{i+k-1'} (x_{i+m} - x_s)}.$$

Finally, when $x_{i+k-1} \leqq t$,

$$Z(t) = (k-1)!Q_{k-1}\left[\frac{(t-x)^{k-1}}{(k-1)!}; x_i, \cdots, x_{i+k-1}\right]$$
$$= (k-1)!\frac{(-1)^{k-1}}{(k-1)!} \quad \text{(by Lemma 2 of [5])}$$
$$= (-1)^{k-1}.$$

We now investigate $w_x(t)$ as a function of x, so put $t = t_0$, a constant. Then

$$w_x(t_0) = \int_{a'}^{t_0} \frac{(t_0 - u)^{k-2}}{(k-2)!} p_x(u)\, du.$$

If $x \leqq t_0$, then it is easy to show that $w_x(t_0) = (t_0 - x)^{k-1}/[(k-1)!]$, whereas if $x > t_0$, $w_x(t_0) = 0$.

We now return to the function $Z(t)$. Since $k \geqq 3$, $Z(t)$ is differentiable in $[a', b']$, and

$$Z'(t) = (k-1)!\,Q_{k-1}\left[\frac{d}{dt}w_x(t); x_i, \cdots, x_{i+k-1}\right]. \tag{3}$$

Now

$$\frac{d}{dt}w_x(t) = 0, \qquad t \leqq x,$$

$$= \frac{(t-x)^{k-2}}{(k-2)!}, \qquad t \geqq x,$$

and if we denote $[(d/dt)w_x(t)]_{t=t_0}$ by $w'_x(t_0)$, then it follows readily that

$$\frac{d^{k-3}}{dx^{k-3}}[w'_x(t_0)] = (-1)^{k-1}(t_0 - x), \quad x \leqq t_0,$$

$$= 0, \qquad x \geqq t_0.$$

This function is 2-convex if k is odd, and 2-concave if k is even. It therefore follows from repeated applications of Theorem 13 of [5] that $w'_x(t_0)$ is $(k-1)$-convex when k is odd, and $(k-1)$-concave when k is even. Hence it follows from (3) that for all t in $[a', b']$, $Z'(t)$ is nonnegative when k is odd, and nonpositive when k is even. Thus $Z(t)$ is a nondecreasing function when k is odd, and a nonincreasing function when k is even, and so the graphs in Figs. 1 and 2 are obtained.

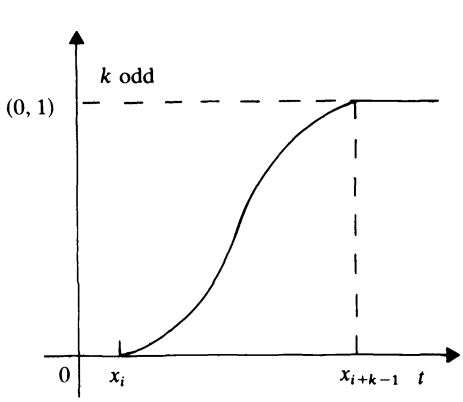


FIG. 1

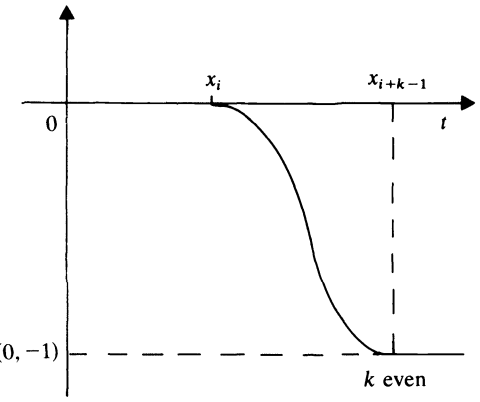


FIG. 2

Considering a $\Gamma(x_{-k+1}, \cdots, x_{n+k-1})$ subdivision of $[a, b]$, we conclude that

$$(k-1)! \sum_{i=-k+1}^{n-1} [Q_{k-1}(w_x(t); x_{i+1}; \cdots, x_{i+k}) - Q_{k-1}(w_x(t); x_i, \cdots, x_{i+k-1})]$$

$$= (k-1)![Q_{k-1}(w_x(t); x_n, \cdots, x_{n+k-1}) - Q_{k-1}(w_x(t); x_{-k+1}, \cdots, x_0)]$$

has the graph shown in Figs. 3 and 4.

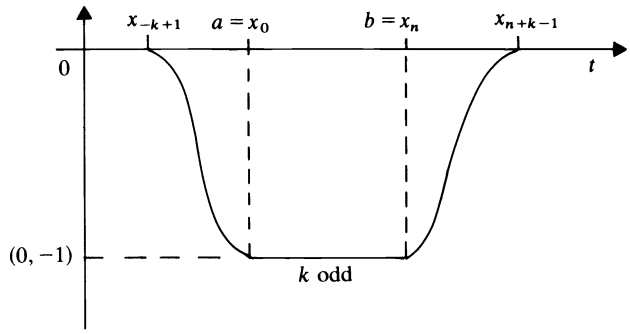


FIG. 3

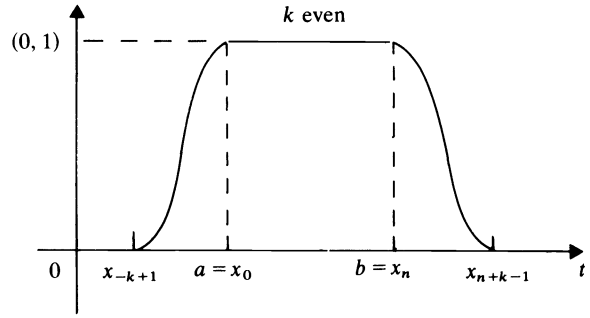


FIG. 4

By noting the constant sign property of

$$Q_{k-1}(w_x(t); x_{i+1}, \cdots, x_{i+k}) - Q_{k-1}(w_x(t); x_i, \cdots, x_{i+k-1}),$$

it follows readily that

$$\sup_{a' \leqq t \leqq b'} (k-1)! \sum_{i=-k+1}^{n-1} \left| [Q_{k-1}(w_x(t); x_{i+1}, \cdots, x_{i+k}) - Q_{k-1}(w_x(t); x_i, \cdots, x_{i+k-1})] \right| = 1. \tag{4}$$

THEOREM 1. *The function g defined by* (1) *belongs to $BV_k[a', b']$.*

Proof. We consider any $\Gamma(x_{-k+1}, \cdots, x_{n+k-1})$ subdivision of $[a, b]$. Then, using (1) we can write an approximating sum for $V_k[g; a', b']$ in the form

$$\sum_{i=-k+1}^{n-1} |Q_{k-1}(g; x_{i+1}, \cdots, x_{i+k}) - Q_{k-1}(g; x_i, \cdots, x_{i+k-1})|$$

$$= \sum_{i=-k+1}^{n-1} (k-1)! u_i L[Q_{k-1}(w_x(t); x_{i+1}, \cdots, x_{i+k}) - Q_{k-1}(w_x(t); x_i, \cdots, x_{i+k-1})],$$

where $u_i = \pm 1$, the value of u_i being chosen so that

$$u_i L[Q_{k-1}(w_x(t); x_{i+1}, \cdots, x_{i+k}) - Q_{k-1}(w_x(t); x_i, \cdots, x_{i+k-1})]$$

is nonnegative. Since L is a bounded linear functional on $C[a, b]$ there exists a constant M such that

$$|L(f)| \leqq M\|f\|$$

for each $f \in C[a, b]$, where

$$\|f\| = \sup_{a \leqq x \leqq b} |f(x)|.$$

Consequently,

$$\begin{aligned} &\sum_{i=-k+1}^{n-1} |Q_{k-1}(g; x_{i+1}, \cdots, x_{i+k}) - Q_{k-1}(g; x_i, \cdots, x_{i+k-1})| \\ &\quad \leqq M \sup_{a \leqq t \leqq b} (k-1)! \sum_{i=-k+1}^{n-1} |u_i| \, |Q_{k-1}(w_x(t); x_{i+1}, \cdots, x_{i+k}) \\ &\qquad\qquad - Q_{k-1}(w_x(t); x_i, \cdots, x_{i+k-1})| \\ &\quad = M \quad \text{(by using (4))}. \end{aligned} \tag{5}$$

It now follows that $g \in BV_k[a', b']$.

We conclude this section by stating two theorems that will be required at a later stage.

THEOREM 2. *Let f and all of its derivatives $f^{(r)}$, $r = 1, 2, \cdots, k-1$, be continuous on $[a, b]$. If $g \in BV_k[a, b]$ and vanishes when $x \leqq a$ and $x \geqq b$, then*

$$(k-1)! \int_a^b f(x) \frac{d^k g(x)}{dx^{k-1}}$$

exists, and

$$(k-1)! \int_a^b f(x) \frac{d^k g(x)}{dx^{k-1}} = (-1)^{k-1} \int_a^b f^{(k-1)}(x)\, dg(x).$$

Proof. We first observe that the conditions given guarantee the existence of the RS_k integral. We now consider any $\Gamma_h(x_{-k+1}, \cdots, x_{n+k-1})$ subdivision of $[a, b]$, so that each subinterval $x_i - x_{i-1}$ is of length h. If we write Δ^s instead of Δ_h^s, which is defined in [6, § 5], then the approximating sum $S(\Gamma_h, f, g)$ for $\int_a^b f(x)[d^k g(x)/dx^{k-1}]$ can be rearranged in the form

$$s(\Gamma_h, f, g) = \frac{(-1)^{k-1}}{(k-1)!} \sum_{i=0}^{n-1} \frac{\Delta^{k-1} f(x_{i-k+2})}{h^{k-1}} \Delta g(x_i) + R,$$

where

$$R = \frac{1}{(k-1)!} \sum_{m=1}^{k-1} (-1)^{m-1} \left[\frac{\Delta^{m-1} f(x_{n+1-m})}{h^{m-1}} \frac{\Delta^{k-m} g(x_n)}{h^{k-m}} - \frac{\Delta^{m-1} f(x_{-k+2})}{h^{m-1}} \frac{\Delta^{k-m} g(x_{-k+m})}{h^{k-m}} \right].$$

Thus, using the result

$$\Delta^s f(x) = h^k f^{(s)}(\xi), \quad \text{where } x < \xi < x + sh,$$

we obtain

$$S(\Gamma_h, f, g) = \frac{(-1)^{k-1}}{(k-1)!} \sum_{i=0}^{n-1} \frac{\Delta^{k-1} f(x_{i-k+2})}{h^{k-1}} \Delta g(x_i) + R$$

$$= \frac{(-1)^{k-1}}{(k-1)!} \sum_{i=0}^{n-1} f^{(k-1)}(x_i) \Delta g(x_i) + R$$

$$+ \frac{(-1)^{k-1}}{(k-1)!} \sum_{i=0}^{n-1} [f^{(k-1)}(\eta_i) - f^{(k-1)}(x_i)] \Delta g(x_i),$$

where $x_{i-k+2} < \eta_i < x_{i+1}$, $i = 0, 1, \cdots, n-1$.

It can now be readily seen that the limit of the right hand side of the last equation as h approaches zero is

$$\frac{(-1)^{k-1}}{(k-1)!} \int_a^b f^{(k-1)}(x)\, dg(x),$$

as required.

THEOREM 3. *If f is of bounded variation on $[a', b']$, and g has a continuous $(k-1)$th derivative on $[a', b']$, then*

$$(k-1)!(M_1) \int_a^b f(x) \frac{d^k g(x)}{dx^{k-1}}$$

exists, and equals

$$\int_a^b f(x)\, dg^{(k-1)}(x).$$

Proof. We consider a $\Gamma(x_{-k+1}, \cdots, x_{n+k-1})$ subdivision of $[a, b]$, and write

$$(k-1)! Q_{k-1}(g; x_i, \cdots, x_{i+k-1}) = g^{(k-1)}(x_i) + \{g^{(k-1)}(\eta_i) - g^{(k-1)}(x_i)\},$$

where $x_i < \eta_i < x_{i+k}$. Therefore,

$$\left| Q_{k-1}(g; x_i, \cdots, x_{i+k-1}) - \frac{g^{(k-1)}(x_i)}{(k-1)!} \right| = \left| \frac{g^{(k-1)}(\eta_i) - g^{(k-1)}(x_i)}{(k-1)!} \right|.$$

Hence, an approximating sum for the $M_1 RS_k$ integral can be written as

$$(k-1)! \sum_{i=-k+1}^{n-1} f(x_{i+1})[Q_{k-1}(g; x_{i+1}, \cdots, x_{i+k}) - Q_{k-1}(g; x_i, \cdots, x_{i+k-1})]$$

$$= \sum_{i=-k+1}^{n-1} f(x_{i+1})[g^{(k-1)}(x_{i+1}) - g^{(k-1)}(x_i)]$$

$$+ \sum_{i=-k+1}^{n-1} f(x_{i+1})[\{g^{(k-1)}(\eta_{i+1}) - g^{(k-1)}(x_{i+1})\} - \{g^{(k-1)}(\eta_i) - g^{(k-1)}(x_i)\}]$$

$$= \sum_{i=0}^{n-1} f(x_{i+1})[g^{(k-1)}(x_{i+1}) - g^{(k-1)}(x_i)] + \sum_{i=-k+1}^{-1} f(x_{i+1})[g^{(k-1)}(x_{i+1}) - g^{(k-1)}(x_i)]$$

$$+ \sum_{i=-k+1}^{n-1} [f(x_{i+1}) - f(x_{i+2})][g^{(k-1)}(\eta_{i+1}) - g^{(k-1)}(x_{i+1})]$$

$$+ f(x_{n+1})[g^{(k-1)}(\eta_n) - g^{(k-1)}(x_n)] - f(x_{-k+2})[g^{(k-1)}(\eta_{-k+1}) - g^{(k-1)}(x_{-k+1})].$$

Letting $\|\Gamma\|$ approach zero in the last equation, and using the uniform continuity of $g^{(k-1)}$ and the bounded variation of f, gives us the required result.

Representation theorem. We are now in a position to present a representation theorem for bounded linear functionals on $C[a, b]$. We recall the definition of the MRS_k integral—see Definition 5 of [6].

THEOREM 4. *Let L be a bounded linear functional on $C[a, b]$. Then, for any natural number k, there exists a function $g \in BV_k[a', b']$ such that*

$$L(f) = (M)\int_a^b f(x)\frac{d^k g(x)}{dx^{k-1}}$$

for all $f \in C[a, b]$.

Furthermore,

$$\|L\| = |D_+^{k-1}g(a) - D_-^{k-1}g(a)| + V_k(g; a, b) + |D_+^{k-1}g(b) - D_-^{k-1}g(b)|.$$

Conversely,

$$(M)\int_a^b f(x)\frac{d^k g(x)}{dx^{k-1}}$$

formed with a fixed function $g \in BV_k[a', b']$ defines a bounded linear functional on $C[a, b]$.

Proof. Assume that k is odd. A similar analysis will be applicable if k is even. We consider any $\Gamma(x_{-k+1}, \cdots, x_{n+k-1})$ subdivision of $[a, b]$, and define

$$\xi_{-k+1} = \cdots = \xi_{-1} = a, \qquad x_i \leqq \xi_i \leqq x_{i+k}, \qquad i = 0, 1, \cdots, n-k,$$

and

$$\xi_{n-k+1} = \cdots = \xi_{n-1} = b.$$

We now define an approximating function ϕ for f.

$$\phi(t) = -(k-1)! \sum_{i=-k+1}^{n-1} f(\xi_i)\{Q_{k-1}(w_x(t); x_{i+1}, \cdots, x_{i+k}) - Q_{k-1}(w_x(t); x_i, \cdots, x_{i+k-1})\}, \qquad a' \leqq t \leqq b'.$$

We now compare f and ϕ on $[a, b]$; so assume that

$$a \leqq x_m \leqq t \leqq x_{m+1} \leqq b.$$

Then, according to Fig. 1,

$$\begin{aligned}
-\frac{\phi(t)}{(k-1)!} = &-f(\xi_m)Q_{k-1}(w_x(t); x_m, \cdots, x_{m+k-1}) \\
&+f(\xi_{m-1})\{Q_{k-1}(w_x(t); x_m, \cdots, x_{m+k-1}) \\
&\qquad - Q_{k-1}(w_x(t); x_{m-1}, \cdots, x_{m+k-2})\} \\
&+ \\
&\vdots \\
&+f(\xi_{m-k+1})\{Q_{k-1}(w_x(t); x_{m-k+2}, \cdots, x_{m+1}) \\
&\qquad - Q_{k-1}(w_x(t); x_{m-k+1}, \cdots, x_m)\}.
\end{aligned}$$

Therefore,

$$\frac{f(t)-\phi(t)}{(k-1)!} = Q_{k-1}(w_x(t); x_m, \cdots, x_{m+k-1})\{f(\xi_{m-1})-f(\xi_m)\}+\cdots$$
$$+Q_{k-1}(w_x(t); x_{m-k+2}, \cdots, x_{m+1})\{f(\xi_{m-k+1})-f(\xi_{m-k+2})\}$$
$$+Q_{k-1}(w_x(t); x_{m-k+1}, \cdots, x_m)\{f(t)-f(\xi_{m-k+1})\},$$

since, when k is odd, $Q_{k-1}(w_x(t); x_{m-k+1}, \cdots, x_m) = 1/(k-1)!$, $x_m \leqq t \leqq x_{m+1}$. Since L is bounded there exists M such that $|L(f)| \leqq M \sup_x |f(x)|$ for all f in $C[a, b]$. Because f is uniformly continuous on $[a, b]$, for each $\varepsilon > 0$, there exists $\delta(\varepsilon)$ such that $\sup_{a \leqq t \leqq b} |\phi(t)-f(t)| < \varepsilon/M$ whenever $\|\Gamma\| < \delta(\varepsilon)$. Hence, using g as in Definition 3, we obtain

$$\left|L(f)-\sum_{i=-k+1}^{n-1} f(\xi_i)\{Q_{k-1}(g; x_{i+1}, \cdots, x_{i+k})-Q_{k-1}(g; x_i, \cdots, x_{i+k-1})\}\right|$$
$$= |L(f)-L(\phi)| \leqq M \sup_{a \leqq t \leqq b} |f(t)-\phi(t)|$$
$$< \varepsilon \quad \text{whenever } \|\Gamma\| < \delta(\varepsilon).$$

Hence,

$$L(f) = (M)\int_a^b f(x)\frac{d^k g(x)}{dx^{k-1}}.$$

We now obtain the stated value for $\|L\|$. Using the approximating sum of $(M)\int_a^b f(x)[d^k g(x)/dx^{k-1}]$, and writing $\|f\| = \sup_{a \leqq x \leqq b} |f(x)|$, we obtain

$$\left|f(a)\{Q_{k-1}(g; x_0, \cdots, x_{k-1})-Q_{k-1}(g; x_{-k+1}, \cdots, x_0)\}\right.$$
$$+\sum_{i=0}^{n-k} f(\xi_i)\{Q_{k-1}(g; x_{i+1}, \cdots, x_{i+k})-Q_{k-1}(g; x_i, \cdots, x_{i+k-1})\}$$
$$\left.+f(b)\{Q_{k-1}(g; x_n, \cdots, x_{n+k-1})-Q_{k-1}(g; x_{n-k+1}, \cdots, x_n)\}\right|$$
$$\leqq \|f\|\Big\{|Q_{k-1}(g; x_0, \cdots, x_{k-1})-Q_{k-1}(g; x_{-k+1}, \cdots, x_0)| \tag{6}$$
$$+\sum_{i=0}^{n-k}|Q_{k-1}(g; x_{i+1}, \cdots, x_{i+k})-Q_{k-1}(g; x_i, \cdots, x_{i+k-1})|$$
$$+|Q_{k-1}(g; x_n, \cdots, x_{n+k-1})-Q_{k-1}(g; x_{n-k+1}, \cdots, x_n)|\Big\}$$
$$\leqq \|f\|\{|Q_{k-1}(g; x_0, \cdots, x_{k-1})-Q_{k-1}(g; x_{-k+1}, \cdots, x_0)|$$
$$+V_k(g; a, b)+|Q_{k-1}(g; x_n, \cdots, x_{n+k-1})-Q_{k-1}(g; x_{n-k+1}, \cdots, x_n)|\}.$$

Since $g \in BV_k[a', b']$, g has right and left $(k-1)$th Riemann* derivatives at a and b, and letting $\|\Gamma\|$ approach zero in (6) gives

$$\left|(M)\int_a^b f(x)\frac{d^k g(x)}{dx^{k-1}}\right| \leqq \|f\|\{|D_+^{k-1}g(a)-D_-^{k-1}g(a)|+V_k(g; a, b)$$
$$+|D_+^{k-1}g(b)-D_-^{k-1}g(b)|\}.$$

Consequently,

$$\|L\| \leqq |D_+^{k-1}g(a) - D_-^{k-1}g(a)| + V_k(g; a, b) + |D_+^{k-1}g(b) - D_-^{k-1}g(b)|. \tag{7}$$

Using (5), we have

$$\begin{aligned}
&|Q_{k-1}(g; x_{-k+1}, \cdots, x_0) - Q_{k-1}(g; x_0, \cdots, x_{k-1})| \\
&\quad + \sum_{i=0}^{n-k} |Q_{k-1}(g; x_{i+1}, \cdots, x_{i+k}) - Q_{k-1}(g; x_i, \cdots, x_{i+k-1})| \\
&\quad + |Q_{k-1}(g; x_n, \cdots, x_{n+k-1}) - Q_{k-1}(g; x_{n-k+1}, \cdots, x_n)| \\
&\leqq \sum_{i=-k+1}^{n-1} |Q_{k-1}(g; x_{i+1}, \cdots, x_{i+k}) - Q_{k-1}(g; x_i, \cdots, x_{i+k-1})| \\
&\leqq \|L\|.
\end{aligned}$$

Thus,

$$|D_+^{k-1}g(a) - D_-^{k-1}g(a)| + V_k(g; a, b) + |D_+^{k-1}g(b) - D_-^{k-1}g(b)| \leqq \|L\|,$$

and this result combined with (7) gives the stated value for $\|L\|$.

For the converse, we comment that $(M)\int_a^b f(x)[d^k g(x)/dx^{k-1}]$ is linear in f, and its boundedness follows from the inequality

$$\begin{aligned}
\left|(M)\int_a^b f(x)\frac{d^k g(x)}{dx^{k-1}}\right| &\leqq \|f\|\{|D_+^{k-1}g(a) - D_-^{k-1}g(a)| + V_k(g; a, b) \\
&\quad + |D_+^{k-1}g(b) - D_-^{k-1}g(b)|\}.
\end{aligned}$$

Remarks. 1. The proof of Theorem 4 also serves to show that a bounded linear functional on $C[a, b]$ can be represented as an RS_k integral; and conversely that an RS_k integral formed with a fixed function $g \in BV_k[a', b']$ is a bounded linear functional on $C[a, b]$. It is, however, more convenient to use the MRS_k integral when determining $\|L\|$.

2. In [8], Webb shows that a bounded linear functional L defined on the wider class $Q_0[a, b]$ of quasi-continuous functions anchored at a, has the Hellinger-integral representation

$$L(f) = \int_a^b \frac{df(x)\, dg(x)}{du(x)},$$

where u is an increasing function on $[a, b]$, and $g \in BV_u[a, b]$, the definition of which appears in [4, § 1]. Using the notation of [4], if $g_u^-(a)$ and $g_u^+(b)$ exist, then the Hellinger-type integral introduced by Webb can be expressed as a modified RS_2 integral—see [4, Thm. 5.2].

Representation of bounded linear functionals on $C^n[a, b]$. We now use the Riesz representation theorem to obtain a representation for bounded linear functionals on the space of functions having continuous nth derivatives on $[a, b]$. Accordingly, let $C^n[a, b]$ denote the space of functions having continuous nth derivatives on $[a, b]$; in particular, denote $C^0[a, b]$ by $C[a, b]$. Furthermore, let $\|f\|_0$ denote $\sup_{a \leqq x \leqq b} |f(x)|$ and let $\|f\|_n$ denote $\sum_{r=0}^n \sup_{a \leqq x \leqq b} |f^{(r)}(x)|$. Let L be a

bounded linear functional on $C^n[a, b]$, L being bounded in the nth sense; that is, there exists a constant M such that

$$|L(f)| \leqq M\|f\|_n \quad \text{for all } f \in C^n[a, b].$$

We now define a linear functional T on $C[a, b]$ by

$$\begin{aligned} T(\phi) &= L\Big(\int_a^x \int_a^{x_{n-1}} \cdots \int_a^{x_1} \phi(x_0)\, dx_0 \cdots dx_{n-1}\Big) \\ &= L\Big(\int_a^x \frac{(x-t)^{n-1}}{(n-1)!}\phi(t)\, dt\Big), \qquad n \geqq 1. \end{aligned} \tag{8}$$

Then T is certainly linear, and since L is bounded in the nth sense,

$$|T(\phi)| \text{ is certainly bounded for all } \phi \in C[a, b].$$

Hence T is a bounded linear functional on $C[a, b]$, and so by the Riesz representation theorem, there exists $g \in BV[a, b]$ such that

$$T(\phi) = \int_a^b \phi(x)\, dg(x), \qquad \phi \in C[a, b].$$

Therefore, using (8), Taylor's theorem, integration-by-parts for Stieltjes integrals, Theorem 3 and [6, § 4], we have

$$\begin{aligned} L(f) - \sum_{s=0}^{n-1} \frac{f^{(s)}(a)}{s!} L\{(x-a)^s\} &= L\Big(\int_a^x \frac{(x-t)^{n-1}}{(n-1)!} f^{(n)}(t)\, dt\Big) \\ &= T(f^{(n)}) = \int_a^b f^{(n)}(x)\, dg(x) \\ &= f^{(n)}(b)g(b) - f^{(n)}(a)g(a) - \int_a^b g(x)\, df^{(n)}(x) \\ &= f^{(n)}(b)g(b) - f^{(n)}(a)g(a) \\ &\quad - n!(M_1)\int_a^b g(x)\frac{d^{n+1}f(x)}{dx^n}. \end{aligned}$$

Conversely, $(M_1)\int_a^b g(x)[d^{n+1}f(x)/dx^n]$ is a bounded linear function on $C^n[a, b]$ since

$$\begin{aligned} n!\left|(M_1)\int_a^b g(x)\frac{d^{n+1}f(x)}{dx^n}\right| &\leqq |f^{(n)}(b)g(b) - f^{(n)}(a)g(a)| + \left|\int_a^b f^{(n)}(x)\, dg(x)\right| \\ &\leqq \|f\|_n\{|g(b)| + |g(a)| + V(g; a, b)\}. \end{aligned}$$

We summarize the previous discussion in the following:

THEOREM 5. *Let L be a bounded linear functional on $C^n[a, b]$, L being bounded in the n-th sense. Then there exists a function g of bounded variation on $[a, b]$ such that*

$$\begin{aligned} L(f) = \sum_{s=0}^{n-1} \frac{f^{(s)}(a)}{s!} L\{(x-a)^s\} + f^{(n)}(b)g(b) - f^{(n)}(a)g(a) \\ - n!(M_1)\int_a^b g(x)\frac{d^{n+1}f(x)}{dx^n}. \end{aligned}$$

Conversely, $(M_1)\ \int_a^b g(x)[d^{n+1}f(x)/dx^n]$, *formed with a fixed function g of bounded variation, is a linear functional on* $C^n[a, b]$, *and is bounded in the n-th sense.*

Generalized functions. Let H be defined on $[a, b]$ by

$$H(x) = 0 \quad \text{when } x > c,$$
$$= 1 \quad \text{when } x > c, \quad \text{where } a < c < b.$$

Note that $H \notin BV_k[a, b]$, $k \geqq 2$. It is then not difficult to show, using [6, § 4], that when $f^{(k-1)}$ is continuous at $c \in (a, b)$, $(f, H) \in M_3RS_k[a, b]$, and

$$(M_3) \int_a^b f(x) \frac{d^k H(x)}{dx^{k-1}} = \frac{(-1)^{k-1}}{(k-1)!} f^{(k-1)}(c).$$

In view of this result, for example, it is not surprising that our generalized Riemann–Stieltjes integrals should appear in the context of generalized functions. We conclude with some brief observations.

As in [1, § 1.2], let K be the set of all real functions ϕ with continuous derivatives of all orders, and with bounded support. We will further assume that the support of each function is a closed interval.

Thus, if $k \geqq 1$, and g is a fixed function of bounded kth variation on $[a', b']$, then the RS_k integral $\int_a^b \phi(x)[d^k g(x)/dx^{k-1}]$, $\phi \in K$, is a bounded linear functional on K and can thus be interpreted as a generalized function—see [1, § 1.3]. We develop this idea a little further and in particular discuss the derivative properties of generalized functions, namely, that if T is a generalized function, and n is a positive integer, then the nth derivative of T is given by

$$T^{(n)}(\phi) = (-1)^n T(\phi^{(n)}), \qquad \phi \in K.$$

Let g be a fixed function belonging to $\bigcap_{k=1}^{\infty} BV_k[a, b]$, and suppose further that it vanishes when $x \leqq a$ and $x \geqq b$. Then, according to [5, Thm. 20], $g \in C^{(\infty)}[a, b]$. Consequently $g \in K$. Now, according to Theorem 2, when k is any positive integer,

$$k! \int_a^b \phi(x) \frac{d^{k+1} g(x)}{dx^k} = (-1)^k \int_a^b \phi^{(k)}(x)\, dg(x)$$

for all $\phi \in K$, and so, if $T(\phi) = \int_a^b \phi(x)\, dg(x)$, the kth derivative of T is given by the $(k+1)$th order Riemann–Stieltjes integral, $k! \int_a^b \phi(x) d^{k+1} g(x)/dx^k$; that is,

$$T^{(k)}(\phi) = k! \int_a^b \phi(x) \frac{d^{k+1} g(x)}{dx^k}.$$

We summarize the previous discussion in

THEOREM 6. *Whenever a distribution T can be represented in the form*

$$T(\phi) = \int_a^b \phi(x)\, dg(x),$$

where $\phi \in K$, $g \in \bigcap_{k=1}^{\infty} BV_k[a, b]$, *and g vanishes when* $x \leqq a$ *and* $x \geqq b$, *then*

$$T^{(k)}(\phi) = k! \int_a^b \phi(x) \frac{d^{k+1} g(x)}{dx^k}, \qquad k \geqq 0.$$

REFERENCES

[1] I. M. GEL'FAND AND G. E. SHILOV, *Generalized Functions*, Vol. 1, Academic Press, New York and London, 1964.
[2] F. RIESZ AND B. SZ-NAGY, *Functional Analysis*, Frederick Ungar, New York, 1955.
[3] A. M. RUSSELL, *Extensions of the Riemann–Stieltjes integral*, M.Sc. thesis, University of Melbourne, Melbourne, Australia, 1963.
[4] ———, *Functions of bounded second variation and Stieltjes-Type integrals*, J. London Math. Soc. (2), 2 (1970), pp. 193–208.
[5] ———, *Functions of bounded kth variation*, Proc. London Math. Soc. (3), 26 (1973), pp. 547–563.
[6] ———, *Stieltjes-Type Integrals*, J. Austral. Math. Soc., 20 (1975), pp. 431–448.
[7] L. SCHWARTZ, *Théorie des distributions*, Hermann, Paris, 1966.
[8] J. R. WEBB, *A Hellinger integral representation for bounded linear functionals*, Pacific J. Math., 20 (1967), pp. 327–337.

SIAM J. MATH. ANAL.
Vol. 8, No. 5, October 1977

MONOTONE APPROXIMATION BY SPLINES*

RONALD A. DE VORE†

Abstract. We prove Jackson type estimates for the approximation of monotone nondecreasing functions by monotone nondecreasing splines with equally spaced knots. Our results are of the same order as the Jackson type estimates for unconstrained approximation by splines with equally spaced knots.

1. Introduction. We are interested in how well we can approximate a monotone nondecreasing function by monotone nondecreasing splines. For $r, n \geqq 1$, let $\mathcal{S}(r, n)$ denote the space of splines of order r (degree $r-1$) with knots $\{i/n\}_0^n$, i.e., $S \in \mathcal{S}(r, n)$ if and only if $S^{(r-2)}$ is continuous on $[0, 1]$ and on each interval $[i/n, (i+1)/n]$, $i = 0, 1, \cdots, n-1$, S is a polynomial of degree $\leqq r-1$. If f is a monotone, nondecreasing function on $[0, 1]$ ($f \uparrow$), then we define the error of monotone approximation by splines to be

$$E_n^*(f, r) \equiv \inf_{S \in \mathcal{S}^*(r, n)} \|f - S\|,$$

where $\|\cdot\|$ is the supremum norm on $[0, 1]$ and $\mathcal{S}^*(r, n)$ is the set of those splines S in $\mathcal{S}(r, n)$ with $S \uparrow$. The question then is how fast does $E_n^*(f, r) \to 0$, $n \to \infty$, in relation to the smoothness of f? Our main result is the following theorem.

THEOREM 1. *There is a constant $C > 0$, depending only on r, such that whenever $f \uparrow$ and $f^{(k)}$ is continuous, $0 \leqq k \leqq r-1$, then*

$$E_n^*(f, r) \leqq C n^{-k} \omega(f^{(k)}, n^{-1}), \qquad n = 1, 2, \cdots. \tag{1.1}$$

This is a Jackson type theorem for monotone approximation by splines. It is exactly the same as the Jackson type theorem for unrestricted spline approximation.

Theorem 1 shows that monotone approximation by splines is as efficient as the unrestricted approximation by splines, at least in the sense of Jackson type estimates of the form (1.1). There is some deficiency in (1.1) however, in that it is preferable to give the Jackson type estimates in terms of the rth order modulus of smoothness, $\omega_r(f, t)$, rather than just the first order as in (1.1). The rth order modulus of smoothness is needed to completely characterize the degree of approximation by splines in the unrestricted case in terms of both direct and inverse theorems (see K. Scherer [9]). We have more to say on this in the remarks section (§ 6).

Theorem 1 is already known in the case of approximation by step functions or piecewise linear functions ($r = 1, 2$) and also in the general case $r \geqq 1$ if $k = 0$ or 1. Here, the variation diminishing splines of Schoenberg are monotone when f is and they also provide the estimate (1.1) in case $k = 0, 1$ (see M. Marsden [7] and De

* Received by the editors April 25, 1975, and in revised form April 14, 1976.

† Department of Mathematics, Oakland University, Rochester, Michigan. This work was supported by the National Science Foundation under Grant GP 19620.

Vore [3]). The variation diminishing splines can not give the result (1.1) for $k \geqq 2$ because they are positive operators and hence saturated. This limitation exists not only for variation diminishing splines but for any linear method of approximation since any such method would have to preserve positivity (of the derivative of f) and hence be restricted in its degree of approximation by the saturation phenomena for positive linear operators. Thus, we must go to nonlinear techniques to prove the general case of (1.1) and this makes the situation more difficult. For example, there is no easy proof of (1.1) even for quadratic or cubic splines.

This situation is paralleled in monotone approximation by polynomials. For this problem, G. G. Lorentz and K. Zeller [6] and G. G. Lorentz [5] have given Jackson type theorems of the same form as (1.1), for $k = 0, 1$ (here n is the degree of the approximating polynomials), but these results have not been extended to $k \geqq 2$. It is possible to use the techniques and results of this paper to prove the higher order Jackson theorems for monotone approximation by polynomials. This is given in the next article of this journal [10].

Our proof of (1.1) is somewhat complicated by the presence of many constants whose actual values are usually not important but they are sometimes used in the definition of other constants. We will use the following conventions in labeling constants. Constants that appear in inequalities for general splines, e.g. B-splines, will be denoted by α_1, α_2, etc. Constants that appear in the approximation of f or its derivatives by splines will be denoted by C_1, C_2, etc. Constants that appear in upper estimates for f' will be denoted by A_1, A_2, etc., while constants appearing in lower estimates for f' will be denoted by $B_1, B_2, \cdots$.

2. *B*-Splines. We will on several occasions have to make local corrections of splines. This is best done by using splines with small support, such as the B-splines. Let $t_j = j/n$, for $j = 0, 1, \cdots, n$, $t_j = 0$ for $j < 0$, and $t_j = 1$, for $j > n$. If $M(x; t) = r(t-x)_+^{r-1}$ then the B-splines of order r are given by

$$N_{i,n,r}(x) = M(x; t_i, \cdots, t_{i+r}), \qquad -r+1 \leqq i \leqq n-1,$$

where the notation means that for fixed x we take the rth divided difference of M with respect to the variable t at the points $t_i, \cdots, t_{i+r}$. We mention now some properties of B-splines which can be found either in [1] or [2].

The B-splines have minimal support. For each i, $N_{i,n,r}$ vanishes outside of (t_i, t_{i+r}) and is strictly positive on (t_i, t_{i+r}). With our normalization we have

$$\int_{-\infty}^{\infty} N_{i,n,r}(x)\, dx = 1.$$

Actually, we will be more interested in the B-splines of order $r-1$ (degree $r-2$) since these will be used in the approximation of f'. Accordingly, let us introduce the notation that $N_i = N_{i,n,r-1}$ with the n and $r-1$ being understood. There are constants $\alpha_1 \leqq 1$ and $\alpha_2 \geqq 1$, which depend only on r, such that

$$N_i(x) \geqq \alpha_1 n, \qquad t_i + n^{-1} \leqq x \leqq t_{i+r-1} - n^{-1}, \tag{2.1}$$

$$N_i(x) \leqq \alpha_2 n, \qquad -\infty < x < \infty. \tag{2.2}$$

The B-splines N_i form a basis for $\mathcal{S}(r-1, n)$. If $S \in \mathcal{S}(r-1, n)$, then there are unique constants $(a_i)_{-r+2}^{n-1}$, such that $S = \sum a_i N_i$. Also, there is a constant $\alpha_3 \geqq 1$, which depends only on r, such that

$$|a_i| \leqq \alpha_3 n^{-1} \sup_{t_i \leqq x \leqq t_{i+r-1}} |S(x)|. \tag{2.3}$$

This last inequality follows from the C. deBoor–G. Fix formula for quasi-interpolation [1] which gives a formula for a_i in terms of derivatives of S on (t_i, t_{i+r-1}). We use Markov's inequality to replace derivatives of S by the supremum of S over (t_i, t_{i+r-1}).

3. Interpolation techniques. There is a very useful idea in approximation that to prove a result like (1.1), it is frequently enough to prove the result for only the end point, which in our case is when f has a bounded rth derivative. This is accomplished by using an interpolation argument to derive the general result from the end point result. The argument relies on replacing the arbitrary function f by a function which has a bounded rth derivative (controlled by the smoothness of f) and approximates f well. More precisely, if $\varepsilon > 0$, $f \in C^{(k)}[0, 1]$, and $k < r$, then there is a function g_ε with the following properties:

$$\|f - g_\varepsilon\| \leqq C_1 \varepsilon^k \omega(f^{(k)}, \varepsilon), \tag{3.1}$$

$$\|g_\varepsilon^{(r)}\| \leqq C_1 \varepsilon^{k-r} \omega(f^{(k)}, \varepsilon) \tag{3.2}$$

where $C_1 \geqq 1$ is a constant depending only on r (see e.g. G Freud and V. Popov [4]).

In the case of monotone approximation, in order to use this technique directly, we would need to know that when $f \uparrow$, then the functions g_ε can also be chosen to be nondecreasing. This does not follow from the Freud–Popov construction and it is not known whether this is actually the case. However, it still will be useful to use this interpolation technique in some of our proofs. We will also use the fact that when $k \geqq 1$, g_ε can be chosen to satisfy

$$\|f' - g_\varepsilon'\| \leqq C_1 \varepsilon^{k-1} \omega(f^{(k)}, \varepsilon). \tag{3.3}$$

In fact, the g_ε given by Freud–Popov already satisfies (3.3).

Let us point out how (3.1) and (3.2) can be used in the proof of the unrestricted version of (1.1), since some of our later arguments are based on this approach. The idea follows V. Popov and Bl. Sendov [8]. We prove first that if g is any function with $\|g^{(r)}\| \leqq M$, then there is a spline $S \in \mathcal{S}(r, n)$, such that

$$\|g - S\| \leqq CMn^{-r}, \tag{3.4}$$

with C depending only on r. This is proved by establishing the more general statement that for each $j = 1, 2, \cdots, r$, there is an $S \in \mathcal{S}(j, n)$, with $\|g^{(r-j)} - S_j\| \leqq 2^j r^j M n^{-j}$. For $j = 1$, the function S_1 can be taken as $S_1(x) = g^{(r-1)}(i/n)$, $x \in [in^{-1}, (i+1)n^{-1}$, $i = 0, 1, \cdots, n-1$.

Suppose then that we have shown the existence of a spline $S_j \in \mathcal{S}(j, n)$ with $\|g^{(r-j)} - S_j\| \leqq 2^j r^j M n^{-j}$. Let $y_i = rin^{-1}$, $i = 0, 1, \cdots, \lambda$, with $\lambda = [n/r]$. Define

$$a_{i,j} = \int_{y_i}^{y_{i+1}} \{g^{(r-j)}(t) - S_j(t)\}\, dt, \qquad i = 0, 1, \cdots, \lambda - 1.$$

We can take

$$S_{j+1}(x) = \int_0^x \left\{ S_j(t) + \sum_{i=0}^{\lambda-1} a_{i,j} N_{ir,n,j}(t) \right\} dt + g^{(r-j-1)}(0).$$

The spline S_{j+1} is in $\mathcal{S}(j+1, n)$ and $S_{j+1}(y_i) = g^{(r-j-1)}(y_i)$, $0 \leqq i \leqq \lambda$. Hence,

$$\begin{aligned} |g^{(r-j-1)}(x) - S_{j+1}(x)| &= \left| \int_{y_i}^x (g^{(r-j)}(t) - S_j(t))\, dt \right| + |a_{i,j}| \\ &\leqq 2^{j+1} r^j M n^{-j} |y_{i+1} - y_i| \leqq 2^{j+1} r^{j+1} M n^{-j-1}, \qquad x \in [y_i, y_{i+1}). \end{aligned}$$

The same estimate holds on $[y_\lambda, 1]$. This shows (3.4).

Now, we take $\varepsilon = n^{-1}$ and g_ε as a function which satisfies (3.1) and (3.2). Let S be the spline which satisfies (3.4) for $g = g_\varepsilon$. Then,

$$\begin{aligned} \|f - S\| &\leqq \|f - g_\varepsilon\| + \|g_\varepsilon - S\| \leqq (C_1 + CC_1) n^{-k} \omega(f^{(k)}, n^{-1}) \\ &\leqq C_2 n^{-k} \omega(f^{(k)}, n^{-1}) \end{aligned} \tag{3.5}$$

with $C_2 > C_1$ a constant depending only on r. This is the unrestricted analogue of (1.1).

Because of (3.3) and the way S is constructed, we also have the estimate

$$\|f' - S'\| \leqq C_2 n^{-k+1} \omega(f^{(k)}, n^{-1}). \tag{3.6}$$

The spline S has a piecewise $(r-1)$st derivative which is of course a step function. The jumps in this step function are controlled because of (3.2) and the construction of S. Namely,

$$\left| \text{jump } S^{(r-1)}\left(\frac{j}{n}\right) \right| \leqq C_2 n^{r-k-1} \omega(f^{(k)}, n^{-1}) \tag{3.7}$$

for $j = 1, \cdots, n-1$. In some sense, we have come full circle since the spline S can be used as the function g_ε when $\varepsilon = n^{-1}$ except that S does not have a continuous rth derivative, but instead we have (3.7). These results about the spline S in unrestricted approximation will be used later.

4. Decomposition of monotone functions. As we have observed in the Introduction, the estimate (1.1) is already known for $k = 0, 1$ and so we will assume from here on that $k \geqq 2$ and $r > 2$. In particular, f' is then continuous.

Note that if f' is strictly positive on $[0, 1]$ then the spline S introduced in the previous section will be monotone nondecreasing when n is sufficiently large because of (3.6). This spline also approximates with the correct order to give (1.1) because of (3.5). Thus, we see that the real difficulty in proving Theorem 1 will be when f' has zeros. The idea then is to decompose f in such a way that we have good

control over the derivative of f. This is done by isolating certain kinds of intervals on which f' is small, while on the remaining intervals, f' is large at a suitable number of places. We begin by introducing four types of intervals. The function f' will be small on the intervals of type 1, 2, and 3, while on the type 4 intervals f' will be large at least some of the time.

Let $A_1 = 100r^2 2^{r^4} C_2 \alpha_1^{-1} \alpha_2^2$, where $C_2 \geq 1$ is the constant that appears in (3.5)–(3.7) and $\alpha_1 \leq 1$ and $\alpha_2 \geq 1$ are the constants that appear in (2.1) and (2.2). The constant A_1 depends only on r. Let $\varepsilon_n = n^{-k}\omega(f^{(k)}, n^{-1})$. An interval I is said to be of type 1 if

$$I = [i_1 n^{-1}, i_2 n^{-1}], \text{ with } i_1, i_2 \text{ integers}, i_2 - i_1 \geq r^2, \text{ and } I \subseteq [0, 1], \tag{4.1}$$

$$f'(x) \leq A_1 n\varepsilon_n, \qquad x \in I, \tag{4.2}$$

$$\text{if } J = [j_1 n^{-1}, j_2 n^{-1}], \text{ with } j_1, j_2 \text{ integers}, I \subseteq J \text{ and } f'(x) \leq A_1 n\varepsilon_n, x \in J, \text{ then } J = I. \tag{4.3}$$

The condition (4.3) guarantees that I is a maximal interval on which (4.2) holds.

Let $A_2 = \alpha_3 A_1^2$, where α_3 is the constant of (2.3). Then $A_2 > A_1$. If $I = [i_1 n^{-1}, i_2 n^{-1}]$ is an interval of type 1, then let $j_1 \leq i_1$ be the smallest integer such that

$$f'(x) \leq A_2 n\varepsilon_n, \qquad x \in [j_1 n^{-1}, i_1 n^{-1}]. \tag{4.4}$$

Similarly, let $j_2 \geq i_2$ be the largest integer such that

$$f'(x) \leq A_2 n\varepsilon_n, \qquad x \in [i_2 n^{-1}, j_2 n^{-1}]. \tag{4.5}$$

We call the intervals $[j_1 n^{-1}, i_1 n^{-1}]$ and $[i_2 n^{-1}, j_2 n^{-1}]$, intervals of type 2.

When we remove the intervals of type 1 and type 2 from $[0, 1]$, then we are left with a finite number of intervals. Such a left over interval is said to be of type 3 or type 4 according to whether the length of the interval is less than $2r^2 n^{-1}$ or greater than or equal to $2r^2 n^{-1}$ respectively. It turns out that f' is also small on intervals of type 3. This follows from our first lemma which shows that if f' is small on an interval it is also small on adjacent intervals.

LEMMA 1. *Suppose* $0 \leq a \leq 1 - n^{-1}$, *and*

$$f'(x) \leq M, \qquad x \in [a, a + n^{-1}]. \tag{4.6}$$

Then, for any integer $l > 1$, *we have*

$$f'(x) \leq 2^{lk^2}(M + n\varepsilon_n), \qquad x \in [a - ln^{-1}, a + (l+1)n^{-1}] \cap [0, 1]. \tag{4.7}$$

Proof. We derive the estimate for $[a + n^{-1}, a + (l+1)n^{-1}] \cap [0, 1]$. The other case is proved in the same way. Suppose first that $x \in [a + n^{-1}, a + n^{-1} + k^{-1}n^{-1}]$. We use standard notation $\Delta_h^k(g, x)$ to denote the kth difference of g with step size h at the point x. Take $h = k^{-1}(x - a)$. So, $a + kh = x$ and $a + (k-1)h \leq a + n^{-1}$. Hence, with $g = f'$, we have

$$\begin{aligned} g(x) = g(a + kh) &\leq |\Delta_h^k(g, a)| + |\Delta_h^k(g, a) - g(a + kh)| \\ &\leq |\Delta_h^k(g, a)| + (2^k - 1)M, \end{aligned} \tag{4.8}$$

where we used the fact that the second term in absolute values only involves values of g on $[a, a + n^{-1}]$, where (4.6) holds.

Now for any y, we have $\Delta_h^{k-1}(g, y) = h^{k-1}g^{(k-1)}(\xi) = h^{k-1}f^{(k)}(\xi)$, with $y < \xi < y + kh$. Thus,

$$|\Delta_h^k(g, a)| = |\Delta_h^{k-1}(g, a+h) - \Delta_h^{k-1}(g, a)| \leqq h^{k-1}|f^{(k)}(\xi_1) - f^{(k)}(\xi_2)|,$$

with ξ_1 and ξ_2 in $[a, a + n^{-1} + k^{-1}n^{-1}]$. Hence, because $h = k^{-1}n^{-1}$, we have

$$\begin{aligned} |\Delta_h^k(g, a)| &\leqq k^{-k+1}n^{-k+1}\omega(f^{(k)}, n^{-1} + k^{-1}n^{-1}) \\ &\leqq 2k^{-k+1}n^{-k+1}\omega(f^{(k)}, n^{-1}) \leqq n\varepsilon_n. \end{aligned}$$

Here we have used the fact that $k \geqq 2$. If we use this inequality back in (4.8), we find that

$$f'(x) = g(x) \leqq (2^k - 1)M + n\varepsilon_n \leqq 2^k M + n\varepsilon_n, \qquad x \in [a, a + n^{-1} + k^{-1}n^{-1}].$$

This extends our original inequality (4.6) to the larger interval $[a, a + n^{-1} + k^{-1}n^{-1}]$. Now, we repeat this procedure lk times to find

$$f'(x) \leqq 2^{lk^2}M + (2^{lk^2-1} + \cdots + 2^k + 1)n\varepsilon_n \leqq 2^{lk^2}(M + n\varepsilon_n),$$

as desired.

As an immediate consequence of Lemma 1, we have the following lemma which shows that f' is small on intervals of type 3.

LEMMA 2. *There is a constant $A_3 > 0$, which depends only on r, such that for any interval I of type 3, we have*

$$f'(x) \leqq A_3 n\varepsilon_n, \qquad x \in I. \tag{4.9}$$

Proof. Since I is an interval of type 3, it must be adjacent to either an interval of type 1 or of type 2. Suppose that $[a, b]$ is this adjacent interval and assume that $[a, b]$ is to the right of I, so that $a \in \bar{I}$. The other case is handled in the same way. On the interval $[a, a + n^{-1}]$, we have

$$f'(x) \leqq A_2 n\varepsilon_n, \qquad x \in [a, a + n^{-1}],$$

because of (4.2) if $[a, b]$ is of type 1 and because of (4.4) if $[a, b]$ is of type 2. We know from Lemma 1 that this inequality can be extended to adjacent intervals. Indeed, Lemma 1 gives

$$f'(x) \leqq 2^{2r^2k^2}(A_2 n\varepsilon_n + n\varepsilon_n) \leqq 2^{2r^4}(A_2 + 1)n\varepsilon_n,$$

because I has length $< 2r^2n^{-1}$. This proves the lemma with $A_3 = 2^{2r^4}(A_2 + 1)$.

While f' is small on intervals of type 3, on intervals of type 4 there are always places where f' is suitably large as our next lemma shows.

LEMMA 3. *Let $B_2 = 2^{-r^4}A_2$. Suppose $I = [i_1 n^{-1}, i_2 n^{-1}]$ is an interval of type 4. Consider the intervals $J_\nu = [(i_1 + \nu r)n^{-1}, (i_1 + (\nu + 1)r)n^{-1}]$, $\nu = 0, 1, \cdots, r-1$. If $i_1 > 0$, then for some value of ν, we have*

$$f'(x) \geqq B_2 n\varepsilon_n, \qquad x \in J_\nu. \tag{4.10}$$

Similarly, consider the intervals $J'_\nu = [(i_2 - (\nu + 1)r)n^{-1}, (i_2 - \nu r)n^{-1}]$, $\nu = 0, 1, \cdots, r-1$. If $i_2 < n$, then there is a value of ν such that

$$f'(x) \geqq B_2 n\varepsilon_n, \qquad x \in J'_\nu. \tag{4.11}$$

Proof. We will prove (4.10). The proof of (4.11) is the same. First observe that we cannot have $f'(x) \leqq B_2 n\varepsilon_n$, $x \in J_\nu$, for any $\nu = 0, 1, \cdots, r-1$. Otherwise, we could use Lemma 1 to extend this inequality to the left and find

$$f'(x) \leqq 2^{r2k^2}(B_2 n\varepsilon_n + n\varepsilon_n) \leqq 2^{r^4} B_2 n\varepsilon_n, \qquad x \varepsilon J_0. \tag{4.12}$$

Here, we used the facts that $B_2 \geqq 1$ and $k \leqq r-1$. Now, (4.12) and the fact that $2^{r^4} B_2 = A_2$ shows that $f'(x) \leqq A_2 n\varepsilon_n$ which means J_0 should be contained in the type 2 interval immediately to the left of J_0 (here is where we need $i_1 > 0$). This is a contradiction.

So, now we know that for each ν, there are points $\xi_\nu \in J_\nu$ for which $f'(\xi_\nu) \geqq B_2 n\varepsilon_n$. On the other hand, suppose that for each ν, there are points $\xi'_\nu \in J_\nu$, with $f'(\xi'_\nu) < B_2 n\varepsilon_n$. Again, we must find a contradiction. This is done as follows. By the continuity of f', there is for each ν a point $x_\nu \in J_\nu$ such that $f'(x_\nu) = B_2 n\varepsilon_n$. Also on J_0, there is a point $x_0 \in J_0$ with $f'(x_0) = A_2 n\varepsilon_n$. Otherwise we would have $f'(x) < A_2 n\varepsilon_n$, $x \in J_0$, which again would put J_0 in the interval of type 2 immediately to the left of J_0.

Now, we want to compute the divided difference of $g = f'$ at the points x_ν. There exist points $\eta_1, \eta_2 \in [i_1 n^{-1}, (i_1 + r^2)n^{-1}]$ with

$$\begin{aligned} g^{(k-1)}(\eta_1) &= (k-1)!\, g[x_0, x_1, \cdots, x_{k-1}] \\ &= (k-1)!(A_2 - B_2) n\varepsilon_n \prod_{i=1}^{k-1} (x_0 - x_i)^{-1}, \end{aligned}$$

and likewise

$$g^{(k-1)}(\eta_2) = (k-1)!\, g[x_1, x_2, \cdots, x_k] = 0.$$

Hence,

$$|A_2 - B_2| \leqq ((k-1)! n\varepsilon_n)^{-1} |g^{(k-1)}(\eta_1) - g^{(k-1)}(\eta_2)| \prod_{i=1}^{k-1} (x_i - x_0). \tag{4.13}$$

On the other hand,

$$\begin{aligned} |g^{(k-1)}(\eta_1) - g^{(k-1)}(\eta_2)| &= |f^{(k)}(\eta_1) - f^{(k)}(\eta_2)| \\ &\leqq \omega(f^{(k)}, r^2 n^{-1}) \leqq r^2 \omega(f^{(k)}, n^{-1}) = r^2 n^k \varepsilon_n \end{aligned}$$

and

$$\prod_{i=1}^{k-1} (x_i - x_0) \leqq (r^2)^{k-1} n^{-k+1}.$$

Putting this back in (4.13) gives the estimate

$$|A_2 - B_2| \leqq ((k-1)!)^{-1} r^{2k} \leqq r^{2r}. \tag{4.14}$$

From the very definition of the constants A_1, A_2 and B_2, we find

$$|A_2 - B_2| \geqq A_1 \geqq r^2 2^{r^4} > r^{2r}, \tag{4.15}$$

since $r \geqq 2$. The estimate (4.14) and (4.15) contradict one another and therefore we have proved the lemma.

Besides guaranteeing that f' is large on parts of intervals of type 4, we can also show that f' is large on parts of any interval that does not intersect a type 1 interval.

LEMMA 4. *Let* $B_1 = 2^{-r^4}A_1$. *If* $J = [jn^{-1}, (j+r^2)n^{-1}]$ *is any interval in* $[0, 1]$ *that does not intersect any interval of type* 1, *then for one of the intervals* $J_\nu = [(j+\nu r)n^{-1}, (j+(\nu+1)r)n^{-1}]$, $\nu = 0, 1, \cdots, r-1$, *we have*

$$f'(x) \geqq B_1 n\varepsilon_n, \qquad x \in J_\nu. \tag{4.16}$$

Proof. This proof is almost identical to that of Lemma 3. If $0 < \nu < r-1$, then we cannot have

$$f'(x) \leqq B_1 n\varepsilon_n, \qquad x \in J_\nu,$$

since otherwise this inequality could be extended by using Lemma 1 to give $f'(x) \leqq A_1 n\varepsilon_n$, $x \in J$, which means that J is already an interval of type 1, as is not the case. Thus, for each $\nu = 0, 1, \cdots, r-1$ there is a point $x_\nu \in J$, and $f'(x_\nu) = B_1 n\varepsilon_n$. This means that with $g = f'$, there is an $\eta_1 \in J$, with

$$g^{(k-1)}(\eta_1) = g[x_0, x_1, \cdots, x_{k-1}] = 0.$$

Assume that (4.16) does not hold. Since J does not intersect any type 1 interval, there must be a point $y \in J$ with $f'(y) = A_1 n\varepsilon_n$. This means we can find points $y_1 < y_2 < \cdots < y_k$ in J with $f'(y_i) = B_1 n\varepsilon_n$, $i \neq i_0$ and $f'(y_{i_0}) = A_1 n\varepsilon_n$. So, there is an $\eta_2 \in J$, with

$$g^{(k-1)}(\eta_2) = g[y_1, \cdots, y_k] = (k-1)!(A_1 - B_1)n\varepsilon_n \prod_{\substack{i=1 \\ i \neq i_0}}^{k} (y_{i_0} - y_i)^{-1}.$$

Arguing as in the proof of Lemma 3, we get

$$|A_1 - B_1| \leqq ((k-1)!n\varepsilon_n)^{-1} |g^{(k-1)}(\eta_1) - g^{(k-1)}(\eta_2)| \prod_{\substack{i=1 \\ i \neq i_0}}^{k} (y_i - y_{i_0})^{-1}$$

$$\leqq r^{2k} \leqq r^{2r}.$$

While on the other hand from the value of the constants, we have

$$|A_1 - B_1| \geqq \tfrac{1}{2}A_1 \geqq \tfrac{1}{2}r^2 2^{r^4} \geqq r^{2r}.$$

This is the desired contradiction and the lemma is proved.

Now that we have given the important properties of intervals of types 1–4, we can give our decomposition for f. Let $I_1, \cdots, I_m$ be the intervals of type 4. For each $j = 1, \cdots, m$, let I_j^* denote the closure of the union of I_j with I_j^l and I_j^r where I_j^l is the interval of type 2 adjacent and to the left of I_j and I_j^r is the interval of type 2 adjacent and to the right of I_j. The intervals I_j^l and I_j^r may be empty. If we delete $I_1^*, \cdots, I_m^*$ from $[0, 1]$, then we are left with a finite number of open intervals which we denote by $J_0^*, \cdots, J_m^*$ where J_0^* or J_m^* may be empty.

If I is one of the intervals $I_1^*, \cdots, I_m^*$ or $J_0^*, \cdots, J_m^*$ then we define f_I by

$$f_I(x) = \int_0^x f'(t)\chi_I(t)\, dt,$$

where χ_I is the characteristic function of I. Since the intervals are disjoint,

$$f(x) = f(0) + \sum_1^m f_{I_j^*}(x) + \sum_0^m f_{J_j^*}(x)$$

which is our decomposition of f.

5. Proof of Theorem 1. We will prove Theorem 1 by showing the existence of splines $S_{J_j^*}$, $S_{I_j^*} \in \mathcal{S}^*(r, n)$ with the properties

(5.1) *if* $J_j^* = (a, b)$, *then* $S_{J_j^*}(x) = f_{J_j^*}(x)$, *for* $x \notin [a - rn^{-1}, b + rn^{-1}]$,

(5.2) $\|f_{J_j^*} - S_{J_j^*}\| \leqq C\varepsilon_n$ *with* C *depending only on* r,

(5.3) *if* $I_j^* = [a, b]$, *then* $S_{I_j^*}(x) = f_{I_j^*}(x)$, *for* $x \notin [a - 2r^2n^{-1}, b + 2r^2n^{-1}]$,

(5.4) $\|f_{I_j^*} - S_{I_j^*}\| \leqq C\varepsilon_n$ *with* C *depending only on* r.

Before proceeding to the proofs of (5.1)–(5.4), let's first see how they give the theorem. Recall first that each of the intervals I_j^* and J_j^* has length $\geqq r^2/n$. Therefore, for a given $x \in [0, 1]$, it follows from (5.1) and (5.3) that there are at most five intervals I among $I_1^*, \cdots, I_m^*$ and $J_0^*, \cdots, J_m^*$ with $S_I(x) \neq f_I(x)$. Hence, if $S = f(0) + \sum S_{I_j^*} + \sum S_{J_j^*}$ then $S \uparrow$ and

$$\begin{aligned} \|f - S\| &\leqq 5 \max \left(\sup_j \|f_{I_j^*} - S_{I_j^*}\|, \sup_j \|f_{J_j^*} - S_{J_j^*}\|\right) \\ &\leqq 5C\varepsilon_n \end{aligned}$$

which gives the estimate (1.1).

Proof of (5.1)–(5.2). This is the easier of the two. Let $I = (i_1 n^{-1}, i_2 n^{-1})$ be one of the intervals J_j^* and write $i_2 - i_1 = \lambda(r-1) + \mu$, where λ and μ are integers with $\lambda \geqq 1$ and $0 \leqq \mu < r - 1$. Define $x_\nu = (i_1 + \nu(r-1))n^{-1}$, $\nu = 0, 1, \cdots, \lambda + 1$, and

$$b_\nu = f_I(x_\nu) - f_I(x_{\nu-1}), \qquad \nu = 1, 2, \cdots, \lambda + 1.$$

Let N_ν be the B-splines of *order* $r-1$ (*degree* $r-2$). Recall that N_ν is normalized to have integral one. Define

$$S_I(x) = \int_0^x \left[\sum_{\nu=0}^{\lambda} b_{\nu+1} N_{i_1+\nu(r-1)}(t)\right] dt.$$

Since each B-spline is nonnegative and the numbers $b_\nu \geqq 0$, we have $S_I \uparrow$. Since $f_I' = 0$ for x not in I, we have $S_I(x) = 0 = f_I(x)$, for $x \leqq x_0$ and $S_I(x) = S_I(x_{\lambda+1}) = f_I(x_{\lambda+1}) = f_I(x)$ for $x \geqq x_{\lambda+1}$. This shows that property (5.1) is satisfied for S_I.

For the estimate (5.2), we need only observe that $S_I(x_\nu) = f_I(x_\nu)$, $\nu = 0, 1, \cdots, \lambda + 1$. Thus, if $x_\nu \leqq x \leqq x_{\nu+1}$, then

$$|f_I(x) - S_I(x)| \leqq |f_I(x) - f_I(x_\nu)| + |S_I(x) - S_I(x_\nu)| \leqq 2b_{\nu+1} \tag{5.5}$$

because both f_I and S_I are nondecreasing. Now, the interval I is a union of intervals of type 1, type 2, and type 3. On any such interval, we have the estimate

$$f'(t) \leqq An\varepsilon_n,$$

where A is the maximum of the three constants A_1, A_2, and A_3 appearing in (4.2),

(4.4), and (4.9), respectively. Integrating this last inequality gives

$$b_{\nu+1} \leqq \int_{x_\nu}^{x_{\nu+1}} f'(t)\,dt \leqq (r-1)n^{-1}An\varepsilon_n \leqq (r-1)A\varepsilon_n,$$

where we have used the fact that $x_{\nu+1}-x_\nu=(r-1)n^{-1}$. Using this inequality back in (5.5) gives (5.2) provided that $C \geqq 2(r-1)A$.

Proof of (5.3)–(5.4). Now let $I=[i_1n^{-1}, i_2n^{-1}]$ be one of the intervals I_j^*, $j=1, 2, \cdots, m$. The construction of the spline S_I is more complicated in this case and it may be beneficial to sketch the idea of the construction before actually embarking on details.

We start with a spline S_1 which approximates f according to (3.5)–(3.7). Since we want to approximate f_I and f_I' vanishes outside of I, we need to modify S_1. We do this by working with the B-spline representation of S_1' and deleting all terms that do not contribute to S_1' on I. In this way, we get a new spline S_2' which agrees with S_1' on I but vanishes once we get a little away from I. Integrating S_2' gives a new spline S_2. Fortunately, the intervals immediately to the right and left of I are of type 1 and so f' is small on these intervals according to (4.2). This means that S_2' will be small on these intervals and so S_2 is a good approximation to f_I.

Unfortunately, the spline S_2 is not necessarily monotone nondecreasing. However, we do know that S_2' is not too negative. On I, $S_2'=S_1'$ which is controlled by (3.6). Outside of I, S_2' is small as mentioned above. What we do then is add a spline to S_2' to pull it up so it can not be negative. We then integrate to get a new spline S_3 which is monotone nondecreasing, sure enough, but we may have introduced too much error to still have (5.4).

The final step is to make a correction on each interval of length r^2n^{-1} to prevent the error from building up. This is done by using the fact that I is of type 4 and hence f' and so S_3' will be big at suitable places because of Lemma 3. What we do is pull down S_3' however much is necessary but still keep a positive derivative. The resulting spline when integrated will satisfy (5.3) and (5.4).

Now, to the actual details. We consider the case when I is strictly interior to $[0, 1]$ and so $r^2 \leqq i_1$ and $i_2 \leqq n-r^2$. When I contains one of the end points, the proof is similar and in fact somewhat easier, so we do not repeat the details. Let $S_1 \in \mathcal{S}(r, n)$ be a spline which satisfies (3.5)–(3.7) and let $S_1' = \sum a_\nu N_\nu$ be the B-spline representation of S_1'. We define

$$S_2(x) = \int_0^x \sum_{i_1-r+2}^{i_2-1} a_\nu N_\nu(t)\,dt. \tag{5.6}$$

We want to see that S_2 approximates f_I well and so we must show that S_2' is small outside of I. The interval $E_1=[(i_1-2r)n^{-1}, i_1n^{-1}]$ is contained in an interval of type 1 because each of the intervals I_j^* is the union of I_j with the left and right adjacent intervals of type 2. Also recall that intervals of type 1 have length $\geqq r^2n^{-1} > 2rn^{-1}$.

From (4.2) and (3.6), we find that

$$|S_1'(x)| \leqq |f'(x)| + |f'(x)-S_1'(x)| \leqq (A_1+C_2)n\varepsilon_n, \qquad x \in E_1. \tag{5.7}$$

Similarly, for $E_2=[i_2, (i_2+2r)n^{-1}]$,

$$|S_1'(x)|\leqq(A_1+C_2)n\varepsilon_n, \qquad x\in E_2. \tag{5.8}$$

These last two estimates can be used to estimate a_ν when $\nu\in F=\{\mu: i_1-2r+4\leqq\mu\leqq i_1-r+1 \text{ or } i_2\leqq\mu\leqq i_2+r-3\}$. For these values of ν, $(t_\nu, t_{\nu+r-1})\subseteq E_1\cup E_2$ and so from (5.7), (5.8), and (2.3), we find

$$|a_\nu|\leqq\alpha_3(A_1+C_2)\varepsilon_n, \qquad \nu\in F. \tag{5.9}$$

Now, we use (5.9) with (2.2) to get the following estimate on $E_3=[(i_1-r+2)n^{-1}, i_1n^{-1}]\cup[i_2n^{-1}, (i_2+r-2)n^{-1}]$:

$$|S_2'(x)-S_1'(x)|=|\sum_{\nu\in F}a_\nu N_\nu(x)|\leqq 2r\alpha_2\alpha_3(A_1+C_2)n\varepsilon_n, \qquad x\in E_3,$$

where we have used the fact that there are less than $2r$ integers ν in F. This last inequality together with (5.7) and (5.8) gives

$$\begin{aligned}|S_2'(x)|&\leqq|S_1'(x)|+|S_2'(x)-S_1'(x)|\\&\leqq 3r\alpha_2\alpha_3(A_1+C_2)n\varepsilon_n, \qquad x\in E_3\end{aligned} \tag{5.10}$$

where we have also used the facts that $\alpha_2, \alpha_3\geqq 1$.

The estimate (5.10) shows that S_2' is small when $x\in E_3$. When $x\in I$, then $S_2'(x)=S_1'(x)$ and for all other values of x, $S_2'(x)=0$. Therefore, since $f_I'=0$, $x\notin I$ and $f_I=f$, $x\in I$, we have

$$\begin{aligned}\|f_I-S_2\|&\leqq 2\|f-S_1\|+\int_{E_3}|S_2'(x)|\,dx\\&\leqq[2C_2+6r^2\alpha_2\alpha_3(A_1+C_2)]\varepsilon_n,\end{aligned} \tag{5.11}$$

where we have used the fact that $|E_3|\leqq 2rn^{-1}$.

The estimate (5.11) shows that S_2 is a good approximation to f_I. Unfortunately, S_2 is not necessarily nondecreasing and so we must make a correction. We know that S_2' is not too negative because of (3.6) and (5.10). Namely,

$$\begin{aligned}S_2'(x)=S_1'(x)&\geqq f'(x)-|S_1'(x)-f'(x)|\geqq f'(x)-C_2n\varepsilon_n\\&\geqq -C_2n\varepsilon_n, \qquad x\in I.\end{aligned} \tag{5.12}$$

$$S_2'(x)\geqq -3r\alpha_2\alpha_3(A_1+C_2)n\varepsilon_n, \qquad x\in E_3. \tag{5.13}$$

Of course, $S_2'(x)=0$, for all other values of x.

Let's define

$$S_3(x)=\int_0^x (S_2'(t)+T_2(t))\,dt$$

where

$$T_2=\gamma_1(N_{i_1-r+1}+N_{i_1-r+2}+N_{i_2-1}+N_{i_2})+\gamma_2\sum_{i_1-1}^{i_2-r+2}N_\nu$$

with $\gamma_1=3r\alpha_1^{-1}\alpha_2\alpha_3(A_1+C_2)\varepsilon_n$ and $\gamma_2=\alpha_1^{-1}C_2\varepsilon_n$. Because of (5.12), (5.13) and

(2.1), we see that $S_3'(x) \geqq 0$, $x \in [0, 1]$. Thus, S_3 is nondecreasing but we may have added too much error to still have a good approximation.

We want to modify S_3 to prevent the error from building up. First, let's see how much S_3 differs from S_2. Since $S_3' = S_2' + T_2$, what we have added in T_2 is the cause of the error in S_3 and we will have to subtract this at suitable places to prevent a buildup of error.

We can write $T_2 = T_{2,1} + T_{2,2}$, where $T_{2,1}$ is the sum of the B-splines in T_2 which have coefficients γ_1 (there are four of these) and $T_{2,2}$ is the remaining sum, which consists of all the B-splines with coefficients γ_2.

Let's first correct for $T_{2,1}$. By Lemma 3, (4.10), there is an interval $[l_0 n^{-1}, (l_0 + r)n^{-1}]$, with l_0 an integer, on which $f'(x) \geqq B_2 n^{-k+1} \omega(f^{(k)}, n^{-1})$. This interval is contained in $[a, a + r^2 n^{-1}]$ where $[a, b]$ is the interval of type 4 that makes up part of I. Recall that intervals of type 4 necessarily have length $\geqq 2r^2 n^{-1}$. We will subtract $4\gamma_1 N_{l_0}$ as our correction for $T_{2,1}$.

For the spline $T_{2,2}$, we have no control over the number of terms that appear and so we do not have the luxury of making just one correction. Instead, we will have to make a correction on each interval of length $r^2 n^{-1}$. This is done as follows. Let $i_2 = i_1 + \lambda r^2 + \mu$, with $\lambda > 1$ and $0 \leqq \mu < r^2$ and define $x_\nu = (i_1 + \nu r^2) n^{-1}$, $\nu = -1, 0, \cdots, \lambda + 2$. From Lemma 4, it follows that for each $\nu = 1, 2, \cdots, \lambda - 1$, there is an interval $[l_\nu n^{-1}, (l_\nu + r)n^{-1}] \subseteq [x_{\nu-1}, x_\nu]$, on which $f'(x) \geqq B_1 n \varepsilon_n$. Our correction on $[x_{\nu-1}, x_\nu]$ will be the spline $r^2 \gamma_2 N_{l_\nu}$, when $1 \leqq \nu \leqq \lambda - 2$.

Now, γ_2 appears a total of $\lambda r^2 + \mu + 4 - r$ times in the definition of $T_{2,2}$. We have already taken care of $(\lambda - 1)r^2$ of these terms. So, we have yet to take care of $s = r^2 + \mu - r + 4 \leqq 2r^2 - r + 4$ terms. We correct for these with the spline $s\gamma_2 N_{l_{\lambda-1}}$.

This corrects for all the error from T_2; however, we would also like our new spline to agree with f_I when $x \geqq x_{\lambda+2}$. To do this, let $[l_\lambda n^{-1}, (l_\lambda + r)n^{-1}]$ be the interval guaranteed by Lemma 3, (4.11). This interval is contained in $[(i_2 - r^2)n^{-1}, i_2 n^{-1}]$. The spline $\gamma_3 N_{l_\lambda}$, with $\gamma_3 = S_2(x_{\lambda+2}) - f_I(x_{\lambda+2})$ is our correction in this case

Thus, our total correction will be the spline

$$T_3 = 4\gamma_1 N_{l_0} + r^2 \gamma_2 \sum_{1}^{\lambda-2} N_{l_\nu} + s\gamma_2 N_{l_{\lambda-1}} + \gamma_3 N_{l_\lambda}.$$

So, we define

$$S_I(x) = S_3(x) - \int_0^x T_3(t)\, dt.$$

This is the spline that will satisfy (5.3) and (5.4).

The verification of (5.3) is quite easy. If $x \notin [x_{-1}, x_{\lambda+2}]$, then $f_I'(x) = 0 = S_I'(x)$. Also, in substracting T_3, we have accomplished two things, taking away exactly the error introduced by T_2 and then forcing an interpolation at $x_{\lambda+2}$. Thus, when $x \leqq x_{-1}$, $S_I(x) = 0 = f_I(x)$, and when $x \geqq x_{\lambda+2}$, $S_I(x) = S_I(x_{\lambda+2}) = f_I(x_{\lambda+2}) = f_I(x)$. Therefore, (5.3) is satisfied.

In order to check (5.4), we must check that indeed we have prevented the error from building up. Now, $S_I' = S_2' + T_2 - T_3$, and we already know by (5.11) that

$$\left\| f_I(x) - \int_0^x S_2'(t)\, dt \right\| = \|f_I - S_2\| \leqq C\varepsilon_n,$$

with C a constant depending only on r. Hence, we need only show that

$$\left|\int_0^x (T_2(t)-T_3(t))\,dt\right| \leqq C\varepsilon_n, \qquad x\in[x_{-1}, x_{\lambda+2}], \tag{5.14}$$

again with C a constant depending only on r. Now, for $x\in[x_{-1}, x_{\lambda+2}]$, the two integrals in (5.14) can differ by at most $4\gamma_1+2r^2\gamma_2+|\gamma_3|$, because any of the terms involving γ_2 in T_2 are taken care of by a corresponding term with coefficient γ_2 in T_3, within a length of $2r^2n^{-1}$. Therefore,

$$\left|\int_0^x (T_2(t)-T_3(t))\,dt\right| \leqq 4\gamma_1+2r^2\gamma_2+|\gamma_3| \leqq C\varepsilon_n, \qquad x\in[x_{-1}, x_{\lambda+2}],$$

where we have used the definitions of the constants γ_1 and γ_2, and the fact that $|\gamma_3|\leqq\|f_I-S_2\|$, which in turn is estimated by (5.11). Thus, we have shown (5.14) and so property (5.4) is verified.

We have one last task and that is to show that S_I is nondecreasing or, what amounts to the same thing, that $S_I'\geqq 0$. Now, $S_I'=S_3'-T_3$ and we already know that $S_3'\geqq 0$ and so we need only check where T_3 is not zero, namely, the intervals $[l_\nu n^{-1}, (l_\nu+r-1)n^{-1}]$, $\nu=0, 1, \cdots, \lambda$.

First, when $\nu=0$, we have

$$\begin{aligned} S_3'(x) &\geqq S_2'(x)\geqq f'(x)-C_2n\varepsilon_n \\ &\geqq (B_2-C_2)n\varepsilon_n, \qquad x\in[l_0n^{-1}, (l_0+r^{-1})n^{-1}], \end{aligned} \tag{5.15}$$

where in the second inequality we used (5.12) and in the last inequality, we used the fact that $[l_0n^{-1}, (l_o+r)n^{-1}]$ is the interval guaranteed by Lemma 3 to satisfy (4.10). We know that on $[l_0n^{-1}(l_0+r)n^{-1}]$, we have a contribution in T_3 from N_{l_0} but we may also have a contribution from some N_{l_ν}, $1\leqq\nu\leqq\lambda-1$. However, we have at most two such contributions, and so

$$\begin{aligned} |T_3(x)| &\leqq (4\gamma_1+r^2\gamma_2)\alpha_2 n \\ &\leqq (B_2-C_2)n\varepsilon_n \leqq S_3'(x), \qquad x\in[l_0n^{-1}, (l_0+r-1)n^{-1}], \end{aligned}$$

where the first inequality uses (2.2), the second inequality uses the values of the various constants, in particular that $C_2<A_1$, $r\geqq 3$, $\alpha_1<1$, and $\alpha_2, \alpha_3\geqq 1$, and the last inequality uses (5.15). This shows that $S_I'(x)=S_3'(x)-T_3(x)\geqq 0$ on $[l_0n^{-1}, (l_0+r-1)n^{-1}]$.

When $\nu=\lambda$, we argue as we did when $\nu=0$ to find that

$$\begin{aligned} |T_3(x)| &\leqq (|\gamma_3|+r^2\gamma_2)\alpha_2 n \leqq (B_2-C_2)n\varepsilon_n \\ &\leqq S_3'(x), \qquad x\in[l_\lambda n^{-1}, (l_\lambda+r-1)n^{-1}], \end{aligned}$$

because $|\gamma_3|\leqq\|f-S_2\|$, which in turn is estimated by (5.11).

When $\nu=1, \cdots, \lambda-2$, we argue in the same way. Now,

$$S_3'(x)\geqq (B_1-C_2)n\varepsilon_n, \qquad [x\in l_\nu n^{-1}, (l_\nu+r-1)n^{-1}],$$

because of (5.12) and Lemma 4. Also, we need only check T_3 on that part of $[l_\nu n^{-1}, (l_\nu+r-1)n^{-1}]$ which does not intersect either $[l_0n^{-1}, (l_0+r-1)n^{-1}]$ or

$[l_\lambda n^{-1}, (l_\lambda + r - 1)n^{-1}]$ and on this part we have

$$|T_3(x)| = r^2\gamma_2 N_{l_\nu}(x) \leqq r^2\gamma_2\alpha_2 n \leqq (B_1 - C_2)n\varepsilon_n$$
$$\leqq S_3'(x),$$

again because of the values of the various constants. This shows that $S_I'(x) \geqq 0$, for $x \in [l_\nu n^{-1}, (l_\nu + r - 1)n^{-1}]$, when $\nu = 1, \cdots, \lambda - 2$.

The proof that $S_I'(x) \geqq 0$, $x \in [l_{\lambda-1} n^{-1}, (l_{\lambda-1} + r - 1)n^{-1}]$ is exactly the same except that we use the additional fact that $s \leqq 3r^2$.

6. Remarks. We have mentioned in the Introduction that it is preferable to get the Jackson estimates of the form

$$E_n^*(f, r) \leqq C\omega_r(f, n^{-1}) \tag{6.1}$$

where ω_r is the rth order modulus of smoothness of f. The reason (6.1) is preferable is that then we would have the inverse theorem that if ω is an rth order modulus of smoothness and $E_n^*(f, r) = O(\omega(n^{-1}))$, $n \to \infty$, then $f \uparrow$ and $\omega_r(f, t) = O(\omega(t))$. This is because of inverse theorems for approximation by splines with equally spaced knots (see e.g. K. Scherer [9] for the case $\omega_r(t) = t^\theta$).

The reason that we do not get this result with our technique is that we approximate f' and then integrate to get our approximation to f. We could put our estimates in the form

$$E_n^*(f, r) \leqq Cn^{-1}\omega_{r-1}(f', n^{-1}) \tag{6.2}$$

when f' is continuous but even this does not reduce to (6.1) in this case.

In order to prove (6.1), it would be enough to show that whenever $f \uparrow$, and $\varepsilon > 0$, then there is a function $g_\varepsilon \uparrow$ with

$$\|f - g_\varepsilon\| \leqq C\omega_r(f, \varepsilon)$$

and

$$\|g_\varepsilon^{(r)}\| \leqq C\varepsilon^{-r}\omega_r(f, \varepsilon).$$

The key is that the functions g_ε be monotone nondecreasing, since otherwise the existence of the functions g_ε is already known as we have used in our proof. Once the existence of the functions g_ε are known then we would only need to use the fact that we can approximate g_ε with $\varepsilon = n^{-1}$, by a spline S in $\mathcal{S}^*(r, n)$ with error

$$\|g_\varepsilon - S\| \leqq Cn^{-r}\|g_\varepsilon^{(r)}\| \leqq C\omega_r(f, n^{-1})$$

because of our Theorem 1. This would give

$$\|f - S\| \leqq \|f - g_\varepsilon\| + \|g_\varepsilon - S\| \leqq C\omega_r(f, n^{-1}).$$

This in turn gives (6.1).

Acknowledgment. The author would like to thank Professor C. deBoor and Professor G. G. Lorentz, who both gave a careful reading of the original manuscript and suggested several improvements, especially in § 4.

REFERENCES

[1] C. DEBOOR AND G. J. FIX, *Spline approximation by quasi-interpolants*, J. Approximation Theory, 8 (1973), pp. 19–45.

[2] H. B. CURRY AND I. J. SCHOENBERG, *On Pólya frequency functions, IV. The fundamental splines and their limits*, J. Analyse Math., 17 (1966), pp. 71–107.

[3] R. DEVORE, *Approximation of Continuous Functions by Positive Linear Operators*, Springer Lecture Notes in Mathematics, no. 393, Springer–Verlag, Berlin, 1972.

[4] G. FREUD AND V. POPOV, *On approximation by spline functions*, Proc. of the Conference on the Constructive Theory of Functions, Akademia Kiado, Budapest, 1969, pp. 163–172.

[5] G. G. LORENTZ, *Monotone approximation*, Inequalities, III, Academic Press, New York, 1971, pp. 201–205.

[6] G. G. LORENTZ AND K. ZELLER, *Degree of approximation by monotone polynomials*, J. Approximation Theory, 1 (1968), pp. 501–504.

[7] M. MARSDEN, *An identity for spline functions with applications to variation diminishing spline approximation*, Ibid., 3 (1970), pp. 7–49.

[8] V. POPOV AND BL. K. SENDOV, *The classes that are characterized by the best approximations by spline functions*, Math. Notes, 8 (1970), pp. 550–557.

[9] K. SCHERER, *On the best approximation of continuous functions by splines*, SIAM J. Numer. Anal., 7 (1970), pp. 418–423.

[10] R. DE VORE, *Monotone approximation by polynomials*, this Journal, 8 (1977), pp. 906–921.

SIAM J. MATH. ANAL.
Vol. 8, No. 5, October 1977

MONOTONE APPROXIMATION BY POLYNOMIALS*

RONALD A. DE VORE†

Abstract. We prove Jackson type estimates for the approximation of monotone functions by monotone polynomials. The results are given in terms of the modulus of continuity of $f^{(k)}$, for any $k \geqq 0$. The estimates are of the same order as for the unconstrained approximation by polynomials.

1. Introduction. In the preceding paper [2], we have developed Jackson type theorems for monotone approximation by splines. Here, we want to give similar results for monotone approximation by algebraic polynomials.

Let Π_n denote the space of algebraic polynomials of degree $\leqq n$ and Π_n^*, the set of those polynomials in Π_n which are monotone nondecreasing on $[0, 1]$. If $f \in C[0, 1]$ is monotone nondecreasing on $[0, 1]$ $(f \uparrow)$, then we define the error of monotone approximation of f by polynomials of degree $\leqq n$ by

$$E_n^*(f) = \inf_{P \in \Pi_n^*} \|f - P\|,$$

with $\|\cdot\|$ the supremeum norm on $[0, 1]$.

Our main result is the following theorem which gives an estimate for $E_n^*(f)$ in terms of the smoothness of f.

THEOREM 1. *If $k \geqq 0$ and $f^{(k)} \in C[0, 1]$, $f \uparrow$, then for $n \geqq k + 1$*

$$E_n^*(f) \leqq C n^{-k} \omega(f^{(k)}, n^{-1}), \tag{1.1}$$

where C is a constant that depends only on k.

Thus (1.1) is the same as the classical Jackson theorem for unconstrained approximation by polynomials, and shows that at least in this sense there is no loss in the degree of approximation caused by the monotone constraint. We should remark that there are known examples with a loss in the degree of monotone approximation given by G. G. Lorentz and K. Zeller [6], and the author [1].

The cases $k = 0$, and $k = 1$ of Theorem 1 have been obtained previously by G. G. Lorentz and K. Zeller [5] and G. G. Lorentz [4], respectively. They have used linear methods of approximation in their cases. This is not possible in the general case since such a sequence of linear operators would have to preserve the positivity of f' and hence be restricted in their effectiveness of approximation by the saturation phenomena for positive operators.

The proof of Theorem 1 relies heavily on the results and techniques of [2]. In fact, the proof is developed in much the same way with the major exception being the fact that there is no direct analogue of B-splines. Instead, we have to construct polynomials which are large on a given interval and fall off fast outside of this interval. These polynomials are constructed in § 2.

---

* Received by the editors April 25, 1975, and in revised form April 14, 1976.

† Department of Mathematics, Oakland University, Rochester, Michigan. This work was supported by the National Science Foundation under Grant GP 19620.

Using the results of [2] on the degree of monotone approximation by splines, we can simplify the kind of functions for which we need to prove Theorem 1. In fact, it will be enough to show the following simpler case of Theorem 1.

THEOREM 2. *If $k \geqq 2$ and $f^{(k)}$ is absolutely continuous with $\|f^{(k+1)}\|_{L^\infty[0,1]} \leqq 1$, then*

$$E_n^*(f) \leqq Cn^{-k-1}, \qquad n \geqq N, \tag{1.2}$$

where C and N are constants that depend only on k.

Let us observe why it is enough to prove Theorem 2. Assuming that we have proved Theorem 2 and f is an arbitrary function in $C^{(k)}[0, 1]$, then by Theorem 1 of [2], there is a spline $S \in \mathcal{S}(k+2, n)$ (following the notation of [2]), with $S \uparrow$ and

$$\|f - S\| \leqq C' n^{-k} \omega(f^{(k)}, n^{-1}), \tag{1.3}$$

where C' depends only on k. From (1.3), it follows that when $|t| \leqq n^{-1}$

$$|\Delta_t^{k+1}(S, x)| \leqq |\Delta_t^{k+1}(f, x)| + 2^{k+1} C' n^{-k} \omega(f^{(k)}, n^{-1}) \leqq C'' n^{-k} \omega(f^{(k)}, n^{-1}). \tag{1.4}$$

Now, S is a spline of degree $k+1$, and so $S^{(k+1)}$ is a step function. If $\|S^{(k+1)}\| = \gamma$, then there is an interval $[a, b]$ of length n^{-1} on which $S^{(k+1)}$ either equals γ or $-\gamma$. If it equals γ and $t = (k+1)^{-1} n^{-1}$, then

$$\Delta_t^{k+1}(S, a) = \int_a^{a+t} \int_{x_k}^{x_k+t} \cdots \int_{x_1}^{x_1+t} S^{(k+1)}(x)\, dx\, dx_1 \cdots dx_k \geqq \gamma t^{k+1}. \tag{1.5}$$

This together with (1.4) shows that

$$\|S^{(k+1)}\| = |\gamma| \leqq Dn\omega(f^{(k)}, n^{-1}),$$

with D a constant depending only on k. This same result holds for $-\gamma$.

Now, by Theorem 2, there is a polynomial $P \in \Pi_n^*$, such that

$$\|S - P\| \leqq CDn^{-k} \omega(f^{(k)}, n^{-1})$$

and thus

$$\|f - P\| \leqq \|f - S\| + \|S - P\| \leqq (C' + CD) n^{-k} \omega(f^{(k)}, n^{-1}),$$

which gives Theorem 1 for those values of $n \geqq N$.

We need also to check the values of n with $k+1 \leqq n < N$. For this, let Q be a polynomial of degree $\leqq n$ with

$$\|f' - Q\| \leqq C(n-1)^{-k+1} \omega(f^{(k)}, (n-1)^{-1}) \leqq C' n^{-k} \omega(f^{(k)}, n^{-1}).$$

The existence of such polynomials Q follows from the usual unconstrained Jackson theorems. The polynomial

$$P(x) = f(0) + \int_0^x (Q(t) + C' n^{-k} \omega(f^{(k)}, n^{-1}))\, dt$$

is monotone nondecreasing and we have

$$\|f - Q\| \leqq \|f - Q\| + C' n^{-k} \omega(f^{(k)}, n^{-1}).$$

Thus, we have (1.2) also for the values of n with $k+1 \leqq n < N$. This then shows that Theorem 2 implies Theorem 1. Thus in the sequel, it will be enough to prove Theorem 2 and so we can restrict ourselves to functions f with $\|f^{(k+1)}\|_\infty \leqq 1$.

There were certain constants that played an important role in the statement and results of [2]. In this paper, we will sometimes have to redefine these constants to fit the needs of polynomial approximation. We will use the same symbolization for these constants. This is allowable because the constants that are redefined all had only the restriction on them that certain inequalities hold. In the case that the constants are $\geqq 1$, these inequalities will still hold if we redefine the constants to be larger, which we do. When the constants are $\leqq 1$, these inequalities will still hold if we redefine the constants to be smaller, which we do.

2. Some special polynomials. We need to construct some polynomials which mimic the B-splines. These polynomials will be used for corrections in the same way that the B-splines were used in [2]. Let $T_m(x) = \cos m(\arccos x)$ be the Chebyshev polynomials of degree m. If m is odd, then $T_m(0) = 0$. Near the end of this section, we will prescribe an even positive integer r, which will depend only on k and will be larger than $4k+2$. Thus any constants that appear and depend only on r will in turn depend only on k. Let m be the largest odd integer such that $mr \leqq n$, and define

$$Q_{m,r}(x) = \int_{-1}^{x} c_{m,r}(m^{-1}t^{-1}T_m(t))^r \, dt,$$

with $c_{m,r}$ a normalizing constant chosen so that $Q_{m,r}(1) = 1$.

We want first to estimate $c_{m,r}$. If $|t| \leqq m^{-1}$, then $|m^{-1}t^{-1}T_m(t)| \geqq 2\pi^{-1} \geqq 2^{-1}$, and so

$$\int_{-1}^{1} (m^{-1}t^{-1}T_m(t))^r \, dt \geqq 2^{-r+1}m^{-1}.$$

Also, since $|T_m(t)| \leqq 1$, $-1 \leqq t \leqq 1$, we have

$$\int_{|t|\geqq m^{-1}} (m^{-1}t^{-1}T_m(t))^r \, dt \leqq 2m^{-1}\int_{1}^{m} t^{-r} \, dt \leqq 2m^{-1},$$

since $r \geqq 4$. Now, $|m^{-1}t^{-1}T_m(t)| \leqq 1$, $-1 \leqq t \leqq 1$, and so

$$\int_{|t|\leqq m^{-1}} (m^{-1}t^{-1}T_m(t))^r \, dt \leqq 2m^{-1}.$$

These last three inequalities show that

$$\frac{m}{4} \leqq c_{m,r} \leqq 2^{r-1}m. \tag{2.1}$$

This last inequality for $c_{m,r}$, together with the fact that $|T_m(t)| \leqq 1$, $-1 \leqq t \leqq 1$, shows that

$$|Q_{m,r}(x)| \leqq \left|\frac{mx}{2}\right|^{-r+1}, \qquad -1 \leqq x \leqq 0, \tag{2.2}$$

$$|1-Q_{m,r}(x)| \leqq \left|\frac{mx}{2}\right|^{-r+1}, \qquad 0 \leqq x \leqq 1. \tag{2.3}$$

If $I=[a, b]$ is an interval contained in $[0, 1]$ with length $\geqq rn^{-1}$, then denote by $|I|$, the length of I and define

$$\lambda_I(x) = c_I(Q_{m,r}(x-a')-Q_{m,r}(x-b'))+n^{-r}|I|^{-1},$$

with $a'=a-8rn^{-1}$, $b'=b+8rn^{-1}$, and c_I a normalizing constant chosen so that

$$\int_0^1 \lambda_I(t)\,dt = 1.$$

λ_I is then a polynomial of degree $\leqq n-1$, which as we shall see is large on I and falls off fast outside of I. First, we want to show that $c_I \sim |I|^{-1}$.

When $n \geqq 3r$, then $m \geqq 2$ and $8rn^{-1} \geqq 4m^{-1}$. We will only consider values of n larger than $3r$ throughout but this is permissible since r depends only on k. If we use (2.2) and (2.3), we find

$$Q_{m,r}(x-a')-Q_{m,r}(x-b') \geqq 1-2^{-r+1}-2^{-r+1} \geqq \tfrac{1}{2}, \qquad x \in I.$$

Also, since $|Q_{m,r}(x)| \leqq 1$, $-1 \leqq x \leqq 1$, we have

$$\tfrac{1}{2} \leqq Q_{m,r}(x-a')-Q_{m,r}(x-b') \leqq 2, \qquad x \in I. \tag{2.4}$$

We can also estimate outside of I. If $x \in [0, 1]$ and $\delta' = \operatorname{dist}(x, [a', b']) \geqq n^{-1}$, then $\operatorname{dist}(x, I) \leqq 9r\delta'$. Thus,

$$\begin{aligned} &|Q_{m,r}(x-a')-Q_{m,r}(x-b')| \\ &\leqq \left(\frac{m\delta'}{2}\right)^{-r+1} \leqq \left(\frac{m \operatorname{dist}(x, I)}{18r}\right)^{-r+1}, \qquad \text{when } \operatorname{dist}(x, I) \geqq 9rn^{-1}, \end{aligned} \tag{2.5}$$

because of (2.2) and (2.3). Since $|I| \geqq rn^{-1}$, we have

$$\frac{|I|}{2} \leqq \int_0^1 (Q_{m,r}(x-a')-Q_{m,r}(x-b'))\,dx \leqq (38+9(4r)^r)|I| \tag{2.6}$$

where the left side was estimated using (2.4) and the right side was estimated by considering the integral over two sets. The first set is where $\operatorname{dist}(x, I) \leqq 9rn^{-1}$, on which we used the facts that $|Q_{m,r}(x)| \leqq 1$, $x \in [-1, 1]$, and that this set has measure $\leqq 19|I|$. The integral over the second set S was estimated by using (2.5), to see that

$$\begin{aligned} \int_S (Q_{m,r}(x-a')-Q_{m,r}(x-b'))\,dx &\leqq 2\left(\frac{18r}{m}\right)^{r-1} \int_{9rn^{-1}}^1 t^{-r+1}\,dt \\ &\leqq 9(4r)^{r-1}m^{-1} \leqq 9(4r)^r|I|, \end{aligned}$$

where we have used the fact that $|I| \geqq rn^{-1}$ and $n \leqq 2rm$.

The estimate (2.6) shows that there are constants a_1, $a_2 > 0$, such that

$$a_1|I|^{-1} \leqq c_I \leqq a_2|I|^{-1}; \tag{2.7}$$

so $c_I \sim |I|^{-1}$ as we have previously stated.

If J is any interval we define

$$d_n(x, J) = \max(1, n \operatorname{dist}(x, J)).$$

Now, because of (2.5) and (2.7), it follows that if we choose α_2 of equation (2.2) of [2] large enough, then

$$\lambda_I(x) \leqq \alpha_2 |I|^{-1} (d_n(x, I))^{-r+1}, \quad \text{for } x \in [0, 1]. \tag{2.8}$$

Note that this inequality automatically holds when $\operatorname{dist}(x, I) \leqq 9rn^{-1}$, because $|Q_{m,r}(x)| \leqq 1$, $x \in [-1, 1]$.

We will also need an estimate for $\lambda_I(x)$ from below. Namely the constant α_1 of equation (2.1) of [2] can be chosen small enough that

$$\alpha_1 |I|^{-1} (d_n(x, I))^{-r} \leqq \lambda_I(x), \qquad x \in [0, 1]. \tag{2.9}$$

This estimate holds for $x \in I$, because of (2.4). In the same way that we have proved (2.4), we can show that (2.9) holds when $\operatorname{dist}(x, I) \leqq rn^{-1}$. Also, (2.9) automatically holds when $\operatorname{dist}(x, I) \geqq \frac{1}{4}$, because of the term $|I|^{-1} n^{-r}$ that appears in the definition of λ_I.

To see that (2.9) holds when $rn^{-1} \leqq \operatorname{dist}(x, I) \leqq \frac{1}{4}$, let $\xi'_\nu = \cos((3\nu + 1)/3)\pi m^{-1}$ and $\xi''_\nu = \cos((3\nu - 1)/3)\pi m^{-1}$. Then, $|T_m(x)| \geqq \frac{1}{2}$, for $x \in [\xi'_\nu, \xi''_\nu]$, and $|\xi'_\nu - \xi''_\nu| \geqq m^{-1}$, whenever $[\xi'_\nu, \xi''_\nu] \subseteq [-\frac{1}{2}, \frac{1}{2}]$. Suppose that $x < a$ and $\delta = \operatorname{dist}(x, I) = |x - a|$. Then the interval $[\delta, \delta + 8rn^{-1}]$ contains a set S, which consists of the parts of the intervals $[\xi'_\nu, \xi''_\nu]$ which intersect $[\delta, \delta + 8rn^{-1}]$ and the set S has measure $|S| \geqq m^{-1}$. Therefore,

$$\begin{aligned}
\int_{x-b'}^{x-a'} c_{m,r}(m^{-1}t^{-1}T_m(t))^r \, dt &\geqq \int_{\delta}^{\delta + 8rn^{-1}} c_{m,r}(m^{-1}t^{-1}T_m(t))^r \, dt \\
&\geqq \int_S c_{m,r}(2mt)^{-r} \, dt \geqq (2m)^{-r}(\delta + 8rn^{-1})^{-r}|S| c_{m,r} \\
&\geqq C(n\delta)^{-r}
\end{aligned}$$

where C is a constant that depends only on r. Here, we have used our estimate for $c_{m,r}$ in (2.1), the definition of m, and the fact that $\delta \geqq rn^{-1}$.

This last estimate and our estimate for c_I in (2.7) show that when $x < a$ and $\operatorname{dist}(x, I) \geqq rn^{-1}$

$$\lambda_I(x) \geqq c_I \int_{x-b'}^{x-a'} c_{m,r}(m^{-1}t^{-1}T_m(t))^r \, dt \geqq \alpha_1 |I|^{-1} (d_n(x, I))^{-r}.$$

The same estimate holds when $x > b$ and $\operatorname{dist}(x, I) \geqq rn^{-1}$, and so we have proved (2.9).

The polynomials λ_I will be used to correct the derivatives of our approximating polynomials. The primitive

$$\Lambda_I(x) = \int_0^x \lambda_I(t) \, dt$$

will therefore be the correction to the approximating polynomials themselves.

Because of our normalization, $\Lambda_I(1)=1$. It follows from (2.8) that

$$|\Lambda_I(x)|\leqq\alpha_2|I|^{-1}\int_0^x (d_n(t,I))^{-r+1}\,dt\leqq 2\alpha_2|I|^{-1}n^{-1}(d_n(x,I))^{-r+2},\qquad x\leqq a. \tag{2.10}$$

Similarly,

$$|1-\Lambda_I(x)|\leqq 2\alpha_2|I|^{-1}n^{-1}(d_n(x,I))^{-r+2},\qquad x\geqq b. \tag{2.11}$$

On I, we have $|\lambda_I(t)|\leqq\alpha_2|I|^{-1}$, and so

$$|\Lambda_I(x)-\Lambda_I(y)|\leqq\left|\int_x^y \lambda_I(t)\,dt\right|\leqq\alpha_2|I|^{-1}|x-y|\quad\text{when } x,\,y\in I. \tag{2.12}$$

We will need one other correcting polynomial. Let k be the integer in Theorem 1, $k\geqq 2$, and now let m be chosen so that it is the largest odd integer with $m(2k+2)\leqq n-2$. Consider,

$$\phi_{m,r}(x)=(r^2(2n)^{-2}-x^2)(m^{-1}x^{-1}T_m(x))^{2k+4}$$

which is a polynomial of degree $\leqq n-1$ that is positive on $[-r(2n)^{-1}, r(2n)^{-1}]$, and negative outside of this interval. We first want an estimate for the integral of $\phi_{m,r}$ over $[-1, 1]$.

When $|x|\leqq kn^{-1}\leqq m^{-1}$, then $|m^{-1}x^{-1}T_m(x)|\geqq\frac{1}{2}$. Also, since $r\geqq 4k$, we have $(r^2(2n)^{-2}-x^2)\geqq 8^{-1}r^2n^{-2}$, when $|x|\leqq kn^{-1}$. Hence,

$$\int_{-r(2n)^{-1}}^{r(2n)^{-1}}\phi_{m,r}(x)\,dx\geqq 8^{-1}r^2n^{-2}2^{-2k-4}(2kn^{-1})=2^{-2k-6}r^2kn^{-3}. \tag{2.13}$$

We can also estimate the integral over the set $S=[-1,1]-[-r(2n)^{-1}, r(2n)^{-1}]$. Now, $|T_m(x)|\leqq 1$, $x\in[-1,1]$, and $\phi_{m,r}$ is negative on S, and so

$$\left|\int_S \phi_{m,r}(x)\,dx\right|\leqq 2\int_{r(2n)^{-1}}^1 x^2(mx)^{-2k-4}\,dx$$

$$\leqq 2m^{-2k-4}(2nr^{-1})^{2k+1}\leqq 2\left(\frac{2n}{mr}\right)^{2k+1}\left(\frac{n}{m}\right)^3 n^{-3}.$$

Since k is fixed and $n\leqq 4mk$, we can choose r sufficiently large, depending only on k, so that $r\geqq 4k+2$, and

$$\left|\int_S \phi_{m,r}(x)\,dx\right|\leqq\frac{1}{2}\int_{-r(2n)^{-1}}^{r(2n)^{-1}}\phi_{m,r}(x)\,dx. \tag{2.14}$$

We fix this value of r for the rest of the paper.

If I is an interval contained in $[0, 1]$ of length rn^{-1} and midpoint a, then define

$$\phi_I(x)=d_I\phi_{m,r}(x-a),\qquad \int_0^1\phi_I(x)\,dx=1.$$

Because of (2.13) and (2.14), it follows that there are constants a_1', $a_2'>0$, such that

$$a_1'n^3\leqq d_I\leqq a_2'n^3.$$

We can also require that α_2 is chosen large enough that when $\operatorname{dist}(x, I) \geqq 2rn^{-1}$, then

$$(2.15)\quad |\phi_I(x)| \leqq C|r^r(2n)^{-2} - (x-a)^2|(m \operatorname{dist}(x, I))^{-2k-4} n^3 \leqq \alpha_2 n(d_n(x, I))^{-2k-2},$$

where we have used the facts that $d_n(x, I) \geqq 1$, for all x and $|r^2(2n)^{-2} - (x-a)^2| \leqq 2(\operatorname{dist}(x, I))^2$, when $\operatorname{dist}(x, I) \geqq 2rn^{-1}$. If α_2 is large enough then this inequality will also hold when $\operatorname{dist}(x, I) \leqq 2rn^{-1}$, because in this case $|\phi_I(x)| \leqq Cn$, with C depending only on k. Therefore,

$$(2.16)\quad |\phi_I(x)| \leqq \alpha_2 n(d_n(x, I))^{-2k-2}, \qquad x \in [0, 1].$$

Define the primitive

$$\Phi_I(x) = \int_0^x \phi_I(t)\, dt,$$

which is a polynomial of degree $\leqq n$. From (2.16), it follows that α_2 can be chosen so large that

$$(2.17)\quad |\Phi_I(x)| \leqq \alpha_2 (d_n(x, I))^{-2k-1}, \qquad x \leqq a,$$

$$(2.18)\quad |1 - \Phi_I(x)| \leqq \alpha_2 (d_n(x, I))^{-2k-1}, \qquad x \geqq a,$$

where as before, a is the midpoint of I.

3. A decomposition of f. In [2], we have given a decomposition for monotone functions which we will also use here. We decompose $[0, 1]$ into a union of certain pairwise disjoint intervals I_j^*, $j = 1, 2, \cdots, m$ and J_j^*, $j = 0, 1, \cdots, m$, with I_j^* to the right of J_{j-1}^* and to the left of J_j^*. For any of these intervals I, we define

$$f_I(x) = \int_0^x f'(t)\chi_I(t)\, dt,$$

where χ_I is the characteristic function of the interval I. Then our decomposition for f is

$$(3.1)\quad f(x) = f(0) + \sum_{j=0}^{m} f_{J_j^*} + \sum_{j=1}^{m} f_{I_j^*}.$$

Each of the functions f_I is monotone nondecreasing and our proof of Theorem 2 will be to approximate each f_I by monotone polynomials to get our approximation to f.

The intervals J_j^* and I_j^* have special properties that we summarize. Recall that to prove Theorem 2, we need only consider functions with $\|f^{(k+1)}\|_{L^\infty[0,1]} = 1$. Each interval J_j^* has length $\geqq r^2 n^{-1}$ and

$$(3.2)\quad \textit{if } I \textit{ is any of the intervals } J_j^*, \textit{ then } f_I'(x) = f'(x) \leqq An^{-k},\ x \in I, \textit{ with } A \textit{ a constant depending only on } k.$$

On the other hand, each interval I_j^* has length $\geqq 2r^2 n^{-1}$ and

$$(3.3)\quad \textit{if } I \textit{ is any of the intervals } I_j^*, \textit{ then there is an interval } [l_0 n^{-1}, (l_0 + r)n^{-1}], \textit{ contained in } I \textit{ on which } f_I'(x) = f'(x) \geqq B_2 n^{-k}.$$

Also, if we denote $I=[i_1n^{-1}, i_2n^{-1}]$, $i_2-i_1=\lambda r^2+\mu$, with $\lambda>1$ and $0\leqq\mu<r^2$, and $x_\nu=(i_1+\nu r^2)n^{-1}$, then

(3.4) *if I is any of the intervals I_j^*, then for each $1\leqq\nu\leqq\lambda$, there is an interval $[l_\nu n^{-1}, (l_\nu+r)n^{-1}]$ contained in $[x_{\nu-1}, x_\nu]$ on which $f_I'(x)=f'(x)\geqq B_1n^{-k}$.*

We should remark that the actual statements of the results (3.3) and (3.4) in [2] (these are Lemmas 3 and 4 in [2]) are stated with $n^{-k+1}\omega(f^{(k)}, n^{-1})$ in place of n^{-k}. However, the proof goes over exactly the same with n^{-k}.

The constant B_1 of (3.14) is equal to $2^{-r^4}A_1=100r^2C_2\alpha_1^{-1}\alpha_2^2$, where $A_1=100r^22^{r^4}C_2\alpha_1^{-1}\alpha_2^2$. As we have remarked earlier, the constants α_1, α_2, C_2 were introduced in [2] so that certain inequalities hold. We have redefined α_1 and α_2 in § 2, preserving the original inequalities, and requiring that some new inequalities hold. In a similar vein, we will redefine C_2 in § 5, so that a new inequality holds while retaining the old inequalities that involved C_2. The constant $B_1=2^{-r^4}\alpha_3A_1^2$ where we will use the same value of α_3 as in [2]. The only importance in the value of α_3 for this paper is that it is bigger than 1.

If I is one of the intervals I_j^*, then we have a control over f' immediately to the right and left of I, because of (3.2). This estimate was actually given in a more precise way in [2] and we will need this more precise version:

(3.5) *if $I=[a, b]$ is one of the intervals I_j^*, then $f'(x)\leqq A_1n^{-k}$, $x\in([a-rn^{-1}, a]\cup[b, b+rn^{-1}])\cap[0, 1]$.*

The intervals $[a-rn^{-1}, a]\cap[0, 1]$ and $[b, b+rn^{-1}]\cap[0, 1]$ are contained in what we called intervals of type 1 in [2].

4. Approximation of the functions $f_{I_j^*}$. We can approximate the functions $f_{I_j^*}$ using the technique of G. G. Lorentz and K. Zeller [5]. Let I be one of the intervals J_j^*. The function f_I is in Lip 1, in fact

$$|f'(x)|\leqq An^{-k}, \qquad \text{a.e. in } [0, 1], \tag{4.1}$$

because of (3.2). Thus, the Lorentz–Zeller theorem gives that there is a polynomial $P_I\in\Pi_n^*$, such that

$$|f_I(x)-P_I(x)|\leqq Cn^{-k-1}, \qquad x\in[0, 1], \tag{4.2}$$

with C an absolute constant. We want to observe more, namely that P_I is a better approximation away from the interval I, due to the fact that $f_I'(x)=0$, outside of I.

LEMMA 1. *If I is one of the intervals J_j^*, $j=0, 1, \cdots, m$, then there is a polynomial $P_I\in\Pi_n^*$, such that*

$$|f_I(x)-P_I(x)|\leqq CAn^{-k-1}(d_n(x, I))^{-k}, \qquad x\in[0, 1], \tag{4.3}$$

with C depending only on k.

Proof. The basic idea is to go to the trigonometric case via the substitution $x=\cos\theta$, and then use the Jackson operators. Let r be the integer defined in § 2, and K_n the Jackson kernel

$$K_n(t)=c_n\left(\frac{\sin(n't/2)}{\sin(t/2)}\right)^{2r+2}, \qquad \int_{-\pi}^{\pi}K_n(t)\,dt=1, \tag{4.4}$$

where n' is chosen as the largest integer such that $(2r+2)n' \leqq n$. So, K_n is a trigonometric polynomial of degree $\leqq n$, and we have the following estimates for the moments of K_n (see G. G. Lorentz [3, p. 57]):

$$\int_{-\pi}^{\pi} |t|^j K_n(t)\, dt \leqq C_1 n^{-j}, \qquad j = 0, 1, \cdots, 2r, \tag{4.5}$$

with C_1 a constant that depends only on k.

If h is a 2π periodic function, we define

$$L_n(h, \theta) = \int_{-\pi}^{\pi} h(\theta + t) K_n(t)\, dt.$$

It will be notationally more convenient to work on $[-\frac{1}{2}, \frac{1}{2}]$, then on $[0, 1]$ and so we introduce $\tilde{f}_I(x) = f_I(x + \frac{1}{2})$. Let $g(\theta) = \tilde{f}_I(\cos\theta)$, and define

$$g_n(\theta) = \begin{cases} g(k\pi(n')^{-1}), & \nu\pi(n')^{-1} \leqq \theta \leqq (\nu+1)\pi(n')^{-1}, \quad \nu = 0, \cdots, n'-1, \\ g_n(-\theta), & \text{for } \theta < 0. \end{cases}$$

Since the function g_n is even, we have that $L_n(g_n, \theta)$ is an even trigonometric polynomial of degree $\leqq n$. Hence, $\tilde{P}_I(x) = L_n(g_n, \arccos x)$ is an algebraic polynomial of degree $\leqq n$. Lorentz and Zeller have only used the operators L_n when $r = 1$, but the proof of the monotonicity of $\tilde{P}_I$ and the verification of (4.2) with $P_I(x) = \tilde{P}_I(x - \frac{1}{2})$ is exactly the same in the general case.

We need to get a better estimate than (4.2) outside of I. If S is any set, let $\tilde{S} = \{x : x + \frac{1}{2} \in S\}$ and $\tilde{S}' = \{\theta : \cos\theta \in S\}$. Note first that $|g_n(\theta) - g(\theta)| \leqq CA\pi n^{-k-1}$ for all θ and $g_n(\theta) = g(\theta)$ if $\operatorname{dist}(\theta, \tilde{I}') \geqq \pi n^{-1}$. Hence, if we take $\theta \notin \tilde{I}'$ and let $\delta = \operatorname{dist}(\theta, \tilde{I}')$ and assume that $\delta \geqq \pi n^{-1}$, then we have

$$|g_n(\theta + t) - g(\theta)| \leqq |g_n(\theta + t) - g(\theta + t)| + |g(\theta + t) - g(\theta)|$$

and so $|g_n(\theta + t) - g(\theta)| = 0$, $|t| \leqq \delta$ and $\leqq CAn^{-k}|t|$ when $|t| > \delta$. This gives

$$\left| \int_{-\pi}^{\pi} (g_n(\theta + t) - g(\theta)) K_n(t)\, dt \right| \leqq CAn^{-k} \int_{|t| \geqq \delta} |t| K_n(t)\, dt$$

$$\leqq CAn^{-k}\delta^{-r+1} \int_{-\pi}^{\pi} |t|^r K_n(t)\, dt \leqq CAn^{-r-k}\delta^{-r+1},$$

because of (4.5). Translating this to $x \notin \tilde{I}$, using $\operatorname{dist}(\theta, \tilde{I}') \geqq \operatorname{dist}(x, \tilde{I})$, we find

$$|\tilde{f}_I(x) - \tilde{P}_I(x)| \leqq CAn^{-k-1}(d_n(x, \tilde{I}))^{-r+1}, \qquad x \in [-1, 1],$$

with C a constant depending only on k, where we have also used (4.2). When we restate this last inequality in terms of f_I and P_I and make the simple observations that $d_n(x, I) \geqq 1$ and $r \geqq k + 1$, we get (4.3).

5. Approximation of the functions $f_{I_j}^*$. The approximation of the functions $f_{I_j}^*$ is more complicated. In this section, we will use standard techniques to get a good polynomial approximation to $f_{I_j}^*$ but this polynomial may not be monotone and so we will have to make corrections to this polynomial in the next section. Again, it is more convenient to work on $[-\frac{1}{2}, \frac{1}{2}]$ than on $[0, 1]$. Let $I = [a, b]$ be one of the

intervals I_j^*, and denote as before $\tilde{I}=\{x: x+\frac{1}{2}\in I\}$, $\tilde{f}_I(x)=f_I(x+\frac{1}{2})$, and $g(\theta)=\tilde{f}_I(\cos\theta)$.

We will approximate first with the Jackson operators of order r. Let

$$M_n(g,\theta)=-\int_{-\pi}^{\pi}\left(\sum_1^r(-1)^\nu\binom{r}{\nu}g(\theta+\nu t)\right)K_n(t)\,dt,$$

where K_n is the kernel of (4.4). Then $M_n(g,\theta)$ is an even trigonometric polynomial of degree $\leqq n$, and so $\tilde{P}(x)=M_n(g,\arccos x)$ is an algebraic polynomial of degree $\leqq n$. The polynomial $P(x)=\tilde{P}(x-\frac{1}{2})$ is a good approximation to f_I.

Let $E=E_1\cup E_2$, where $E_1=[a-rn^{-1},\ a+rn^{-1}]\cap[0,1]$ and $E_2=[b-rn^{-1},\ b+rn^{-1}]\cap[0,1]$. The following lemma establishes the approximating properties of P and in the process redefines the constant C_2 of [2].

LEMMA 2. *The constant C_2 can be chosen so large that*

$$|f_I(x)-P(x)|\leqq C_2A_1n^{-k-1}(d_n(x,I))^{-r},\qquad x\in[0,1], \tag{5.1}$$

$$|f_I'(x)-P'(x)|\leqq C_2(A_1(d_n(x,E))^{-r}+(d_n(x,I))^{-r})n^{-k},\qquad x\in[0,1]. \tag{5.2}$$

Proof. It will be important to observe that our choice of C_2 does not depend on any of the other constants, particularly A_1 and A_2. Throughout the proof C and C' denote constants that depend on k but are independent of all of the other constants.

If $I=[0,1]$, then (5.1) and (5.2) follow from the usual Jackson theorems on the simultaneous approximation of a function and its derivatives. Hence, we can assume that $I\neq[0,1]$. This will allow us to control the derivatives of f provided n is sufficiently large. Indeed, we know that $|f'(x)|\leqq A_1n^{-k}$ on an interval $\mathcal{I}$ of length $\geqq n^{-1}$ because of (3.5). This and the fact that $|f^{(k+1)}(x)|\leqq 1$ on $\mathcal{I}$ give that $|f^{(i)}(x)|\leqq CA_1n^{-k+i-1}$, $x\in\mathcal{I}$, $i=1,2,\cdots,k+1$, with C a constant depending only on k.

Now that we have $f^{(i)}$ controlled on $\mathcal{I}$, it is easy to get an estimate for all x. For example, if $x_0\in\mathcal{I}$, we have

$$|f^{(k)}(x)|\leqq|f^{(k)}(x_0)|+\left|\int_{x_0}^{x}f^{(k+1)}(t)\,dt\right|\leqq(CA_1n^{-1}+1)$$

where we used the fact that $\|f^{(k+1)}\|_\infty=1$. Continuing in this way, we get

$$\|f^{(i)}\|\leqq(C'A_1n^{-1}+1),\qquad i=1,2,\cdots,k+1,$$

with C' depending only on k. Thus for n sufficiently large $C'A_1n^{-1}\leqq 1$ and we have $\|f^{(i)}\|\leqq 2$, $i=1,2,\cdots,k+1$.

We now proceed to the actual proof of (5.1) and (5.3). We need only verify these inequalities for n sufficiently large since they hold automatically for $n\leqq N$. The proof consists of showing the corresponding result for the approximation of g by $M_n(g)$. To do this, we first need some estimates for $\Delta_t^r(g,\theta)$. If S is any subset of $[0,1]$, then as in § 4, we denote $\tilde{S}=\{x: x+\frac{1}{2}\in S\}$, and $\tilde{S}'=\{\theta:\cos\theta\in\tilde{S}\}$.

Since f_I is constant outside of I, we have

$$\Delta_t^r(g,\theta)=0,\quad\text{when } [\theta,\theta+rt]\cap\tilde{I}'=\phi. \tag{5.3}$$

Also, since $f_I(x)=f(x)-f(a)$ on I, we have

$$|\Delta_t^r(g,\theta)|\leqq 2^r|\Delta_t^{k+1}(g,\theta)|\leqq C\|f^{(k+1)}\|_\infty t^{k+1}\leqq Ct^{k+1}, \quad \text{when } [\theta,\theta+rt]\subseteq \tilde{I}' \tag{5.4}$$

with C a constant depending only on k. In the second inequality, we used the fact that the $(k+1)$st derivative of g can be expressed in terms of the $f^{(i)}$, $i=1,\cdots,k+1$ and we know $\|f^{(i)}\|\leqq 2$, $i=1,2,\cdots,k+1$.

We also need an estimate for $\Delta_t^r(g,\theta)$ in the remaining case. Let $F=([a-rn^{-1},a]\cup[b,b+rn^{-1}])\cap[0,1]$. Because of (3.5), we know that

$$|f_I(x)-(f(x)-f(a))|\leqq\int_F f'(t)\,dt\leqq 2rA_1n^{-k-1}, \qquad x\in F, \tag{5.5}$$

where we have used the fact that the two intervals that make up F each have length rn^{-1}, and $f_I'(x)=0$, $x\in F$.

Let $\tilde{f}(x)=f(x+\frac{1}{2})$ and $h(\theta)=\tilde{f}(\cos\theta)$. It follows from (5.5) that

$$|g(\theta)-(h(\theta)-f(a))|\leqq 2rA_1n^{-k-1}, \qquad \theta\in\tilde{F}'. \tag{5.6}$$

Now, suppose that $[\theta,\theta+rt]$ is not contained in either $\tilde{I}'$, or $\mathscr{C}\tilde{I}'$, and $|t|\leqq n^{-1}$. Then, $\theta\in\tilde{E}'$ and for each value of ν either $\theta+\nu t\in\tilde{F}'$ or $\theta+\nu t\in\tilde{I}'$, and so

$$\begin{aligned}|\Delta_t^r(g,\theta)|&\leqq\left|\sum_0^r(-1)^\nu\binom{r}{\nu}h(\theta+\nu t)\right|+\left|\sum{}'(-1)^\nu\binom{r}{\nu}(g(\theta+\nu t)-h(\theta+\nu t)+f(a))\right|\\ &\leqq|\Delta_t^r(h,\theta)|+2^{r+1}rA_1n^{-k-1}\leqq CA_1n^{-k-1}, \qquad \theta\in\tilde{E}',\quad |t|\leqq n^{-1},\end{aligned} \tag{5.7}$$

with C a constant depending only on k and the set E introduced at the beginning of this section. The $\sum'$ indicates that this sum is taken only over those ν for which $\theta+\nu t$ is in F'. This sum was estimated using (5.6).

The estimate (5.7) holds for all other values of θ, if $|t|\leqq n^{-1}$, because in the other cases either (5.3) or (5.4) holds. Hence,

$$\omega_r(g,n^{-1})\leqq CA_1n^{-k-1},$$

again with C a constant depending only on k. Hence for any t and θ,

$$|\Delta_t^r(g,\theta)|\leqq\omega_r(g,t)\leqq(1+nt)^r\omega_r(g,n^{-1})\leqq CA_1n^{-k-1}(1+nt)^r. \tag{5.8}$$

Now, the estimates (5.3) and (5.8), can be used to prove (5.1). First for any θ, we have

$$\begin{aligned}|g(\theta)-M_n(g,\theta)|&\leqq\int_{-\pi}^{\pi}|\Delta_t^r(g,\theta)|K_n(t)\,dt\\ &\leqq CA_1n^{-k-1}\int_{-\pi}^{\pi}(1+nt)^rK_n(t)\,dt\leqq C'A_1n^{-k-1},\end{aligned} \tag{5.9}$$

with C' a constant depending only on k. Here, we have used (5.8), the fact that K_n has integral 1 and the estimates for the moments of K_n given in (4.5).

We need to improve this estimate when θ is not in $\tilde{I}'$. For such θ, let $\delta = \operatorname{dist}(\theta, \tilde{I}')$. If $\delta > 0$, then

$$
\begin{aligned}
|g(\theta) - M_n(g,\theta)| &\leqq \int_{|t|>\delta r^{-1}} |\Delta_t^r(g,\theta)| K_n(t)\,dt \leqq r^{r-\delta^{-r}} \int_{-\pi}^{\pi} t^r |\Delta_t^r(g,\theta)| K_n(t)\,dt \\
&\leqq r^r C A_1 \delta^{-r} n^{-k-1} \int_{-\pi}^{\pi} t^r (1+nt)^r K_n(t)\,dt \\
&\leqq C' A_1 (n\delta)^{-r} n^{-k-1},
\end{aligned} \tag{5.10}
$$

with C' a constant depending only on k. Here, we have used (5.3), (5.8), and (4.5).

The last inequality coupled with (5.9) shows that

$$
|g(\theta) - M_n(g,\theta)| \leqq C' A_1 n^{-k-1} (d_n(\theta, \tilde{I}'))^{-r}. \tag{5.11}
$$

This together with the fact that $d_n(x, \tilde{I}) \leqq d_n(\theta, \tilde{I}')$, $x = \cos\theta$, gives the estimate (5.1), when everything is restated in terms of f_I and P.

The estimate (5.2) is established in much the same way. In exactly the same way that we have proved (5.3), (5.4), and (5.8), we can show that

$$
\Delta_t^r(g',\theta) = 0, \quad \text{when } [\theta, \theta+rt] \cap \tilde{I}' = \phi, \tag{5.12}
$$

$$
|\Delta_t^r(g',\theta)| \leqq Ct^k, \quad \text{when } [\theta, \theta+rt] \subseteq \tilde{I}', \tag{5.13}
$$

$$
|\Delta_t^r(g',\theta)| \leqq C A_1 n^{-k} (1+nt)^r, \quad \text{for any } t \text{ and } \theta. \tag{5.14}
$$

In exactly the same way that we have established (5.11), we can use (5.12) and (5.14) to see that

$$
|g'(\theta) - (M_n(g))'(\theta)| \leqq C A_1 n^{-k} (d_n(\theta, \tilde{I}'))^{-r}, \quad \text{for any } t \text{ and } \theta. \tag{5.15}
$$

Here, we have used the fact that $M_n(g',\theta) = (M_n(g))'(\theta)$. Restating this last inequality in terms of $\tilde{f}_I$, we find

$$
\begin{aligned}
|\tilde{f}_I'(x) - \tilde{P}'(x)| &\leqq C A_1 n^{-k} (d_n(x, \tilde{I}))^{-r} (1-x^2)^{-1/2} \\
&\leqq 2C A_1 n^{-k} (d_n(x, \tilde{I}))^{-r}, \qquad -\tfrac{1}{2} \leqq x \leqq \tfrac{1}{2}.
\end{aligned} \tag{5.16}
$$

For x in $\tilde{I}$, we can improve this last estimate. When $\theta \in \tilde{I}'$, let $\delta = \operatorname{dist}(\theta, \tilde{E}')$. Then

$$
\begin{aligned}
|g'(\theta) - (M_n(g))'(\theta)| &\leqq \int_{|t| \leqq r^{-1}\delta} |\Delta_t^r(g',\theta)| K_n(t)\,dt \\
&\quad + \int_{|t| \geqq r^{-1}\delta} |\Delta_t^r(g',\theta)| K_n(t)\,dt \\
&= \Sigma_1 + \Sigma_2.
\end{aligned} \tag{5.17}
$$

We use (5.14) on Σ_2 and estimate exactly as in (5.10) to find that

$$
\Sigma_2 \leqq C' A_1 n^{-k} (n\delta)^{-r}. \tag{5.18}
$$

For Σ_1, we use (5.13) to find

$$
\Sigma_1 \leqq C \int_{-\pi}^{\pi} |t|^k K_n(t)\,dt \leqq C' n^{-k}, \tag{5.19}
$$

because of (4.5). These last two estimates when put back into (5.17) give

$$|g'(\theta)-(M_n(g))'(\theta)|\leqq C'(A_1(d_n(\theta,\tilde{E}'))^{-r}+1)n^{-k},\qquad \theta\in\tilde{I}', \tag{5.20}$$

where we have used (5.15) to replace $n\delta$ by $d_n(\theta,\tilde{E}')$ in our estimate (5.18) of Σ_2.

Restating (5.20) in terms of f_I and using the fact that $d_n(x,\tilde{I})=1$, $x\in\tilde{I}$, gives

$$|f_I'(x)-P'(x)|\leqq 2C'(A_1(d_n(x,\tilde{E}))^{-r}+(d_n(x,\tilde{I}))^{-r})n^{-k},\qquad x\in\tilde{I}. \tag{5.21}$$

This inequality also holds for any $x\in[-\frac{1}{2},\frac{1}{2}]$ because of (5.16) and the fact that outside of $\tilde{I}$, $d_n(x,\tilde{E})\leqq d_n(x,\tilde{I})$. Restating (5.21) in terms of f_I and P gives (5.2).

6. Monotone approximation of the functions f_{Ij}^*. We can now give a monotone approximation to the functions f_{Ij}^* analogous to that given in Lemma 1 for f_J^*.

LEMMA 3. *If I is one of the intervals I_j^*, $j=1,\cdots,m$, then there is a polynomial $P_I\in\Pi_n^*$, such that*

$$|f_I(x)-P_I(x)|\leqq Cn^{-k-1}(d_n(x,I))^{-k},\qquad x\in[0,1]. \tag{6.1}$$

Proof. Let P be a polynomial of degree $\leqq n$ which satisfies Lemma 2. Since P need not be nondecreasing, we must make some corrections. The correcting polynomials will not vanish outside I as was the case for splines. Instead, these polynomials will fall off due to the factor $d_n(x,I)$. This means that all our estimates will contain terms involving $d_n(x,I)$. While this complicates matters some, the basic idea is the same as the approximation of f_I by splines given in [2].

While P' is not necessarily positive, we do have from (5.2) that

$$P'(x)\geqq f'(x)-C_2(A_1(d_n(x,E))^{-r}+(d_n(x,I))^{-r})n^{-k},\qquad x\in[0,1]. \tag{6.2}$$

Let $\gamma_1=C_2A_1\alpha_1^{-1}|E_1|n^{-k}=C_2A_1r\alpha_1^{-1}n^{-k-1}$, $\gamma_2=C_2\alpha_1^{-1}|I|n^{-k}$, and define

$$Q_1(x)=\gamma_1(\Lambda_{E_1}(x)+\Lambda_{E_2}(x))+\gamma_2\Lambda_I(x),$$

where E_1 and E_2 are the two intervals that make up E and the Λ polynomials are as defined in § 2. Now, $\Lambda_I'(x)=\lambda_I(x)$ and therefore from (2.9) and (6.2), it follows that

$$P'(x)+Q_1'(x)\geqq f'(x)\geqq 0,\qquad x\in[0,1]. \tag{6.3}$$

However, we may have added too much error and so we must take it away.

As in § 3, let $I=[i_1n^{-1},\ i_2n^{-1}]$, with $i_2-i_1=r^2\lambda+\mu$, $1<\lambda$, $0\leqq\mu<r^2$, and $x_\nu=(i_1+\nu r^2)n^{-1}$, $\nu=0,1,\cdots,\lambda$. From (3.3), we know that there is an interval $[l_0n^{-1},\ (l_0+r)n^{-1}]\subseteq I$, on which

$$P'(x)+Q_1'(x)\geqq f'(x)\geqq B_2n^{-k}. \tag{6.4}$$

Also, from (3.4), we know that for each $1\leqq\nu\leqq\lambda$ there is an interval $[l_\nu n^{-1},\ (l_\nu+r)n^{-1}]\subseteq[x_{\nu-1},x_\nu]$, on which

$$P'(x)+Q_1'(x)\geqq f'(x)\geqq B_1n^{-k}. \tag{6.5}$$

Define

$$a_1=\Lambda_I(x_1)-\Lambda_I(0),\qquad a_\lambda=\Lambda_I(1)-\Lambda_I(x_{\lambda-1}),$$
$$a_\nu=\Lambda_I(x_\nu)-\Lambda_I(x_{\nu-1}),\qquad 2\leqq\nu\leqq\lambda-1.$$

Then, because of (2.10)–(2.12), we have

$$|a_\nu| \leqq 4r^2\alpha_2|I|^{-1}n^{-1}, \qquad \nu = 1, 2, \cdots, \lambda. \tag{6.6}$$

where the r^2 appears because $|x_\nu - x_{\nu-1}| = r^2 n^{-1}$, and the 4 appears because the estimate of a_λ uses (2.12) twice and (2.11) once.

Recall the Φ polynomials introduced in § 2. Let us use the notation $\Phi_\nu = \Phi_{[l_\nu n^{-1},\,(l_\nu+r)n^{-1}]}$ and $\phi_\nu = \phi_{[l_\nu n^{-1},\,(l_\nu+r)n^{-1}]} = \Phi'_\nu$. Define

$$Q_2(x) = 2\gamma_1\Phi_0(x) + \gamma_2 \sum_1^\lambda a_\nu \Phi_\nu(x).$$

The polynomial $P_I = P + Q_1 - Q_2$ will be our approximation to f_I.

First, we want to show that P_I is nondecreasing. The polynomial $\phi_\nu = \Phi'_\nu$ is only positive on the interval $[l_\nu n^{-1}, (l_\nu + r)n^{-1}]$ and so because of (6.3), we need only check that P'_1 is positive on these intervals. Let's first consider $[l_0 n^{-1}, (l_0+r)n^{-1}]$. This interval can intersect at most two of the other intervals $[l_\nu n^{-1}, (l_\nu + r)n^{-1}]$ and so from the definition of Q_2, we find

$$\begin{aligned} Q'_2(x) &\leqq 2\gamma_1\alpha_2 n + \gamma_2(4r^2\alpha_2|I|^{-1}n^{-1})\alpha_2 n \\ &\leqq 2rC_2A_1\alpha_1^{-1}\alpha_2 n^{-k} + 4r^2C_2\alpha_2^2\alpha_1^{-1}n^{-k} \\ &\leqq B_2 n^{-k} \leqq P'(x) + Q'_1(x), \qquad x \in [l_0 n^{-1}, (l_0+r)n^{-1}], \end{aligned}$$

where the first inequality uses (2.16) and (6.6), the second uses the definition of γ_1 and γ_2, the third used the value of $B_2 = \alpha_3 2^{-r^4} A_1^2 \geqq 100r^2\alpha_1^{-1}\alpha_2^2 C_2 A_1$, and the fourth inequality is (6.4). Thus $P'_I(x) \geqq 0$ on $[l_0 n^{-1}, (l_0+r)n^{-1}]$.

For the other intervals $[l_\nu n^{-1}, (l_\nu + r)n^{-1}]$, $1 \leqq \nu \leqq \lambda$, we need only check on the parts of these intervals that don't intersect $[l_0 n^{-1}, (l_0+r)n^{-1}]$. For such an interval, we have

$$\begin{aligned} Q'_2(x) &\leqq \gamma_2|a_\nu|\,|\phi_\nu(x)| \leqq C_2\alpha_1^{-1}|I|n^{-k}(4r^2\alpha_2|I|^{-1}n^{-1})\alpha_2 n \\ &\leqq B_1 n^{-k} \leqq P'(x) + Q'_1(x), \qquad x \in [l_\nu n^{-1}, (l_\nu+r)n^{-1}] \backslash [l_0 n^{-1}, (l_0+r)n^{-1}], \end{aligned}$$

where the second inequality uses the definition of γ_2, (6.6), and (2.16). The third inequality uses the value of $B_1 = 100r^2 C_2\alpha_1^{-1}\alpha_2^2$, and the last inequality is (6.5). This shows that P_I is indeed monotone nondecreasing.

To finish the proof of Lemma 3, we need to verify (6.1). To this end, it is enough to show that

$$|Q_1(x) - Q_2(x)| \leqq Cn^{-k-1}(d_n(x, I))^{-k}, \qquad x \in [0, 1], \tag{6.7}$$

with C depending only on k. Consider first the polynomial $2\gamma_1\Phi_0 - \gamma_1(\Lambda_{E_1} + \Lambda_{E_2})$. If dist $(x, I) \geqq 2rn^{-1}$, then dist $(x, I) \leqq 2$ dist (x, E_1), dist $(x, I) \leqq 2$ dist (x, E_2), and dist $(x, I) \leqq$ dist $(x, [l_0 n^{-1}, (l_0+r)n^{-1}])$. Hence, if we use (2.10) and (2.17), we find that for $x \leqq a - 2rn^{-1}$, $I = [a, b]$, we have

$$\begin{aligned} |2\gamma_1\Phi_0(x) - \gamma_1(\Lambda_{E_1}(x) + \Lambda_{E_2}(x))| &\leqq 6\gamma_1\alpha_2(d_n(x, I))^{-2k-1} \\ &\leqq Cn^{-k-1}(d_n(x, I))^{-k}, \end{aligned} \tag{6.8}$$

where in the first inequality, we have used the facts that $r - 2 \geqq 2k - 1$ and $|E_1| = rn^{-1}$. In the second inequality we used the fact that $\gamma_1 \leqq$ const. n^{-k-1}, with the constant depending only on k.

Similarly, when $x \geqq b + 2rn^{-1}$,

$$|2\gamma_1\Phi_0(x) - \gamma_1(\Lambda_{E_1}(x) + \Lambda_{E_2}(x))| \leqq |2\gamma_1(1-\Phi_0(x)) - \gamma_1(1-\Lambda_{E_1}(x)) - \gamma_1(1-\Lambda_{E_2}(x))| \leqq Cn^{-k-1}(d_n(x, I))^{-k}, \tag{6.9}$$

because of (2.11) and (2.18).

The estimates (6.8) and (6.9) also hold when $\operatorname{dist}(x, I) \leqq 2rn^{-1}$, because the polynomials Φ_0, Λ_{E_1}, and Λ_{E_2} all have supremum norm equal to 1 on $[0, 1]$. Therefore,

$$|2\gamma_1\Phi_0(x) - \gamma_1(\Lambda_{E_1}(x) + \Lambda_{E_2}(x))| \leqq Cn^{-k-1}(d_n(x, I))^{-k}, \qquad x \in [0, 1], \tag{6.10}$$

with C depending only on k.

We will now prove an estimate like (6.10) for the polynomial $\gamma_2(\Lambda_I - \sum_1^\lambda a_\nu\Phi_\nu)$. Let $x_{-1} = 0$, $x_{\lambda+1} = 1$, and $a_0 = 0$. Then, if $x_{\nu_0} \leqq x \leqq x_{\nu_0+1}$,

$$\begin{aligned}\left|\gamma_2\Big(\Lambda_I(x) - \sum_0^\lambda a_\nu\Phi_n(x)\Big)\right| &\leqq \gamma_2\left|\sum_{-1}^{\nu_0} a_\nu(1-\Phi_\nu(x))\right| + \gamma_2|\Lambda_I(x) - \Lambda_I(x_{\nu_0})| \\ &\quad + \gamma_2\left|\sum_{\nu_0+1}^{\lambda} a_\nu\Phi_\nu(x)\right| \\ &= \Sigma_1 + \Sigma_2 + \Sigma_3.\end{aligned} \tag{6.11}$$

To estimate Σ_1, we need only observe that because of (6.6), we have $\gamma_2 \max|a_\nu| \leqq Cn^{-k-1}$, with C depending only on k. Also, if we let $s_\nu = [l_\nu n^{-1}, (l_\nu + r)n^{-1}] \subseteq [x_{\nu-1}, x_\nu] \subseteq I$ then $d_n(x, I) \leqq d_n(x, s_\nu)$, for any $x \in [0, 1]$ and $\nu = 1, \cdots, \lambda$. Hence, from (2.18)

$$\begin{aligned}\Sigma_1 &\leqq \alpha_2\gamma_2 \max|a_\nu| \sum_1^{\nu_0} (d_n(x, s_\nu))^{-2k-1} \\ &\leqq \alpha_2 Cn^{-k-1}(d_n(x, I))^{-k} \sum_1^{\nu_0} (d_n(x, s_\nu))^{-2} \leqq C'n^{-k-1}(d_n(x, I))^{-k},\end{aligned} \tag{6.12}$$

with C' depending only on k. Here, we used the fact that $\sum_1^\lambda (d_n(x, s_\nu))^{-2}$ is uniformly bounded on $[0, 1]$, because $|x_{\nu-1} - x_\nu| = r^2n^{-1}$.

The sum Σ_3 can be estimated in exactly the same way as Σ_1 except that now we use (2.17) in place of (2.18) to find

$$\Sigma_3 \leqq C'n^{-k-1}(d_n(x, I))^{-k}, \qquad x \in [0, 1]. \tag{6.13}$$

If $x \in I$, then Σ_2 is estimated by using (2.12) to find

$$\Sigma_2 \leqq \gamma\alpha_2|I|^{-1}|x - x_{\nu_0}| \leqq C'n^{-k-1}(d_n(x, I))^{-k}, \qquad x \in I,$$

because $d_n(x, I) = 1$, $x \in I$, and $|x - x_{\nu_0}| \leqq 2r^2n^{-1}$. When $x \notin I$, we use either (2.10) or (2.11) as appropriate to find

$$\begin{aligned}\Sigma_2 \leqq \gamma_2|\Lambda_I(x) - \Lambda_I(x_{\nu_0})| &\leqq 2\gamma_2\alpha_2|I|^{-1}n^{-1}(d_n(x, I))^{-r+2} \\ &\leqq C'n^{-k-1}(d_n(x, I))^{-k}, \qquad x \in I,\end{aligned}$$

because $r - 2 \geqq k$.

Putting our estimates for Σ_1, Σ_2, and Σ_3 back into (6.11) gives

$$\left|\gamma_2\left(\Lambda_I(x)-\sum_1^{\lambda} a_\nu \Phi_\nu(x)\right)\right| \leqq 3C' n^{-k-1}(d_n(x, I))^{-k}, \qquad x \in [0, 1]. \tag{6.14}$$

Finally, when we use (6.14) and (6.8), it follows that

$$|Q_1(x)-Q_2(x)| \leqq (3C'+C)n^{-k-1}(d_n(x, I))^{-k}, \qquad x \in [0, 1],$$

and so using (5.1), we have

$$|f_I(x)-P_I(x)| \leqq |f_I(x)-P(x)|+|Q_1(x)-Q_2(x)| \leqq Cn^{-k-1}(d_n(x, I))^{-k}$$

for all $x \in [0, 1]$. with C depending only on k. This proves Lemma 3.

7. Proofs of theorem. It is easy to prove Theorem 2, using the results in Lemmas 1 and 3. If $\|f^{(k+1)}\|_{L^\infty[0,1]} = 1$, then

$$f(x) = f(0)+\sum_0^m f_{J_j^*}(x)+\sum_1^m f_{I_j^*}(x)$$

as in (3.1). Let

$$P(x) = f(0)+\sum_0^m P_{J_j^*}(x)+\sum_1^m P_{I_j^*}(x),$$

where the polynomials $P_{J_j^*}$ are given in Lemma 1 and the polynomials $P_{I_j^*}$ are given in Lemma 3. P is then a monotone polynomial.

Now, $k \geqq 2$, $|J_j^*| \geqq r^2 n^{-1}$, and $|I_j^*| \geqq r^2 n^{-1}$. Hence,

$$\sum_0^m (d_n(x, J_j^*))^{-k}+\sum_1^m (d_n(x, I_j^*))^{-k} \leqq D, \qquad x \in [0, 1],$$

with D depending only on k. Therefore,

$$\begin{aligned}|f(x)-P(x)| &\leqq \sum_0^m |f_{J_j^*}(x)-P_{J_j^*}(x)|+\sum_1^m |f_{I_j^*}(x)-P_{I_j^*}(x)| \\ &\leqq Cn^{-k-1}\left(\sum_0^m (d_n(x, J_j^*))^{-k}+\sum_1^m (d_n(x, I_j^*))^{-k}\right) \leqq CDn^{-k-1},\end{aligned}$$

for any $x \in [0, 1]$, which proves Theorem 2. As we have shown in the introduction, Theorem 1 follows from Theorem 2.

REFERENCES

[1] R. De Vore, *Degree of monotone approximation*, Linear Operators and Approximation II, ISNM 25, P. Butzer, ed., Birkhauser Verlag, Basel and Stuttgart, 1974, pp. 337–351.

[2] ———, *Monotone approximation by splines*, this Journal, 8 (1977), pp. 891–905.

[3] G. G. Lorentz, *Approximation of Functions*, Holt, Reinhart and Winston, New York.

[4] ———, *Monotone approximation*, Inequalities, III, Academic Press, New York, 1972, pp. 201–215.

[5] G. G. Lorentz and K. Zeller, *Degree of approximation by monotone polynomials I*, J. Approximation Theory, 1 (1968), pp. 501–504.

[6] ———, *Degree of approximation by monotone polynomials II*, Ibid., 2 (1969), pp. 265–269.

SIAM J. MATH. ANAL.
Vol. 8, No. 5, October 1977

ERRATA: POLYGAMMA FUNCTIONS OF ARBITRARY ORDER*

NATHANIEL GROSSMAN†

R. B. Paris of the Centre d'Etudes Nucléaires has pointed out to the author that the coefficients in several expansions in the entitled paper were incorrectly calculated. They are given correctly below.

The last term in the braces in the equation for $I^P \log \Gamma(x)$ at the top of p. 369 has a spurious factor $(k+p)^{-1}$. That term should read

$$\sum_{k=2}^{\infty} (-1)^k \zeta(k) B(p+1, k) x^k.$$

The residues used to obtain the asymptotic expansion of $I^P \log \Gamma(x)$ as $|x| \to \infty$ were all incorrectly calculated. There are simple poles at $s = 2, 3, 4, \cdots$ and at $s = -2, -4, -6, \cdots$, and there are double poles at $s = 1, 0, -1, -3, -5, \cdots$. The residues are as follows.

At $s = -2k$ $(k = 1, 2, \cdots)$. Using the functional equation of $\zeta(s)$, we calculate the residue to be

$$R_{-2k} = (-1)^{k+1} x^{-2k} (2\pi)^{-2k-1} \sin \pi p \Gamma(2k-p) \zeta(2k+1).$$

At $s = 1$.

$$R_1 = -\frac{x}{\Gamma(p+2)} \{\log x - \psi(p+2)\}.$$

At $s = 0$.

$$R_0 = -\frac{1}{2\Gamma(p+1)} \{\log 2\pi x - \gamma - \psi(p+1)\}.$$

At $s = 1-2k$ $(k = 1, 2, \cdots)$. Again using the functional equation of $\zeta(s)$, we obtain the residue

$$R_{1-2k} = \frac{(-1)^k}{\pi} (2\pi x)^{1-2k} \frac{\zeta(2k)}{\Gamma(p+2-2k)} \left\{ \log 2\pi x - \psi(p+2-2k) - \frac{\zeta'(2k)}{\zeta(2k)} \right\}.$$

The asymptotic expansion for $I^P \log \Gamma(x)$ follows by summing residues, as in the article. It is more complicated than that originally given.

* This Journal, 7 (1976), pp. 366–372. Received by the editors October 8, 1976.

† Department of Mathematics, University of California, Los Angeles, California 90024.

SIAM J. MATH. ANAL.
Vol. 8, No. 6, November 1977

SOLUTION OF INTERFACE PROBLEMS BY HOMOGENIZATION III*

IVO BABUŠKA†

Abstract. The paper is the third in a series. The first two, *Solution of interface problems by homogenization I, II* are devoted to the study of the linear case. This paper studies the case of strongly nonlinear differential equations.

1. Introduction. In [1] and [2] appear surveys of some problems and applications of the homogenization method, together with extensive lists of references. For an indication of the relevance of the homogenization method in different fields of applications, we refer the reader to the first paper in this series [3] where an extensive list of references is given. In [3] and [4] we prove some basic theorems on the homogenization approach for linear equations. This paper generalizes the homogenization method to nonlinear equations. Starting from the form of results in [3], a simplified proof is presented which also holds for the nonlinear case.

We are concerned with the analysis of the solution of the following differential equation on Ω:

$$-\sum_{i,j=1}^{n} \frac{\partial}{\partial x_j} a_{i,j}(\xi, u_x) \frac{\partial u^H}{\partial x_i} = h,$$

with boundary conditions

$$u = g \quad \text{on } \partial\Omega.$$

Here, $\xi = x/H$ and $u_x = \text{grad } u$. The functions $a_{i,j}(\xi, u_x)$ are periodic with period 1 in ξ, and satisfy some growth and ellipticity conditions. The solution of the problem obviously depends on H. The main result (Theorem 4.1) describes the behavior of the solution for small H with an error estimate. Section 5 introduces an example for $n = 2$ when the coefffifficients $a_{i,j}(\xi, u_x)$ depend only on ξ_1 (i.e., are independent of ξ_2).

2. Background. We denote by R_n the n-dimensional space, $x \equiv (x_1, \cdots, x_n) \in R_n$, with $|x|^2 = \sum_{i=1}^{n} x_i^2$. Let $\Omega \subset R_n$ be a bounded domain and assume that its boundary, $\partial\Omega$, is of Lipschitz type.[1] We say, in short, Ω is a Lipschitz domain. Denote by $S^H(z)$, $z \in R_n$ the cube

$$S^H(z) = \{x \mid |x_i - z_i| < \tfrac{1}{2}H\}$$

and by γ^H the set of all grid points with the mesh size H. That is, γ^H is the set of all points $x \equiv (m_1 H, \cdots, m_n H)$ with m_i an integer. When $z = 0$ and $H = 1$ we will write S instead of $S^1(0)$.

* Received by the editors October 17, 1975, and in final revised form June 1, 1976.

† Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, College Park, Maryland 20742. This research was supported in part by the U.S. Energy Research and Development Administration under Contract AEC AT (40-1)3443. Computer time for this project was supported in part through the facilities of the Computer Science Center of the University of Maryland.

[1] The boundary, $\partial\Omega$, is of Lipschitz type when it may be locally expressed as a Lipschitz function of $n-1$ variables.

Given a function $u(x, \mu)$ defined on $R_n \times R_n$, we will use the notation $\partial u/\partial x_i = D_{x;i}u = u_{x_i}$, etc. For example, $D_{\mu;i}u(x, \mu)$ means the value of the derivative $\partial u/\partial \mu_i$ evaluated at (x, μ). The gradient of u will be denoted by u_x, and $u_x \equiv (u_{x_1}, u_{x_2}, \cdots, u_{x_n})$.

We consider the Sobolev spaces $W_p^j(\Omega)$, $1 < p < \infty$, $j = 0, 1$, and denote by $\|\cdot\|_{W_p^1(\Omega)}$ the norm in $W_p^1(\Omega)$,

$$\|u\|_{W_p^1(\Omega)} = \|u\|_{L_p(\Omega)} + \sum_{i=1}^{n} \|u_{x_i}\|_{L_p(\Omega)}$$

where $L_p(\Omega) = W_p^0(\Omega)$ is the usual space of functions whose pth power is integrable.

It is well known (see, e.g., [5]) that any $u \in W_p^1(\Omega)$ has a trace on $\partial\Omega$. The subspace of $W_p^1(\Omega)$ of all functions with zero trace on $\partial\Omega$ will be denoted by $\mathring{W}_p^1(\Omega)$, and $W_p^{-1}(\Omega)$ will be the space of linear functionals over $\mathring{W}_p^1(\Omega)$.

We also will consider the space of periodic functions with period 1 and understand these functions to be defined on S. Denote the space of these functions by $W_{p,\mathrm{PER}}^1(S)$. Obviously $L_{p,\mathrm{PER}}(S) = L_p(S)$.

Thus far we have considered only scalar functions. We will deal also with vector functions $u \equiv (u^{[1]}, \cdots, u^{[n]})$, with norm defined in the natural way:

$$\|u\|_{W_p^1(\Omega)} = \sum_{l=1}^{n} \|u^{[i]}\|_{W_p^1(\Omega)}.$$

All other notation is extended analogously to vector functions.

Let us formulate a lemma which will be used later.

LEMMA 2.1. *Let $u \in \mathring{W}_p^1(\Omega)$. Then there exists a function $\{u\}_H \in L_p(\Omega)$ such that it is constant on every $S^H(z)$, $z \in \gamma^H$, $S^H(z) \subset \Omega$, $\{u\}_H = 0$ on $\Omega - \bar{\Omega}^H$, where*

$$\Omega^H = \bigcup_{\substack{z \in \gamma^H \\ S^H(z) \subset \Omega}} S^H(z)(\Omega)$$

and for all $H < H_0$

$$\|u - \{u\}_H\|_{L_p(\Omega)} \leqq CH\|u\|_{W_p^1(\Omega)} \tag{2.1}$$

where C is independent of H and u.

Proof. 1) Because $u \in \mathring{W}_p^1(\Omega)$ it can be extended by zero to R_n, preserving its norm, i.e., $\|u\|_{W_p^1(R_n)} = \|u\|_{W_p^1(\Omega)}$. It can be shown by the usual arguments that

$$\|u(x + \Delta) - u(x)\|_{L_p(R_n)} \leqq C|\Delta|\,\|u\|_{W_p^1(\Omega)}. \tag{2.2}$$

The domain Ω is Lipschitz by assumption and hence we can write (by partition of unity arguments)

$$u = \sum_{i=1}^{k} u_i,$$

$$\|u_i\|_{W_p^1(\Omega)} \leqq C\|u\|_{W_p^1(\Omega)},$$

so that the functions $u_i(x+\Delta_i\lambda)$, $|\Delta_i|=1$ have compact support in Ω for any $0<\lambda<H_0$. Thus

$$v^H=\sum_{i=1}^{k} u_i(x+\Delta_i HC)$$

is a function with support in $\bar{\Omega}^H$, where $C>0$ is a properly chosen constant. Inequality (2.2) yields

$$\|v^H-u\|_{L_p(\Omega)}\leqq CH\|u\|_{W_p^1(\Omega)}. \tag{2.3}$$

2) Now we may construct function $\{u\}_H$ so that its value on $S^H(z)$ is the average of v^H over $S^H(z)$. Then it can be shown that

$$\begin{aligned}\|\{u\}_H-u\|_{L_p(\Omega)}&\leqq\|\{u\}_H-v^H\|_{L_p(\Omega)}+\|v^H-u\|_{L_p(\Omega)}\\&\leqq CH\|u\|_{W_p^1(\Omega)},\end{aligned}$$

and the lemma is proved.

3. The boundary value problem.

3.1. The nonlinear Poisson problem. Consider the Poisson boundary value problem for a quasilinear equation on Ω,

$$L(u)=-\sum_{i,j=1}^{n} D_{x;j}a_{i,j}(\xi,\mu)D_{x;i}u(x)=h(x) \tag{3.1}$$

with the boundary condition

$$u=g \quad \text{on } \partial\Omega \tag{3.2}$$

where $\xi=x/H$, $0<H\leqq 1$ and $\mu=u_x$. We are assuming

(i) the functions $a_{i,j}(\xi,\mu)=a_{j,i}(\xi,\mu)$, $i,j=1,\cdots,n$, are defined on $R_n\times R_n$ and are periodic in ξ with period 1;

(ii)

$$\sum_{i,j=1}^{n}|a_{i,j}(\xi,\mu)|\leqq K(1+|\mu|)^{p-2}, \tag{3.3}$$

$$\sum_{l,i,j=1}^{n}|D_{\mu;l}a_{i,j}(\xi,\mu)|\leqq K(1+|\mu|)^{p-3}, \tag{3.4}$$

$$\sum_{i,j,k,l}^{n}|D_{\mu;k}D_{\mu;l}a_{i,j}(\xi,\mu)|\leqq K(1+|\mu|)^{p-4}, \tag{3.5}$$

where K is independent of (ξ,μ) and $p\geqq 2$;

(iii) g in (3.2) is the trace on $\partial\Omega$ of a function $G\in W_p^1(\Omega)$, and $h\in W_p^{-1}(\Omega)$.

Condition (3.3) yields that for any $u\in W_p^1(\Omega)$ $(1/p+1/q=1)$,

$$\begin{aligned}&\left\|\sum_{l=1}^{n} a_{i,j}\left(\frac{x}{H},u_x\right)D_{x;i}u\right\|_{L_q(\Omega)}\\&\qquad\leqq C(1+\|u\|_{W_p^1(\Omega)})^{p-2}\|u\|_{W_p^1(\Omega)},\end{aligned} \tag{3.6}$$

with C independent of H and u. Therefore on $W_p^1(\Omega)\times \mathring{W}_p^1(\Omega)$, the form $B(u, v)$,

$$B(u, v)=\int_\Omega \left[\sum_{i,j=1}^{n} a_{i,j}\left(\frac{x}{H}, u_x\right)D_{x;i}uD_{x;j}v\right] dx, \tag{3.7}$$

is continuous in v and

$$|B(u, v)|\leqq C(1+\|u\|^p_{W^1(\Omega)})^{1/q}\|v\|_{W_p^1(\Omega)}. \tag{3.8}$$

Let us now assume that the problem (3.1) and (3.2) is well posed. More precisely,

(iv) let $\bar{u}\in W_p^1(\Omega)$ such that $\bar{u}=\bar{G}+\bar{w}$, $\bar{G}\in W_p^1(\Omega)$, $\bar{w}\in \mathring{W}_p^1(\Omega)$, and

$$\left|B(\bar{u}, v)-\int_\Omega hv\, dx\right|\leqq \gamma\|v\|_{W_p^1(\Omega)} \tag{3.9}$$

for any $v\in \mathring{W}_p^1(\Omega)$, and $\|G-\bar{G}\|_{W_p^1(\Omega)}\leqq\gamma$, $0<\gamma\leqq\gamma_0$. Then there exists at least one $u\in W_p^1(\Omega)$, $u=G+w$, $w\in \mathring{W}_p^1(\Omega)$ such that

$$B(u, v)=\int_\Omega hv\, dx$$

for any $v\in \mathring{W}_p^1(\Omega)$, and

$$\|\bar{u}-u\|_{W_p^1(\Omega)}\leqq C\gamma^\rho, \qquad 0<\rho\leqq 1, \tag{3.10}$$

where ρ and C depend on γ_0 only. Quasilinear problems satisfying conditions (i)–(iv) are well studied. See [6], [7], or [8].

3.2. The associated problem. Corresponding to the operator (3.1) is another problem which we will call the associated problem. For any given $\sigma\in R_n$ we seek a periodic (vector) function $\chi(x, \sigma)\equiv(\chi^{[1]}(x, \sigma), \cdots, \chi^{[n]}(x, \sigma))\in W_{p,\mathrm{PER}}^1(S)$, such that

$$\int_S \chi(x, \sigma)\, dx=0, \tag{3.11}$$

and

$$-\sum_{i,j} D_{x;j}a_{i,j}(x, \mu)(\delta_i^k+D_{x;i}\chi^{[k]})=0, \qquad k=1, \cdots, n, \tag{3.12}$$

where $\mu=(\mu_1, \cdots, \mu_n)$,

$$\mu_i=\sum_{k=1}^{n}\sigma_k(\delta_i^k+D_{x;i}\chi^{[k]})$$

and δ_i^k is the Kronecker symbol.

Similarly to equation (3.1), equation (3.12) must be understood in the weak sense. That is, we have for $k=1, \cdots, n$

$$\int_S \left[\sum_{i,j=1}^{n} a_{i,j}(x, \mu)(\delta_i^k+D_{x;i}\chi^{[k]})D_{x;j}v\right] dx=0 \tag{3.13}$$

for any $v \in W^1_{p,\mathrm{PER}}(S)$. Obviously, if $\chi(x,\sigma)$ satisfies (3.12), so does $\chi(x,\sigma)+$ constant. Equation (3.11) normalizes this arbitrary vector constant.

We make the following assumptions:

(i) For every $\sigma \in R_n$ there exists at least one solution $\chi(x,\sigma) \in W^1_{p,\mathrm{PER}}(S)$.

(ii) The function $\chi(x,\sigma)$ has two derivatives with respect to σ, i.e.,

$$D_{\sigma;i}\chi(x,\sigma) \in W^1_{p,\mathrm{PER}}(S), \qquad D_{\sigma;i}D_{\sigma;j}\chi(x,\sigma) \in W^1_{p,\mathrm{PER}}(S).$$

(iii) Let $\sigma(x)$ be any function of $x \in S$ such that

$$|D_{x;j}\sigma| \leqq \xi, \qquad 0 < \xi \leqq \xi_0, \quad j = 1, \cdots, n. \tag{3.14}$$

Then $\chi(x,\sigma(x))$, $D_{\sigma;i}\chi(x,\sigma(x))$, and $D_{\sigma;i}D_{\sigma;j}\chi(x,\sigma(x))$ are functions in $W^1_{p,\mathrm{PER}}(S)$, and

$$\|\chi(x,\sigma(x)) - \chi(x,\sigma(0))\|_{W^1_p(S)} \leqq C\xi, \tag{3.15}$$

with the same inequality valid for $D_{\sigma;i}\chi(x,\sigma(x))$ and $D_{\sigma;i}D_{\sigma;j}\chi(x,\sigma(x))$.

The results of [6] can be used for the analysis of problem (3.12). As in the previous section, we postulate only the required properties. Now we define

$$A_{l,k}(\sigma) = \int_S \left[\sum_{i,j=1}^{n} a_{i,j}(x,\mu)(\delta_i^l + D_{x;i}\chi^{[l]}(x,\sigma))(\delta_j^k + D_{x;j}\chi^{[k]}(x,\sigma)) \right] dx. \tag{3.16}$$

Because of (3.13) we have also

$$\begin{aligned} A_{l,k}(\sigma) &= \int_S \left[\sum_{i,j=1}^{n} a_{i,j}(x,\mu)(\delta_i^l + D_{x;i}\chi^{[l]})\delta_j^k \right] dx \\ &= \int_S \left[\sum_i a_{i,k}(x,\mu)(\delta_i^l + D_{x;i}\chi^{[l]}) \right] dx. \end{aligned} \tag{3.17}$$

Equation (3.16) yields immediately that $A_{i,j}(\sigma) = A_{j,i}(\sigma)$ and that ellipticity of the operator L leads to the ellipticity of the operator

$$\bar{L}(\sigma,u) = -\sum_{i,j=1}^{n} D_{x;j}A_{i,j}(\sigma)D_{x;i}. \tag{3.18}$$

Assumptions (3.3)–(3.5), together with assumption (ii) about $\chi^{[l]}$, yield that the functions $A_{l,k}(\sigma)$, $l, k = 1, \cdots, n$ have two derivatives with respect to σ.

3.3. The homogenized problem. In § 3.1 we introduced the problem (3.1) and (3.2) and in § 3.2 concerned ourselves with the associated problem. Now we will formulate the homogenized problem as the Poisson problem

$$\bar{L}(u) = -\sum_{i=1}^{n} D_{x;j}A_{i,j}(u_x)D_{x;i}u = h \tag{3.19}$$

with boundary conditions

$$u = g. \tag{3.20}$$

The functions h and g in (3.19) and (3.20) are the same as in (3.1) and (3.2). We will assume that there exists at least one solution u of (3.19) and (3.20). Later we will assume that the solution is smooth. In applications, our assumptions about the homogenized problem are a consequence of the assumptions about the Poisson problem (3.1) and (3.2) and about the associated problem (3.11).

4. The homogenization.

4.1. Formulation of the problem. The problem (3.1) and (3.2) has rapidly changing coefficients with the scale H. The solution obviously depends on H. Denote it by u^H. The question is what is its behavior when $H \to 0$. This problem is studied in [1]–[4] when the $a_{i,j}$ are independent of μ. Here we study it (by another method) when the coefficients depend on the gradient of the solution.

4.2. The homogenization.

THEOREM 4.1. *Let there exist a solution of the homogenized problem* (3.19) *and* (3.20). *In addition, assume that the solution U has three bounded derivatives on Ω. Then for every $0 < H < H_0$ there exist at least one solution u^H of* (3.1) *and* (3.2), *and*

$$\left\| u^H - U - H \sum_{k=1}^{n} \chi^{[k]}\left(\frac{x}{H}, U_x\right) D_{x;k} U \right\|_{W_p(\Omega)} \leqq C H^{\rho/p} \tag{4.1}$$

where $\chi(x, \sigma)$ is the solution of the associated problem, C is independent of H (it depends on U), and ρ is given in (3.10).

Proof. 1. Define

$$\begin{gathered} w^H(x) = U(x) + H \sum_{l=1}^{n} \chi^{[l]}(\xi, \sigma) D_{x;l} U, \\ \text{with } \xi = x/H \quad \text{and} \quad \sigma = U_x, \end{gathered} \tag{4.2}$$

and compute $D_{x;j} w^H(x)$. We get

$$D_{x;j} w^H = D_{x;j} U + \sum_{l=1}^{n} D_{\xi;j} \chi^{[l]}(\xi, \sigma) D_{x;l} U + H V_j, \tag{4.3}$$

where

$$\begin{aligned} V_j = \sum_{l=1}^{n} \Big[& \chi^{[l]}(\xi, \sigma) D_{x;j} D_{x;l} U \\ & + D_{x;l} U \sum_{k=1}^{n} D_{\sigma;k} \chi^{[l]}(\xi, \sigma) D_{x;j} D_{x;k} U \Big]. \end{aligned} \tag{4.4}$$

Now, using assumptions (ii) and (iii) about the associated problem (in § 3.2) and the fact that U has two bounded derivatives, it readily follows that

$$\|V\|_{W_p^1(\Omega)} \leqq C \tag{4.5}$$

and C does not depend on H.

2. We will show now that

$$\left| B(w^H, v) - \int_\Omega hv\, dx \right| \leqq CH\|v\|_{W^1{}_p(\Omega)} \tag{4.6}$$

for any $v \in \mathring{W}^1_p(\Omega)$, with C independent of H and v and where $B(w^H, v)$ is given by (3.7). We write

$$B(w^H, v) = \Phi_1^H(v) + R_1^H(v)$$

with

$$\Phi_1^H(v) = \int_\Omega \left[\sum_{i,j=1}^n a_{i,j}(\xi, \bar{\mu})(D_{x;i}U + \sum_{l=1}^n D_{x;l}UD_{\xi,i}\chi^{[l]})D_{x;j}v \right] dx \tag{4.7}$$

and

$$\bar{\mu} = U_x + \sum_l \chi_\xi^{[l]}(\xi, \sigma)D_{x;l}U$$

and wish to estimate $R_1^H(v)$. We have

$$w_x^H = \bar{\mu} + HV \ ,$$

and therefore for $2 \leqq p \leqq 3$ and $3 \leqq p < \infty$

$$\begin{aligned} |a_{i,j}(\xi, w_x^H) - a_{i,j}(\xi, \bar{\mu})| &= |a_{i,j}(\xi, \bar{\mu} + HV) - a_{i,j}(\xi, \bar{\mu})| \\ &= \left| \int_0^1 \left(\sum_{k=1}^n D_{\mu;k}a_{i,j}(\xi, \bar{\mu} + tHV)HV_k \right) dt \right| \\ &\leqq CH|V| \int_0^1 (1 + |\bar{\mu} + tHV|)^{p-3}\, dt \\ &\leqq CH|V|(1 + |\bar{\mu}| + H|V|)^{p-3}. \end{aligned}$$

So

$$\begin{aligned} |R_1(v)| &\leqq \left| \int_\Omega \left[\sum_{i,j=1}^n (a_{i,j}(\xi, w_x^H) - a_{i,j}(\xi, \bar{\mu}))(\bar{\mu}_i + HV_i) + a_{i,j}(\xi, \bar{\mu})HV_i \right] D_{x;j}v\, dx \right| \\ &\leqq CH \Bigg[\int_\Omega [(1 + |\bar{\mu}| + H|V|)^{p-3}|V|((|\bar{\mu}| + H|V|) \\ &\qquad + (1 + |\bar{\mu}|)^{p-2}|V|)]^{p/(p-1)}\, dx \Bigg]^{1/q} \|v\|_{W^1_p(\Omega)} \\ &\leqq CH\|v\|_{W^1_p(\Omega)}. \end{aligned} \tag{4.8}$$

3. By (4.8) we may replace $B(w^H, v)$ in (4.6) by Φ_1^H. Thus we will restrict our analysis to Φ_1^H as

$$\Phi_1^H = \Phi_2^H - \Phi_3^H \tag{4.9}$$

where

$$\Phi_2^H(v)=\int_\Omega\left[\sum_{i,j,l}^{n}(D_{x;j}(vD_{x;l}U)a_{i,j}(\xi,\bar{\mu})\cdot(\delta_i^l+D_{\xi,i}\chi^{[l]}(\xi,\sigma))\right]dx, \tag{4.10}$$

$$\Phi_3^H(v)=\int_\Omega\left[\sum_{i,j,l}(vD_{x;j}D_{x;l}U)a_{i,j}(\xi,\bar{\mu})\cdot(\delta_i^l+D_{\xi;i}\chi^{[l]}(\xi,\sigma))\right]dx, \tag{4.11}$$

and we will show that

$$\Phi_2^H(v)=-\int_\Omega\left[\sum_{l,j=1}^{n}(vD_{x;l}UD_{x;j})A_{l,j}(\sigma(x))\right]dx+R_2(v), \tag{4.12}$$

$$\Phi_3(v)=\int_\Omega\left[\sum_{l,j=1}^{n}(vD_{x;j}D_{x;l}U)A_{l,j}(\sigma(x))\right]dx+R_3(v), \tag{4.13}$$

with

$$|R_i(v)|\leqq CH\|v\|_{W_p^1(\Omega)},\qquad i=2,3. \tag{4.14}$$

4. Let us prove (4.12). The proof of (4.13) is similar. Integrating (4.10) by parts and taking into account (3.13), we obtain

$$\begin{aligned}\Phi_2^H(v)&=-\int_\Omega\left[v\sum_{i,j,l}D_{x;l}U\left[\sum_{k=1}^{n}D_{\sigma;k}[a_{i,j}(\xi,\bar{\mu}(\sigma))\cdot(\delta_i^l+D_{\xi;i}\chi^{[l]}(\xi,\sigma))]D_{x;k}D_{x;j}U\right]\right]dx\\&=-\sum_{\substack{z\in\gamma^H(z)\\S^H(z)\subset\Omega}}\int_{S^H(z)}\left[\sum_{i,j,l}\{vD_{x;l}U\}_H\left[\sum_{k-1}^{n}D_{\sigma;k}[a_{ij}(\xi_j,\bar{\mu}(\sigma))\cdot(\delta_i^l+D_{\xi,i}\chi^{[l]}(\xi,\sigma))]D_{x;k}D_{x;j}U\right]\right]dx+R_4(v).\end{aligned} \tag{4.15}$$

Because

$$\left\|\left[\sum_{k=1}^{n}D_{\sigma;k}[a_{i,j}(\xi,\bar{\mu}(\sigma))(\delta_i^l+D_{\xi;i}\chi^{[l]}(\xi,\sigma))]D_{x;k}D_{x;j}U\right]\right\|_{L_q(\Omega)}\leqq C, \tag{4.16}$$

where C is independent of H, we readily get (using Lemma 2.1)

$$|R_4(v)|\leqq CH\|v\|_{W_p^1(\Omega)}. \tag{4.17}$$

Inequality (4.16) is a simple consequence of assumption (ii) of § 3.2 about the associated problem, and of inequalities (3.3)–(3.5). It is also possible to show (by easy, but slightly tedious, computations) that

$$-\sum_{\substack{z\in\gamma^H\\S^H(z)\subset\Omega}}\int_{S^H(z)}\left[\sum_{i,j,l}\{vD_{x;l}U\}_H\left[\sum_{k=1}^{n}D_{\sigma;k}[a_{i,j}(\xi,\bar{\mu}(\sigma))\cdot(\delta_i^l+D_{\xi;i}\chi^{[l]}(\xi,\sigma))]D_{x;k}D_{x;j}U\right]\right]dx \tag{4.18}$$

$$= -\sum_{\substack{z\in\gamma^H \\ S^H(z)\subset R}} \int_{S^H(z)} \Big[\sum_{i,j,l} \{vD_{x;l}U\}_H \Big[\sum_{k=1}^{n} D_{\sigma,k}[a_{i,j}(\xi, \bar{\mu}(\sigma(z)))$$

$$\cdot (\delta_i^l + D_{\xi;i}\chi(\xi, \sigma(z)))]D_{x;k}D_{x;j}U(z)\Big]\Big]\, dx + R_5(v)$$

$$= -\sum_{\substack{z\in\gamma^H \\ S^H(z)\subset\Omega}} \sum_{l}^{n} \{vD_{x;l}U\}_H \int_{S^H(z)} \Big[\sum_{k,j=1}^{n} D_{\sigma;k}A_{l,j}(\sigma(z))D_{x;k}D_{x;j}U(z)\Big]\, dx + R_5(v)$$

$$= -\sum_{\substack{z\in\gamma^H \\ S^H(z)\subset\Omega}} \sum_{l,j=1}^{n} \{vD_{x;l}U\}_H D_{x;j}A_{l,j}((\sigma(x))) + R_5(v) + R_6(v)$$

$$= -\int_\Omega \Big[\sum_{l,j=1}^{n} vD_{x;l}UD_{x;j}A_{l,j}(\sigma(x))\Big]\, dx + R_5(v) + R_6(v) + R_7(v)$$

and that

$$|R_5(v) + R_6(v) + R_7(v)| \leqq CH\|v\|_{W_p^1(\Omega)}. \tag{4.19}$$

Thus we have shown that

$$\begin{aligned}\Phi_1^H &= -\int_\Omega \Big[\sum_{l,j=1}^{n} (vD_{x;l}U)D_{x;j}A_{l,j}(\sigma) \\ &\qquad + (vD_{x;l}D_{x;j}U)A_{l,j}(\sigma)\Big]\, dx + R_8(v) \\ &= \int_\Omega \Big[\sum_{l,j=1}^{n} D_{x;l}UA_{l,j}(\sigma)D_{x;j}v\Big]\, dx + R_8(v),\end{aligned} \tag{4.20}$$

with

$$|R_8(v)| \leqq CH\|v\|_{W_p^1(\Omega)}.$$

This gives in turn that

$$B(w^H, v) = \int_\Omega \Big[\sum_{l,j=1}^{n} A_{l,j}(\sigma)D_{x,l}UD_{x,j}v\Big]\, dx + R(v), \tag{4.21}$$

where

$$|R(v)| \leqq CH\|v\|_{W_p^1(\Omega)}. \tag{4.22}$$

Using (3.19), (4.21) and (4.22) gives (4.6).

5. The function U satisfies the boundary conditions; w^H does not. Let us show that $U - w^H$ on $\partial\Omega$ is a trace of the function G^H, so that $\|G - G^H\|_{W_p^1(\Omega)} \leqq CH^{1/p}$. Here G is the function whose trace is g, as in the assumption (iii) about the Poisson problem (§ 3.1).

Obviously, for any $H > 0$ there exists a smooth function κ^H such that $\kappa^H = 1$ in a neighborhood of $\partial\Omega$, $\kappa^H(x) = 0$ for any x such that $\operatorname{dist}(x, \partial\Omega) \geqq 2H$ and $|\kappa_x^H| \leqq C/H$.

Using (4.2) and the assumptions about $\chi^{[l]}(x, \sigma)$, we readily see that

$$\|(U - w^H)\kappa^H\|_{W_p^1(\Omega)} \leqq CH^{1/p},$$

where C does not depend on H. So for $G^H = G+(U-G)\kappa^H-(U-w^H)\kappa^H$ the trace of G^H on $\partial\Omega$ is the same as that for w^H, and $\|G^H-G\|_{W_p^1} \leqq CH^{1/p}$. Assumption (iv) on the Poisson problem, (§ 3.1) gives (4.1), and so Theorem 4.1 is proved.

Theorem 4.1 and its proof lead to the following.

THEOREM 4.2. *Let U and u^H be the functions in Theorem* 4.1. *Let*

$$\mathcal{T}_j^H(x) = \sum_{i=1}^{n} a_{i,j}(\xi,\mu)D_{x;i}u^H(x) \tag{4.23}$$

$$(\xi = x/H, \quad \mu = u_x^H),$$

$$T_j^H(x) = \frac{1}{H^n}\int_{S^H(x)} \mathcal{T}_j^H(z)\,dz \tag{4.24}$$

with

$$S^H(x) \subset \Omega,$$

and

$$W_j(x) = \sum_{i=1}^{n} A_{ij}(\sigma)D_{x;i}U, \qquad \sigma = U_x. \tag{4.25}$$

Then

$$\|T_j^H(x) - W_j(x)\|_{L_q(\Omega)} \leqq CH^{\rho/p}, \tag{4.26}$$

for any $\bar{\Omega} < \Omega$ such that dist $(\bar{\Omega}, \partial\Omega) > H$.

Proof. $T_j^H(x)$ is defined on $\bar{\Omega}$. Repeating the argument used in (4.18) and (4.20), we get

$$\|\tilde{T}_j^H(x) - W_j(x)\|_{L_q(\bar{\Omega})} \leqq CH \tag{4.27}$$

where

$$\tilde{\mathcal{T}}_j^H(x) = \sum_{l=1}^{n} a_{i,j}(\xi,\tilde{\mu})D_{x,i}w^H(x),$$

$$\xi = x/H, \qquad \tilde{\mu} = w_x^H$$

and

$$\tilde{T}_j^H(x) = \frac{1}{H^n}\int_{S^H(x)} \tilde{\mathcal{T}}_j^H(z)\,dz. \tag{4.28}$$

Because

$$\|w^H - u^H\|_{W_p^1(\Omega)} \leqq CH^{\rho/p},$$

we have

$$\|\tilde{\mathcal{T}}_j^H - \mathcal{T}_j^H\|_{L_q(\bar{\Omega})} \leqq CH^{\rho/\rho}, \tag{4.29}$$

and therefore also

$$\|\tilde{T}_j^H - T_j^H\|_{L_q(\bar{\Omega})} \leqq CH^{\rho/p}. \tag{4.30}$$

Inequality (4.27) and (4.30) give the desired result.

5. Applications.

5.1. Linear equations. Let us assume that coefficients $a_{i,j}(\xi, \mu)$ are independent of μ, i.e., $a_{i,j}(\xi, \mu) = a_{i,j}(\xi)$. Then we may take $p = 2$, and obviously (3.3)–(3.5) are satisfied when $|a_{i,j}(\xi)| \leqq K$.

Assume now that

$$\sum_{i,j=1}^{n} a_{i,j}(\xi)\eta_i\eta_j \geqq M \sum_{i=1}^{n} \eta_i^2, \qquad M > 0. \tag{5.1}$$

Then the operator is elliptic and all assumptions introduced in § 3.1 are satisfied.

The associated problem will split, in this case, into n independent problems with periodic boundary conditions. The existence of these solutions follows from the ellipticity of the operator L. Obviously the homogenized operator $\bar{L}$ now has constant coefficients and its ellipticity follows from (5.1) and (3.16). Therefore, if $\partial\Omega$ is smooth, smooth right hand side and boundary conditions assure that the solution U is smooth, too. Thus all of the assumptions of Theorems 4.1 and 4.2 are satisfied, giving (as a special case) part of the results included in [3].

5.2. The problem of a laminated medium. As an example, consider the two dimensional problem

$$-D_{x;1}a(\xi, u_x)D_{x;1}u - D_{x;2}a(\xi, u_x)D_{x;2}u = h \tag{5.2}$$

in Ω, with $u = g$ on $\partial\Omega$. Assume that with $0 < b < \frac{1}{2}$

$$a(\xi, u_x) = a_1(\|u_x\|^2) > 0$$

for

$$-\tfrac{1}{2} < \xi_1 < -b, \qquad b < \xi_1 < \tfrac{1}{2},$$
$$a(\xi, u_x) = a_2(\|u_x\|^2) > 0$$

for

$$-b < \xi_1 < b,$$

where

$$\|u_x\|^2 = (D_{x;1}u)^2 + (D_{x;2}u)^2.$$

The homogenization approach now consists of solving the associated problem (3.11) for the unknown functions

$$\chi^{[1]}(x, \sigma), \quad \chi^{[2]}(x, \sigma), \qquad \sigma \in R_2.$$

In this case, (3.12) reduces to

$$\begin{aligned} &-D_{x;1}a(x_1, \mu_1^2 + \mu_2^2)(1 + D_x z_{;1}\chi^{[1]}) \\ &\qquad -D_{x;2}a(x_1, \mu_1^2 + \mu_2^2)D_{x;2}\chi^{[1]} = 0, \\ &-D_{x;1}a(x_1, \mu_1^2 + \mu_2^2)D_{x;1}\chi z^{[2]} \\ &\qquad -D_{x;2}a(x_1, \mu_1^2 + \mu_2^2)(1 + D_{x;2}\chi^{[2]}) = 0, \end{aligned} \tag{5.3}$$

$$\begin{aligned} \mu_1 &= \sigma_1(1 + D_{x;1}\chi^{[1]}) + \sigma_2 D_{x;1}\chi^{[2]}, \\ \mu_2 &= \sigma_1 D_{x;2}\chi^{[1]} + \sigma_2(1 + D_{x;2}\chi^{[2]}). \end{aligned} \tag{5.4}$$

These equations are to be understood in a weak sense. For example, (5.3) and (5.4) may be satisfied with $\chi^{[2]}=0$ and $\chi^{[1]}(x_1,\sigma)$ depending on x_1 only and satisfying the ordinary differential equation in x_1,

$$\frac{-d}{dx_1}a\left(x_1,\left(\sigma_1\left(1+\frac{d}{dx_1}\chi^{[1]}\right)\right)^2+\sigma_2^2\right)\left(1+\frac{d}{dx_1}\chi^{[1]}\right)=0. \tag{5.5}$$

We can readily see that

$$\frac{d\chi^{[1]}}{d\chi_1}(\chi,\sigma)=z_1 \quad (\text{resp. } z_2), \quad z_i \text{ constant}$$

in the intervals $b<|x_1|<\frac{1}{2}$ $(-b\leqq x_1\leqq b)$. The problem reduces to solving a nonlinear system for unknown z_1 and z_2. We get

$$A_{1,1}(\sigma)=\int_{-1/2}^{1/2} a(x,\mu)\left(1+\frac{d}{dx}\chi^{[1]}(x,\sigma)\right)dx,$$

$$A_{1,2}(\sigma)=0,$$

$$A_{2,2}(\sigma)=\int_{-1/2}^{1/2} a(x,\mu)\,dx.$$

In [7, § 68–69], an equation of the form (5.1) is studied under the assumptions that

(i)

$$\bar{A}_0+\bar{A}_1 z^{p/2-1}\leqq a(z)\leqq A_0+A_1 z^{p/2-1}, \qquad 0<z<\infty, \tag{5.6}$$

and, $\bar{A}_0$, $\bar{A}_1$, A_0, A_1 are positive numbers,

$$a'(z)\leqq K|z|^{p/2-2} \tag{5.7}$$

(realizing that $z=|\mu|^2$, the right hand sides of (5.6) and (5.7) are equivalent to those of (3.3) and (3.4).);

(ii) The function a is twice continuously differentiable. (This condition is close to (3.5).); and

(iii)

$$a(z)+2z\frac{da}{dz}\geqq A>0. \tag{5.8}$$

Under these hypotheses, all assumptions we introduced earlier are satisfied.

Let us compute now a special example. Assume

$$a_i(z)=\alpha_i, \qquad 0\leqq z\leqq\mu_i,$$

$$a_i(z)=\alpha_i+\beta_i(z^2-\mu_i^2)^\gamma, \qquad \mu_i\leqq z<\infty,$$

$$\alpha_i>0, \quad \beta_i>0, \quad \gamma\geqq 2.$$

All of the assumptions made above are satisfied with $p=2(\gamma+1)$ and $\rho=1/(p-1)$. Taking the torsion problem (see [7]) as a physical interpretation, the values of μ_i are yield points for a composite material.

For

$$b = .2, \quad \alpha_1 = 1.0, \quad \alpha_2 = 10, \quad \mu_1 = 1.5,$$
$$\beta_1 = 5, \quad \beta_2 = .1, \quad \gamma = 2, \quad \mu_2 = 10,$$

the functions $A_{1,1}(\sigma_1, \sigma_2)$ and $A_{2,2}(\sigma_1, \sigma_2)$ are shown in Figs. 1 and 2. Because of symmetry, only $\sigma_1 \geqq 0$ and $\sigma_2 \geqq 0$ are considered.

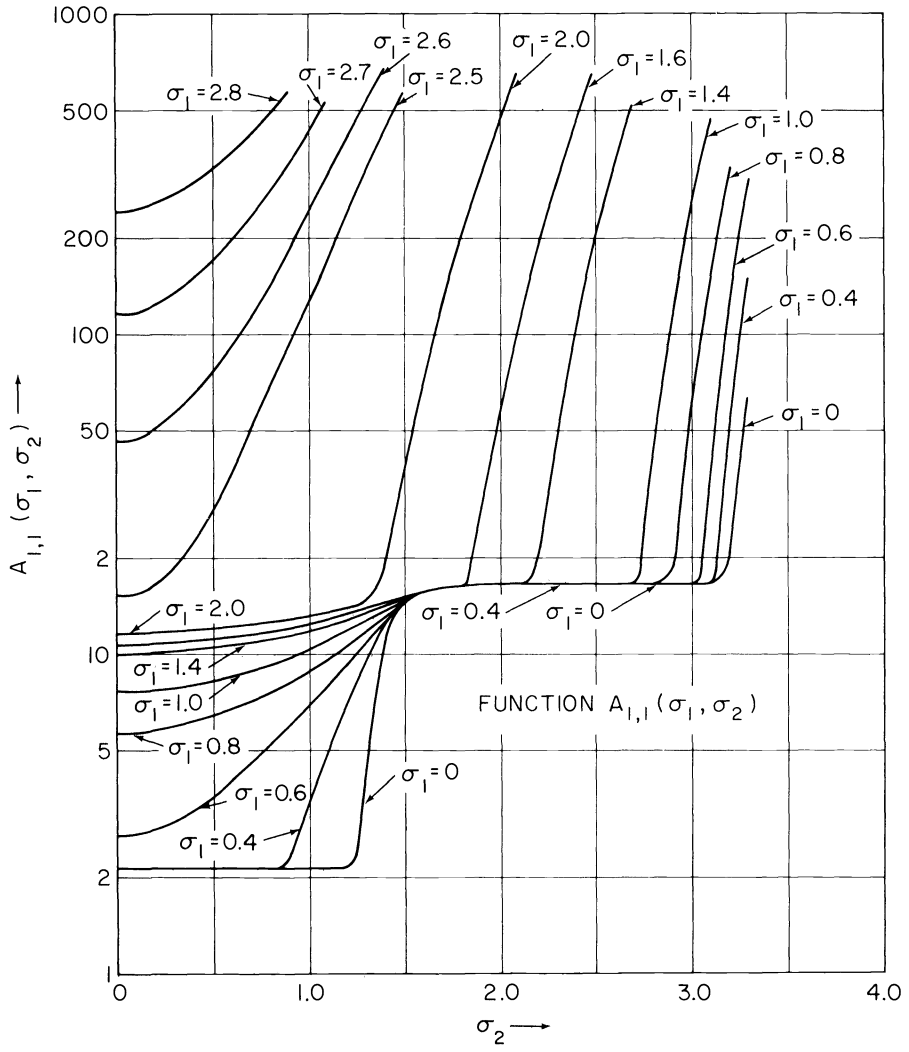


FIG. 1

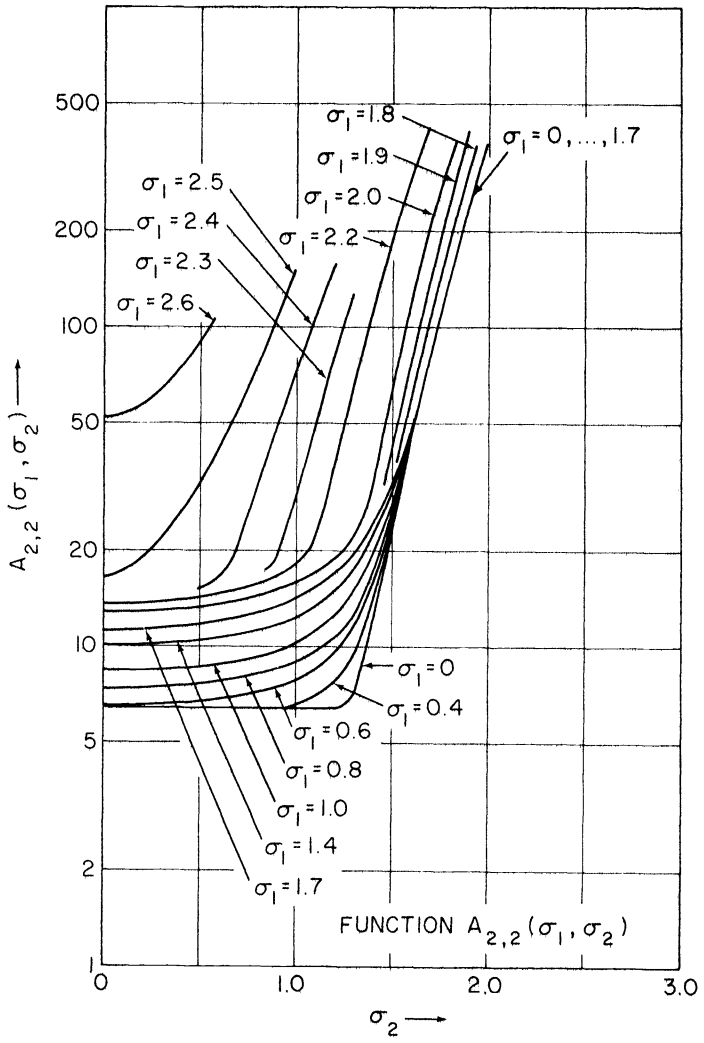


FIG. 2

REFERENCES

[1] I. BABUŠKA, *Solution of problems with interfaces and singularities*, Mathematical Aspects of Finite Elements in Partial Differential Equations, C. de Boor, ed., Academic Press, New York, 1974, pp. 213–279.

[2] ———, *Homogenization and its applications*, Mathematical and Computational Problems, Proceedings of SYNSPADE Conference, B. Hubbard, ed., Academic Press, New York, 1976, to appear.

[3] ———, *Solution of interface problems by homogenization I*, Tech. Note BN-782, Institute for Fluid Dynamics and Applied Mathematics, Univ. of Maryland, 1974; this Journal, 7 (1976), pp. 603–634.

[4] ———, *Solution of interface problems by homogenization II*, Tech. Note BN-787, Institute for Fluid Dynamics and Applied Mathematics, Univ. of Maryland, 1974; this Journal, 7 (1976), pp. 635–645.

[5] J. NEČAS, *Les méthodes directies en théorie des equations elliptiques*, Academia, Prague, 1967.
[6] M. I. VISIK, *Quasi-linear strongly elliptic systems of differential equations in divergence form*, Trans. Moscow Math. Soc., 12 (1963), pp. 140–209.
[7] S. G. MIKHLIN, *The Numerical Performance of Variational Methods*, Walters-Noordhoff, Groningen, the Netherlands, 1971.
[8] G. N. JAKOVLEV, *The first boundary value problem for quasilinear elliptic equations of second order*, Proc. Steklov Inst. Math., 117 (1972), pp. 381–403.

SIAM J. MATH. ANAL.
Vol. 8, No. 6, November 1977

A VOLTERRA EQUATION WITH A NONCONVOLUTION KERNEL*

T. R. KIFFE†

Abstract. This paper is concerned with the asymptotic behavior of solutions of the Volterra integral equation

$$x(t)+\int_0^t a(t,\tau)g(x(\tau))\,d\tau=f(t), \qquad 0\leqq t<\infty.$$

If $x(t)$ is a solution of this equation, the limiting values of $g(x(t))$ are given under various sets of hypotheses on the kernel $a(t,\tau)$ and the functions $g(t)$ and $f(t)$.

1. Introduction. In this paper we consider the asymptotic behavior of bounded solutions of the equation

$$x(t)+\int_0^t a(t,s)g(x(s))\,ds=f(t), \qquad 0\leqq t<\infty, \tag{1.1}$$

where a, g, and f are prescribed real-valued functions. Throughout we will assume that

$$g(x)\in C(-\infty,\infty), \tag{1.2}$$

$$f(t)\in C[0,\infty)\cap BV[0,\infty). \tag{1.3}$$

Closely related to (1.1) is the convolution equation

$$x(t)+\int_0^t b(t-s)g(x(s))\,ds=f(t), \qquad 0\leqq t<\infty, \tag{1.4}$$

which has been studied extensively by several authors.

With regard to (1.4), Levin [2] proved that if $b(t)\in C^1[0,\infty)$, $(-1)^k b^{(k)}(t)\geqq 0$ $(0\leqq t<\infty,\ k=0,1)$, $f(t)\in C^1[0,\infty)$, $\int_0^\infty |f'(t)|\,dt<\infty$, $g(x)\in C(-\infty,\infty)$, $g(0)=0$, and $g(x)$ is monotone nondecreasing, then (1.4) has a bounded solution. He also proved that if, in addition, we assume that $b(t)\in L^1(0,\infty)$, $b(t)$ is not constant on any interval except, possibly $b(t)\equiv 0$ on $T\leqq t<\infty$ for some T, $f'(t)\to 0$ as $t\to\infty$, $g(x)$ is monotone strictly increasing, and $x(t)$ is a solution of (1.4), then $x(\infty)=\lim_{t\to\infty} x(t)$ exists and satisfies

$$\lim_{t\to\infty}\left[x(t)+g(x(t))\int_0^\infty b(s)\,ds\right]=f(\infty). \tag{1.5}$$

Londen [8] generalized these results by showing that if (1.2) and (1.3) are satisfied, $b(t)\geqq 0$, $b'(t)\leqq 0$ and $b(t)\notin L^1(0,\infty)$, then for every bounded solution $x(t)$ of (1.4) we have $g(x(t))\to 0$ as $t\to\infty$. If $b(t)\in L^1(0,\infty)$ he proved that (1.5) holds for bounded solutions of (1.4). In a later paper [9], Londen replaced the

* Received by the editors August 14, 1975, and in revised form April 12, 1976.
† Department of Mathematics, Texas A & M University, College Station, Texas 77843.

hypothesis $b'(t) \leqq 0$ by $b(t)$ is nonincreasing without altering the above conclusions. This result was further improved by Londen [10] when he replaced (1.3) by the condition $f(t) \in L(0, \infty)$, $f(t) \to F$ as $t \to \infty$ and obtained a result similar to (1.5).

Several of the ideas for treating the nonconvolution equation (1.1) arose from an analysis of the asymptotic behavior of solutions of the differentiated equation

$$\begin{cases} x'(t) + \int_0^t a(t, s) g(x(s))\, ds = f(t), & 0 \leqq t < \infty, \\ x(0) = x_0. \end{cases} \tag{1.6}$$

This equation was first studied by Levin [3] and his results were improved by the author [1]. Several of the identities in the present paper, like (3.13), are closely related to similar ones which appeared in these two papers.

Volterra equations with nonconvolution kernels arise naturally in the area of viscoelasticity in the presence of chemical reactions. For more details see [7].

It is the purpose of this paper to study the asymptotic behavior of bounded solutions of (1.1). Specifically, this paper will be concerned with extending the results in [8] to nonconvolution kernels. This paper is a refinement of the author's Ph.D. thesis written at the University of Wisconsin–Madison under the direction of Professor J. A. Nohel.

2. Statement and discussion of results.

THEOREM 1. *Let* $R = \{(t, s) | 0 < t < \infty,\ 0 < s < t\}$ *and* $\bar{R} = \{(t, s) | 0 \leqq t < \infty,\ 0 \leqq s \leqq t\}$. *Suppose*

(i) $a(t, s) \in C(\bar{R})$, $a_t(t, s) \in C(\bar{R})$, $a(t, s) \geqq 0$, $a_t(t, s) \leqq 0$ *for* $(t, s) \in R$,
(ii) $\sup_{0 \leqq t < \infty} a(t, t) = M < \infty$,
(iii) $\int_{t-T}^t a_t(t, s)\, ds + a(t, t) - a(t, t - T) \geqq 0$ *for* $t \geqq T$,
(iv) *there is an* $\eta > 0$ *such that* $\lim_{t \to \infty} \inf \beta_\delta(t) > 0$ *for every* δ, $0 < \delta < \eta$, *where* $\beta_\delta(t) = \inf \{-\int_{\beta_1}^{\beta_2} a_t(t, s)\, ds \,|\, \beta_2 - \beta_1 \geqq \delta,\ t - \eta \leqq \beta_1 < \beta_2 \leqq t\}$,
(v) $\lim_{T \to \infty} \{\sup_{t \geqq T} \int_{t-T}^t a(t, s)\, ds\} = \infty$.
(vi) (1.2) *and* (1.3) *are satisfied,*
(vii) $x(t)$ *is a bounded solution of* (1.1) *for* $0 \leqq t < \infty$.

Then $g(x(t)) \to 0$ *as* $t \to \infty$.

In Theorem 1 we do not assume that every solution of (1.1) is bounded but we only treat the asymptotic behavior of bounded solutions. Hypothesis (i) is the obvious analogue of the hypothesis $b(t) \geqq 0$, $b'(t) \leqq 0$ in (1.4) and we use the notation $a_t(t, s) = \partial a(t, s)/\partial t$. For the convolution case $a(t, s) = b(t - s)$, (i) implies (ii) and (iii). Hypothesis (iv) is not required for (1.4) and (v) is the analogue of $b(t) \notin L'(0, \infty)$ in (1.4).

THEOREM 2. *Let* (i), (ii), (iii), (iv), (vi) *and* (vii) *but not* (v) *of Theorem* 1 *be satisfied. In addition, suppose*

(viii) $\lim_{T \to \infty} \{\sup_{t \geqq T} \int_{t-T}^t a(t, s)\, ds\} = A < \infty$.

Then we have

$$\lim_{t \to \infty} [x(t) + Ag(x(t))] = f(\infty). \tag{2.1}$$

The remarks following Theorem 1 also apply to Theorem 2. Hypothesis (viii) is the analogue of $b(t)\in L^1(0,\infty)$ in (1.4). If one tries to relax hypothesis (iii) of Theorems 1 and 2 the behavior of solutions of (1.1) becomes more complex than in the convolution case. Theorem 3 and especially Theorem 4 below give an indication of what can happen.

THEOREM 3. *Let* (i), (ii), (iv), (vi), (vii) *of Theorem* 1 *be satisfied. In addition, suppose*

(ix) $\int_0^t a_t(t,s)\,ds + a(t,t)\geqq 0$,
(x) $\lim_{T\to\infty}\lim_{t\to\infty}\int_0^{t-T} a(t,s)\,ds = B<\infty$,
(xi) $\int_0^t a(t,s)\,ds\to\infty$ *as* $t\to\infty$.

Then $g(x(t))\to 0$ *as* $t\to\infty$.

THEOREM 4. *Let* (i), (ii), (iv), (vi), *and* (vii) *of Theorem* 1 *and* (ix) *and* (x) *of Theorem* 3 *be satisfied. Suppose that*

(xii) $\int_0^t a(t,s)\,ds\to A<\infty$ *as* $t\to\infty$.

Then we have

$$\begin{aligned} -KB &\leqq \liminf_{t\to\infty}[x(t)+(A-B)g(x(t))-f(t)] \\ &\leqq \limsup_{t\to\infty}[x(t)+(A-B)g(x(t))-f(t)]\leqq KB \end{aligned} \tag{2.2}$$

where K *is given by* (3.1). *In particular, if* $B=0$, *then* (2.1) *is true.*

Examples of kernels $a(t,s)$ which satisfy the hypotheses of Theorems 1 and 2 are easy to construct. We present two such examples.

Example 1. If $a(t,s)=d(t)c(s)b(t-s)$ and $d(t)\in C^1[0,\infty)$, $c(s)\in C^2[0,\infty)$, $b(t)\in C^1[0,\infty)$,

$$\begin{aligned} &d(t)\geqq 0,\quad d'(t)\leqq 0,\quad b(t)\geqq 0,\quad b'(t)\leqq 0 \quad\text{and}\quad [d(t)c(t)]'\geqq 0 \\ &\text{for}\quad 0<t<\infty, \\ &c(s)\geqq 0,\quad c'(s)\geqq 0,\quad c''(s)\leqq 0 \quad\text{for}\quad 0<s<\infty, \\ &c(s)\not\equiv 0,\quad c(\infty)<\infty,\quad d(t)\not\equiv 0,\quad d(\infty)>0. \\ &b(t)\text{ is not constant on any interval except possibly an interval of the form} \\ &[T,\infty)\text{ where } T>0 \text{ and } b(t) \text{ is not identically equal to a constant,} \end{aligned} \tag{2.3}$$

then the hypotheses of Theorem 1 are satisfied if we add $b(t)\notin L^1(0,\infty)$ to (2.3) and the hypotheses of Theorem 2 are satisfied if we add $b(t)\in L^1(0,\infty)$ to (2.3).

Example 2. If $a(t,s)=b(d(t)(t-s))$ and if in addition to the hypotheses on $b(t)$ and $d(t)$ included in (2.3) we assume that

$$[td(t)]'\geqq 0,\qquad td'(t)\to 0 \quad as\ t\to\infty, \tag{2.4}$$

then the hypotheses of Theorem 1 are satisfied if $b(t)\notin L^1(0,\infty)$ and those of Theorem 2 are satisfied if $b(t)\in L^1(0,\infty)$. These examples will be discussed further in § 8.

For the sake of completeness we present a simple set of sufficient conditions under which (1.1) has a bounded solution.

THEOREM 5. *Let* (i), (vi), *and* (ix) *of Theorem* 3 *be satisfied. Suppose*

(xiii) $\inf_{-\infty<x<\infty} G(x)>-\infty$, $\lim_{|x|\to\infty} \sup G(x)=\infty$ *and there is a constant* $M>0$ *such that* $|g(x)|\leqq M[1+|G(x)|]$ *where* $G(x)=\int_0^x g(\xi)\,d\xi$,

(xiv) $f(t)$ *is absolutely continuous and* $\int_0^\infty |f'(t)|\,dt<\infty$.

Then (1.1) *has a bounded solution for* $0\leqq t\leqq\infty$ *and every solution of* (1.1) *is bounded.*

For a much deeper result concerning boundedness we refer to [5, Th. 3].

3. Proof of Theorem 1. By (vii) and (1.2), there is a constant $K<\infty$ such that

$$\sup_{0\leqq t<\infty} |g(x(t))|=K. \tag{3.1}$$

First we wish to establish

$$g(x(t)) \text{ is uniformly continuous for } 0\leqq t<\infty. \tag{3.2}$$

Let $y(t)=\int_0^t a(t,s)g(x(s))\,ds$, so (1.1) becomes

$$x(t)+y(t)=f(t). \tag{3.3}$$

By (i), $y(t)$ is locally absolutely continuous for $0\leqq t<\infty$ and we have

$$y'(t)=a(t,t)g(x(t))+\int_0^t a_t(t,s)g(x(s))\,ds \quad \text{a.e.} \tag{3.4}$$

To see that (3.4) is true, integrate the right side of (3.4), use Fubini's theorem and invoke the absolute continuity of $y(t)$. Also by (ii), (iii), (3.1) and (3.4) we have

$$|y'(t)|\leqq 2MK \quad \text{a.e.} \tag{3.5}$$

so $y(t)$ is uniformly continuous for $0\leqq t<\infty$. By (1.3), $f(t)$ is uniformly continuous, so by (vii), (1.2), and (3.3), (3.2) is established.

Next we wish to show that

$$\sup_{0\leqq T<\infty} \int_0^T a(T,s)[g(x(s))]^2\,ds<\infty \tag{3.6}$$

and

$$\sup_{0\leqq T<\infty} (-1)\int_0^T\int_0^t a_t(t,s)[g(x(t))-g(x(s))]^2\,ds\,dt<\infty. \tag{3.7}$$

Fix $T>0$. By (3.3) and the fact that $y(t)$ is of bounded variation on the interval $[0, T]$, we have

$$\int_0^T g(x(t))\,dx(t)+\int_0^T g(x(t))\,dy(t)=\int_0^T g(x(t))\,df(t). \tag{3.8}$$

Now $\int_0^T g(x(t))\,dx(t)=G(x(T))-G(x(0))$ where $G(x)=\int_0^x g(\xi)\,d\xi$ so, by (vii) and (1.2), there is a constant C, independent of T, such that

$$\left|\int_0^T g(x(t))\,dx(t)\right|\leqq C. \tag{3.9}$$

Also by (1.3) and (3.1) we have

$$\left| \int_0^T g(x(t))\, df(t) \right| \leqq K V_f[0, \infty) \tag{3.10}$$

where $V_f[0, \infty)$ denotes the total variation of $f(t)$. Since T was arbitrary, we have

$$\sup_{0 \leqq T < \infty} \left| \int_0^T g(x(t))\, dy(t) \right| < \infty. \tag{3.11}$$

By (3.4) we have

$$\begin{aligned} \int_0^T g(x(t))\, dy(t) &= \int_0^T a(t, t)[g(x(t))]^2\, dt \\ &\quad + \int_0^T \int_0^t a_t(t, s) g(x(t)) g(x(s))\, ds\, dt \end{aligned} \tag{3.12}$$

which becomes

$$\begin{aligned} \int_0^T g(x(t))\, dy(t) &= \frac{1}{2} \int_0^T a(t, t)[g(x(t))]^2\, dt \\ &\quad + \frac{1}{2} \int_0^T \int_0^t a_t(t, s)[g(x(t))]^2\, ds\, dt \\ &\quad + \frac{1}{2} \int_0^T a(T, s)[g(x(s))]^2\, ds \\ &\quad - \frac{1}{2} \int_0^T \int_0^t a_t(t, s)[g(x(t)) - g(x(s))]^2\, ds\, dt. \end{aligned} \tag{3.13}$$

To see that (3.13) is true, merely expand the last term in (3.13) and use Fubini's theorem. If we let

$$A(t) = \int_0^t a_t(t, s)\, ds + a(t, t) \tag{3.14}$$

we may rewrite (3.13) to get

$$\begin{aligned} \int_0^T g(x(t))\, dy(t) &= \frac{1}{2} \int_0^T A(t)[g(x(t))]^2\, dt \\ &\quad + \frac{1}{2} \int_0^T a(T, s)[g(x(s))]^2\, ds \\ &\quad - \frac{1}{2} \int_0^T \int_0^t a_t(t, s)[g(x(t)) - g(x(s))]^2\, ds\, dt. \end{aligned} \tag{3.15}$$

By using (i), (iii), and (3.11) in (3.15), (3.6) and (3.7) are established.

The following property of solutions of (1.1) will be crucial in the proof of Theorems 1, 2, 3, and 4 below. Suppose that there is a sequence $\{t_n\}$ and real numbers α and α_1 with $\alpha_1 > 0$ such that $g(x(t_n)) \geqq \alpha + \alpha_1$ for all n and $t_n \to \infty$ as

$n \to \infty$. We claim that there is a $\delta_1 > 0$ and a sequence of integers $\{I_m\}$ such that for each m

$$g(x(t)) \geqq \alpha \quad \text{for } t_n - \delta_1 - m\eta \leqq t \leqq t_n \quad \text{and} \quad n > I_m, \tag{3.16}$$

where $\eta > 0$ is given by (iv). To prove (3.16) we first observe that by (3.2) there is a $\delta_1 > 0$ such that

$$g(x(t)) \geqq \alpha + \frac{\alpha_1}{2} \quad \text{for } t_n - \delta_1 \leqq t \leqq t_n. \tag{3.17}$$

Next we claim that there is an integer I_1 such that

$$g(x(t)) \geqq \alpha + \frac{\alpha_1}{4} \quad \text{for } t_n - \delta_1 - \eta \leqq t \leqq t_n \quad \text{and} \quad n \geqq I_1. \tag{3.18}$$

Suppose (3.18) is not true. Then there is a subsequence $\{t_{n_i}\}$ of $\{t_n\}$, a sequence $\{\bar{t}_{n_i}\}$ and a $\gamma > 0$ such that

$$g(x(\bar{t}_{n_i})) \geqq \alpha + \frac{\alpha_1}{4} \quad \text{and} \quad t_{n_i} - \delta_1 - \eta \leqq \bar{t}_{n_i} \leqq t_{n_i} - \delta_1 - \gamma. \tag{3.19}$$

By (3.2), there is a $\delta_2 > 0$ such that

$$\delta_2 < \delta_1, \quad \delta_2 < \gamma \quad \text{and} \quad g(x(t)) < \alpha + \frac{7\alpha_1}{24} \quad \text{for } \bar{t}_{n_i} \leqq t \leqq \bar{t}_{n_i} + \delta_2. \tag{3.20}$$

Again by (3.2) there is a $\delta > 0$ such that

$$\delta < \delta_2, \quad \delta < \eta, \quad \delta < \gamma - \delta_2 \quad \text{and} \quad g(x(t)) \geqq \alpha + \frac{\alpha_1}{3} \quad \text{for } \bar{t}_{n_i} + \delta_2 \leqq t \leqq \bar{t}_{n_i} + \delta_2 + \delta. \tag{3.21}$$

Let $z_{n_i} = \bar{t}_{n_i} + \delta_2$ and $\mu_{n_i} = \min\{t_{n_i}, z_{n_i} + \eta\}$. Consider (3.7) and let $T = \mu_{n_I}$ for some integer I. Then

$$\begin{aligned} &(-1) \int_0^{\mu_{n_I}} \int_0^t a_t(t, s)[g(x(t)) - g(x(s))]^2 \, ds \, dt \\ &\geqq \frac{\alpha_1^2}{36} \sum_{i=1}^{I} \int_{t_{n_i} - \delta_1}^{\mu_{n_i}} (-1) \int_{z_{n_i}}^{z_{n_i} + \delta} a_t(t, s) \, ds \, dt \\ &\geqq \frac{\alpha_1^2}{36} \sum_{i=1}^{I} \int_{t_{n_i} - \delta_1}^{\mu_{n_i}} B_\delta(t) \, dt \end{aligned} \tag{3.22}$$

since, if $t_{n_i} - \delta_1 \leqq t \leqq \mu_{n_i}$, then by (3.20) and (3.21) we have $t - \eta \leqq z_{n_i} \leqq z_{n_i} + \delta < t$. Since $\mu_{n_i} - (t_{n_i} - \delta_1) \geqq \delta_2$ we have, by (iv)

$$\sum_{i=1}^{I} \int_{t_{n_i} - \delta_1}^{\mu_{n_i}} B_\delta(t) \, dt \geqq \delta_2 \sum_{i=1}^{I} \left(\inf_{t_{n_i} - \delta_1 \leqq t} B_\delta(t) \right) \to \infty \tag{3.23}$$

as $I \to \infty$. By (3.22) and (3.23) we have

$$(-1) \int_0^T \int_0^t a_t(t, s)[g(x(t)) - g(x(s))]^2 \, ds \, dt \to \infty \quad \text{as } T \to \infty$$

which contradicts (3.7). This establishes (3.18) and by repeating this argument we have that there is an integer I_2 such that $g(x(t)) \geqq \alpha + \alpha_1/8$ for $t_n - \delta_1 - 2\eta \leqq t \leqq t_n$ and $n \geqq I_2$. Proceeding in this way it is clear that we can construct a sequence $\{I_m\}$ so that (3.16) is satisfied. Similarly one can show that if there is a sequence $\{t_n\}$ and real numbers α and α_1 with $\alpha_1 > 0$ and $g(x(t_n)) \leqq \alpha - \alpha_1$ then there is a $\delta_1 > 0$ and a sequence of integers $\{I_m\}$ such that for each m, $g(x(t)) \leqq \alpha$ for $t_n - \delta_1 - m\eta \leqq t \leqq t_n$ and $n \geqq I_m$.

To complete the proof of Theorem 1, suppose that $g(x(t))$ does not converge to zero as $t \to \infty$. Then, for instance, there exist real numbers $\alpha > 0$ and $\alpha_1 > 0$ and a sequence $\{t_n\}$ such that $t_n \to \infty$ as $n \to \infty$ and $g(x(t_n)) \geqq \alpha + \alpha_1$ for all n. Let $S_m = \delta_1 + m\eta$ where δ_1 and η are given by (3.16). By (iii) $\int_{t-S_m}^{t} a(t, s)\, ds$ is an increasing function of t for each S_m. Hence, by (v) and (3.16), there is a sequence $\{T_m\}$ which is a subsequence of $\{t_n\}$ such that

$$\int_{T_m - S_m}^{T_m} a(t, s)\, ds \to \infty \quad \text{as } m \to \infty \quad \text{and} \quad g(x(t)) \geqq \alpha \quad \text{for } T_m - S_m \leqq t \leqq T_m.$$

Hence

$$\int_0^{T_m} a(T_m, s)[g(x(s))]^2\, ds \geqq \alpha^2 \int_{T_m - S_m}^{T_m} a(T_m, s)\, ds \to \infty \tag{3.24}$$

as $m \to \infty$, which contradicts (3.6). Hence $g(x(t)) \to 0$ as $t \to \infty$.

4. Proof of Theorem 2. Suppose (2.1) is not true. Then, for example, there is a $\delta > 0$ and a sequence $\{t_n\}$ such that $t_n \to \infty$ as $n \to \infty$ and

$$x(t_n) + Ag(x(t_n)) > f(\infty) + \delta \quad \text{for all } n. \tag{4.1}$$

We claim that there is a $\delta_1 < 0$ and a sequence of integers $\{I_m\}$ such that for each m

$$g(x(t)) \geqq g(x(t_n)) - \frac{\delta}{2A} \quad \text{for } t_n - \delta_1 - m\eta \leqq t \leqq t_n, \quad n \geqq I_m, \tag{4.2}$$

where η is given by (iv). We begin proving (4.2) by noticing that (3.1) implies that there is a subsequence $\{t_{n_i}\}$ of $\{t_n\}$ and a real number μ such that $g(x(t_{n_i})) \to \mu$ as $n_i \to \infty$. Without loss of generality we may assume this for the original sequence. Thus there is an α_1 such that $0 < \alpha_1 < \delta/(4A)$ and

$$\mu - \tfrac{1}{2}\alpha_1 \leqq g(x(t_n)) \leqq \mu + \tfrac{1}{2}\alpha_1 \quad \text{for all large } n. \tag{4.3}$$

If we let $\alpha = \mu - 3\alpha_1/2$, (4.3) becomes

$$\alpha + \alpha_1 \leqq g(x(t_n)) \leqq \alpha + 2\alpha_1 \quad \text{for large } n. \tag{4.4}$$

As in the proof of Theorem 1, (3.16) is still true, so (3.16) and (4.4) imply that there is a $\delta_1 > 0$ and a sequence of integers $\{I_m\}$ such that for each m,

$$\begin{aligned} g(x(t)) \geqq \alpha = \mu - \frac{3\alpha_1}{2} = \mu + \frac{1}{2}\alpha_1 - 2\alpha_1 \geqq g(x(t_n)) - 2\alpha_1 \\ \text{for } t_n - \delta_1 - m\eta \leqq t \leqq t_n \quad \text{and} \quad n \geqq I_m. \end{aligned} \tag{4.5}$$

Since $\alpha_1 < \delta/(4A)$, (4.2) is established.

Let $S_m = \delta_1 + m\eta$. By (iii) $\int_{t-S_m}^{t} a(t, s)\,ds$ is an increasing function of t for each S_m, so by (viii) and (4.2) there is a sequence $\{T_m\}$ which is a subsequence of $\{t_m\}$ such that

$$\int_{T_m - S_m}^{T_m} a(T_m, s)\,ds \to A \quad \text{as } m \to \infty \tag{4.6}$$

and

$$g(x(t)) \geqq g(x(T_m)) - \frac{\delta}{2A} \quad \text{for } T_m - S_m \leqq t \leqq T_m. \tag{4.7}$$

For these sequences $\{T_m\}$ and $\{S_m\}$ we have

$$\int_{T_m - S_m}^{T_m} a(T_m, s)\,ds \leqq \int_0^{T_m} a(T_m, s)\,ds \leqq \sup_{t \geqq T_m} \int_{t-T_m}^{t} a(t, s)\,ds \tag{4.8}$$

so by (viii) and (4.6) we have

$$\int_0^{T_m} a(T_m, s)\,ds \to A \quad \text{as } m \to \infty. \tag{4.9}$$

Since $\int_0^{T_m - S_m} a(T_m, s)\,ds = \int_0^{T_m} a(T_m, s)\,ds - \int_{T_m - S_m}^{T_m} a(T_m, s)\,ds$ we have

$$\int_0^{T_m - S_m} a(T_m, s)\,ds \to 0 \quad \text{as } m \to \infty. \tag{4.10}$$

Since

$$\begin{aligned} x(T_m) + Ag(x(T_m)) = &- \int_0^{T_m - S_m} a(T_m, s) g(x(s))\,ds \\ &- \int_{T_m - S_m}^{T_m} a(T_m, s) g(x(s))\,ds + Ag(x(T_m)) + f(T_m) \end{aligned}$$

we have, by (4.7),

$$\begin{aligned} x(T_m) + Ag(x(T_m)) \leqq &- \int_0^{T_m - S_m} a(T_m, s) g(x(s))\,ds \\ &- g(x(T_m)) \int_{T_m - S_m}^{T_m} a(T_m, s)\,ds \\ &+ \frac{\delta}{2A} \int_{T_m - S_m}^{T_m} a(T_m, s)\,ds + Ag(x(T_m)) + f(T_m). \end{aligned} \tag{4.11}$$

By (3.1) and (4.10) the first integral on the right hand side of (4.11) converges to zero as $m \to \infty$. By (4.6) we have

$$-g(x(T_m)) \int_{T_m - S_m}^{T_m} a(T_m, s)\,ds + Ag(x(T_m)) \to 0$$

and

$$\frac{\delta}{2A} \int_{T_m - S_m}^{T_m} a(T_m, s)\,ds \to \frac{\delta}{2} \quad \text{as } m \to \infty.$$

These facts together with (1.3) imply that

$$x(T_m)+Ag(x(T_m))\leqq f(\infty)+\delta \quad \text{for large } m \tag{4.12}$$

which contradicts (4.1). This establishes (2.1) and completes the proof of Theorem 2.

5. Proof of Theorem 3. Hypothesis (ix) is just enough to ensure that all the arguments in the proof of Theorem 1 up to and including (3.16) remain valid. Suppose that $g(x(t))$ does not converge to zero as $t\to\infty$. Then, for example, there is a sequence $\{t_n\}$ and real numbers α and α_1 wtih $\alpha>0$ and $\alpha_1>0$ such that $g(x(t_n))\geqq\alpha+\alpha_1$. As in the proof of Theorem 1, there are sequences $\{S_m\}$ and $\{T_m\}$ such that

$$\int_0^{T_m-S_m} a(T_m, s)\,ds\to B \quad \text{as } m\to\infty \tag{5.1}$$

and

$$g(x(t))\geqq\alpha \quad \text{for } T_m-S_m\leqq t\leqq T_m. \tag{5.2}$$

Since $\int_{T_m-S_m}^{T_m} a(T_m, s)\,ds=\int_0^{T_m} a(T_m, s)\,ds-\int_0^{T_m-S_m} a(T_m, s)\,ds$, we have, by (xi) and (5.1),

$$\int_{T_m-S_m}^{T_m} a(T_m, s)\,ds\to\infty \quad \text{as } m\to\infty. \tag{5.3}$$

Since (3.6) is still true, (5.2) and (5.3) lead to a contradiction as in the proof of Theorem 1.

6. Proof of Theorem 4. If $A=B$, (2.2) follows immediately from (1.1) and (xii). Thus let us consider the case $B<A$. We will establish the last inequality in (2.2). The proof of the first inequality is similar. Suppose it is not true that $\limsup_{t\to\infty}[x(t)+(A-B)g(x(t))-f(t)]\leqq KB$. Then there is a $\delta>0$ and a sequence $\{t_n\}$ such that $t_n\to\infty$ as $n\to\infty$ and

$$x(t_n)+(A-B)g(x(t_n))-f(t_n)>KB+\delta \quad \text{for all } n. \tag{6.1}$$

As in the proof of Theorem 2 we have that there is a $\delta_1>0$ and a sequence of integers $\{I_m\}$ such that for each m

$$g(x(t))\geqq g(x(t_n))-\frac{\delta}{2(A-B)} \quad \text{for } t_n-\delta_1-m\eta\leqq t\leqq t_n,\quad n\geqq I_m. \tag{6.2}$$

Hence by (x) there exist sequences $\{S_m\}$ and $\{T_m\}$ where $S_m=\delta_1+m\eta$ and $\{T_m\}$ is a subsequence of $\{t_n\}$ such that

$$\int_0^{T_m-S_m} a(T_m, s)\,ds\to B \quad \text{as } m\to\infty \tag{6.3}$$

and

$$g(x(t))\geqq g(x(T_m))-\frac{\delta}{2(A-B)} \quad \text{for } T_m-S_m\leqq t\leqq T_m. \tag{6.4}$$

By (xii) and (6.3) we have

$$\int_{T_m-S_m}^{T_m} a(T_m, s)\,ds \to A-B \quad \text{as } m\to\infty. \tag{6.5}$$

By (1.1) and (6.4) we have

$$\begin{aligned} x(T_m)+(A-B)g(x(T_m))-f(T_m) \leqq & -\int_0^{T_m-S_m} a(T_m, s)g(x(s))\,ds \\ & -g(x(T_m))\int_{T_m-S_m}^{T_m} a(T_m, s)\,ds \\ & +\frac{\delta}{2(A-B)}\int_{T_m-S_m}^{T_m} a(T_m, s)\,ds \\ & +(A-B)g(x(T_m)). \end{aligned} \tag{6.6}$$

By (3.1) and (6.3) we have $\limsup_{m\to\infty}(-1)\int_0^{T_m-S_m} a(T_m, s)g(x(s))\,ds \leqq KB$. By (6.5) we have $-g(x(T_m))\int_{T_m-S_m}^{T_m} a(T_m, s)\,ds+(A-B)g(x(T_m))\to 0$ and $[\delta/(2(A-B))]\int_{T_m-S_m}^{T_m} a(T_m, s)\,ds \to \delta/2$ as $m\to\infty$. Thus for large m we have $x(T_m)+(A-B)g(x(T_m))-f(T_m)\leqq KB+\delta$ which contradicts (6.1).

7. Proof of Theorem 5. By a well-known result, (1.1) has a solution $x(t)$ for $0\leqq t\leqq T$ for some $T>0$. Using the notation of Theorem 1, we have, by (3.8)

$$G(x(T))-G(x(0))+\int_0^T g(x(t))\,dy(t)=\int_0^T g(x(t))\,df(t). \tag{7.1}$$

By (i), (ix), (3.14) and (3.15) we have $\int_0^T g(x(t))\,dy(t)\geqq 0$, so by (xiv)

$$G(x(T))-G(x(0))\leqq \int_0^T |g(x(t))|\,|f'(t)|\,dt. \tag{7.2}$$

By (xiii) there is a constant M_1 such that

$$\begin{aligned} |G(x(T))| \leqq & M_1+|G(x(0))|+\int_0^\infty |f'(t)|\,dt \\ & +M\int_0^T |G(x(t))|\,|f'(t)|\,dt. \end{aligned} \tag{7.3}$$

By Gronvall's inequality, there is a constant C, independent of T, such that $|G(x(T))|\leqq C$. By (xiii) there is a constant C_1, again independent of T_1 such that

$$|x(T)|\leqq C_1. \tag{7.4}$$

The usual continuation argument for Volterra equations gives us a solution $x(t)$ of (1.1) for $0\leqq t<\infty$ and $|x(t)|\leqq C_1$ for $0\leqq t<\infty$.

8. Discussion of examples. For Example 1 the verification that the hypotheses of Theorems 1 and 2 are satisfied is fairly straightforward. The verification of

(iii) rests on the fact that if $a_t(t, s)+a_s(t, s)\geqq 0$, then

$$\int_{t-T}^{t} a_t(t, s)\,ds + a(t, t) - a(t, t-T)$$

$$\geqq -\int_{t-T}^{t} a_s(t, s)\,ds + a(t, t) - a(t, t-T) = 0.$$

Since $a_t(t, s) = d'(t)c(s)b(t-s)+d(t)c(s)b'(t-s)$ and $a_s(t, s) = d(t)c'(s)b(t-s)-d(t)c(s)b'(t-s)$ we want $d'(t)c(s)+d(t)c'(s)\geqq 0$ for $0<s<t$. Since $c(s)\geqq 0$, $c'(s)\geqq 0$, and $c''(s)\leqq 0$ we have $d'(t)c(s)+d(t)c'(s)\geqq d'(t)c(t)+d(t)c'(t)$, $0<s<t$; so $[d(t)c(t)]'\geqq 0$ implies $a_t(t, s)+a_s(t, s)\geqq 0$.

For Example 2 we only need to comment about hypothesis (iv). Note that $a_t(t, s) = b'(d(t)(t-s))[d'(t)(t-s)+d(t)]$. Thus

$$-\int_{B_1}^{B_2} a_t(t, s)\,ds = -\int_{B_1}^{B_2} b'(d(t)(t-s))d(t)\,ds - \int_{B_1}^{B_2} b'(d(t)(t-s))d'(t)(t-s)\,ds \tag{8.1}$$

and

$$\int_{B_1}^{B_2} b'(d(t)(t-s))d'(t)(t-s)\,ds \leqq \frac{td'(t)}{d(t)} \int_{B_1}^{B_2} b'(d(t)(t-s))d(t)\,ds.$$

By (2.4) this last term converges to zero as $t\to\infty$. Since $b(t)$ is not constant on any interval except possibly those of the form $[T, \infty)$ where $T>0$ and $b(t)$ is not identically equal to a constant, it is easy to see that if η is chosen small enough, then for every δ, $0<\delta<\eta$, there is a constant c_δ such that $-\int_{B_1}^{B_2} b'(d(t)(t-s))d(t)\,ds \geqq c_\delta > 0$ for $B_2-B_1\geqq\delta$, $t-\eta\leqq B_1<B_2\leqq t$ and t sufficiently large. Thus the same is true for $-\int_{B_1}^{B_2} a_t(t, s)\,ds$.

REFERENCES

[1] T. R. KIFFE, *On nonlinear Volterra equations of nonconvolution type*, J. Differential Equations, 22 (1976), pp. 349–367.

[2] J. J. LEVIN, *The qualitative behavior of a nonlinear Volterra equation*, Proc. Amer. Math. Soc., 16 (1965), pp. 711–718.

[3] ———, *A nonlinear Volterra equation not of convolution type*, J. Differential Equations, 4 (1968), pp. 176–186.

[4] ———, *On a nonlinear Volterra equation*, J. Math. Anal. Appl., 39 (1972), pp. 458–476.

[5] ———, *Remarks on a Volterra equation*, Delay and Functional Differential Equations and Their Applications, Academic Press, New York and London, 1972, pp. 233–255.

[6] ———, *A bound on the solutions of a Volterra equation*, Arch. Rational Mech. Anal., 52 (1973), pp. 339–349.

[7] A. S. LODGE, *Body Tensor Fields in Continuum Mechanics*, Academic Press, New York and London, 1974.

[8] S.-O. LONDEN, *On the solutions of a nonlinear Volterra equation*, J. Math. Anal. Appl., 39 (1972), pp. 564–573.

[9] ———, *On a nonlinear Volterra integral equation*, J. Differential Equations, 14 (1973), pp. 106–120.

[10] ———, *On the asymptotic behavior of the bounded solutions of a nonlinear Volterra equation*, this Journal, 5 (1974), pp. 849–875.

SIAM J. MATH. ANAL.
Vol. 8, No. 6, November 1977

ON AN INTEGRAL EQUATION IN A HILBERT SPACE*

STIG-OLOF LONDEN†

Abstract. We consider the nonlinear Volterra equation

$$u(t)+\int_0^t a(t-\tau)g(u(\tau))\,d\tau \ni f(t), \qquad t\geqq 0, \tag{1.1}$$

where a, g, f are given and u is the unknown function taking values in a real Hilbert space H. The kernel $a(t)$ maps $R^+\to R$ whereas f is a map of $R^+\to H$. The nonlinear function g has its domain and range contained in H.

Making use of the theory of monotone operators we give at first an existence and uniqueness theorem on (1.1). This is followed by a result detailing the asymptotic behavior of solutions of (1.1). Finally we give some applications of our results. The results extend earlier results by Barbu.

1. Introduction.

We consider the nonlinear Volterra equation

$$u(t)+\int_0^t a(t-\tau)g(u(\tau))\,d\tau \ni f(t), \qquad t\geqq 0, \tag{1.1}$$

where a, g, f are given and u is the unknown taking values in a real Hilbert space H. The kernel $a(t)$ is real-valued and defined on R_+, whereas f maps R_+ into H. The nonlinear function g (in general multivalued) has its domain Dg and range Rg contained in H. The integral in (1.1) is to be considered as a Bochner integral.

A solution of (1.1) on an interval $[0, T]$ is a function $u(t)$ defined on $[0, T]$, taking values in H, and satisfying

$$u\in L_2(0, T; H), \tag{1.2}$$

$$u(t)\in Dg \quad \text{a.e. on } (0, T), \tag{1.3}$$

and such that there exists w satisfying

$$w\in L_2(0, T; H), \tag{1.4}$$

$$w(t)\in g(u(t)) \quad \text{a.e. on } (0, T), \tag{1.5}$$

$$u(t)+\int_0^t a(t-\tau)w(\tau)\,d\tau = f(t), \qquad 0\leqq t\leqq T. \tag{1.6}$$

A solution of (1.1) on $[0, \infty)$ is a function $u(t)$ defined for all $t\geqq 0$, taking values in H, and satisfying (1.2)–(1.6) for every $T<\infty$.

This work is structured as follows. We begin § 2 by stating an existence and uniqueness result which is followed by some comments including comparisons to earlier studies. We then formulate Theorem 2 which deals with boundedness and asymptotic behavior of solutions of (1.1). After a few remarks related to Theorem 2 we give some corollaries providing partial extensions of our results. The proofs

* Received by the editors October 31, 1975, and in revised form June 20, 1976.

† Institute of Mathematics, Helsinki University of Technology, SF-02150, Otaniemi, Finland.

of Theorems 1 and 2 are given in §§ 3 and 4 respectively. Corollaries 1 and 2 are proved in § 5. Finally, in § 6, we give some applications of the results.

Our approach when analyzing (1.1) in H relies on combining the properties of maximal monotone operators, see [3], with the use of certain techniques developed in [5], [6] for scalar Volterra equations. The notation used is the usual one; thus g_λ for example denotes the Yosida-approximation of g, that is $g_\lambda = \lambda^{-1}(I - j_\lambda)$ with $j_\lambda = (I + \lambda g)^{-1}$.

2. Statement of results.

THEOREM 1. (a) *Suppose*

(2.1) $a(0) > 0$,

(2.2) $g = \partial\varphi$ *for some lower semicontinuous proper convex function* $\varphi: H \to (-\infty, \infty]$,

(2.3) $f(0) \in D(\varphi)$,

and let, for some $T > 0$,

(2.4) $a(t), f(t)$ *be absolutely continuous on* $[0, T]$ *and such that*

(2.5) $a' \in BV[0, T]$,

(2.6) $f' \in L_2(0, T; H)$.

Then there exists a unique solution u *of* (1.1) *on* $[0, T]$.

(b) *Assume* (2.1)–(2.5) *hold. In addition let*

(2.7) $a(t), f(t)$ *be locally absolutely continuous on* $[0, \infty)$,

(2.8) $f' \in L_2^{\text{loc}}(0, \infty; H)$.

Then there exists a unique solution u *of* (1.1) *on* $[0, \infty)$.

The solution $u(t)$ of course by definition satisfies (1.2)–(1.6). But observe in addition (see (3.63)–(3.67)) that $u(t)$ and $\varphi(u(t))$ are locally absolutely continuous and

$$u'(t) + a(0)w(t) + \int_0^t a'(t-\tau)w(\tau)\,d\tau = f'(t) \tag{2.9}$$

a.e. on the interval of existence. Thus, Theorem 1 in fact asserts the existence of a strong solution of (2.9). If $a'(t) \equiv 0$, then (2.9) reduces to

$$u'(t) + a(0)w(t) = f'(t), \tag{2.10}$$

for which existence and uniqueness of a strong solution are well established under the present hypotheses. See [3, Chap. III].

The motivation for this work came partly from recent results by Barbu [1] who considered (1.1) in the same setting. Comparing our Theorem 1 to Barbu's first theorem we observe that in the latter $a(t)$ is assumed to satisfy

(2.11) $a(t)$ continuous on $[0, \infty)$ and locally absolutely continuous on $(0, \infty)$,

(2.12) $(-1)^k a^{(k)}(t) \geqq 0, \quad k = 0, 1$ a.e. on $t > 0$,

$$\operatorname{Re} \hat{a}(\lambda) > 0, \qquad \operatorname{Re} \lambda > 0, \tag{2.13}$$

where

$$\hat{a}(\lambda) = \int_0^\infty \exp(-\lambda t) a(t)\, dt. \tag{2.14}$$

Thus $a(t)$ is required to be a kernel of positive type before existence can be obtained. Obviously this hypothesis has been abandoned in the present Theorem 1.

The assumptions on g, f in our Theorem 1 and in the existence part of Theorem 1 of [1] are identical.

Note that we obtain uniquenesss without imposing strict monotonicity on g (which is done in [1]).

Finally it should be pointed out that Barbu obtains some existence results (i.e. Theorems 3 and 4 of [1]) which do not require g to be the subdifferential of a convex function. These results do however require other conditions to be satisfied. Theorem 3 of [1] for example requires, for some positive constants c_1, c_2

$$\|g(u)\|_{v'} \leq c_1 \|u\|_v, \quad (g(u), u)_{v'v} \geq c_2 \|u\|_v^2, \quad u \in V,$$

with $g\colon V \to V'$, V a Hilbert space and $V \subset H \subset V'$.

THEOREM 2. *Let* (2.1), (2.7), (2.8) *hold. In addition assume*

$$a(t) \geq 0, \qquad t \geq 0, \tag{2.15}$$

$$a'(t) \leq 0 \quad \textit{a.e. on } t \geq 0, \tag{2.16}$$

$$a \in L_1(0, \infty), \tag{2.17}$$

$$\sum_{n=0}^{\infty} \left[\int_n^{n+1} \|f'\|^2 \, d\tau \right]^{1/2} < \infty. \tag{2.18}$$

Let $u(t)$ be a solution of (1.1) *on* $[0, \infty)$ *and suppose there exists a locally absolutely continuous function $\varphi(t)$ such that*

$$\varphi'(t) = \langle w(t), u'(t) \rangle \quad \textit{a.e. on } t > 0, \tag{2.19}$$

and such that for some $q < 2$

$$\varphi(t) \geqq -\alpha \|u(t)\|^q - \beta, \tag{2.20}$$

for all $t \geqq 0$, and some constants α, β. Then

$$\sup_{t \in R^+} \int_t^{t+1} \|w(\tau)\|^2 \, d\tau < \infty, \tag{2.21}$$

$$\sup_{t \in R^+} \|u(t)\| < \infty, \tag{2.22}$$

$$u' \in L_2(0, \infty; H). \tag{2.23}$$

Observe that (2.19), (2.21), (2.23) together imply that φ varies "slower and slower" when $t \to \infty$.

Commenting on the hypothesis of Theorem 2 we observe at first that once existence of u is assumed then the monotonicity of g may be dropped; only the existence of a function φ satisfying (2.19), (2.20) is needed. Note that the condition (2.2) and the existence of a function $u(t)$ satisfying (1.2)–(1.5) together imply that (2.19), (2.20) hold; the latter has $q = 1$. (Recall that the asymptotic result of Barbu, that is Theorem 2 of [1], does assume (2.2).) Also observe that we do not, in order to obtain (2.21)–(2.23) impose anything like $\varphi(u) \to +\infty$ for $|u| \to \infty$.

Secondly, we note that (2.15), (2.16) are customary assumptions made when analyzing (1.1) with $H = R$. Observe, however, that the strongest results obtained in this scalar case require less smoothness on the kernel; that is only $a \geqq 0$, a nonincreasing. See for example [4], [6]. It is very likely possible to weaken (2.15), (2.16) in this direction but $w(t)$ must then be taken continuous. Note, however, Corollaries 3 and 4 below.

In Theorem 2 we assume $a \in L_1(0, \infty)$. This allows us to impose only a rather weak hypothesis on the behavior of φ, namely (2.20). The case when $a \notin L_1(0, \infty)$, $a(\infty) = 0$, is included in Corollary 1 where instead we have to take φ bounded from below.

The assumption (2.18) is somewhat stronger than the corresponding one, which is $f' \in L_1(0, \infty)$, usually made when working with the scalar version of (1.1). The condition (2.18) constitutes the price we pay for working in H, and in particular for not using the unrealistic condition $\|g(u(t))\| \leqq K[1 + \varphi(u(t))]$, (compare with [8]).

Theorem 2 and Corollary 1, when applied to the case $H = R$, extend results obtained in [4], [6]. This follows from the fact that only (2.20), (2.24) are assumed about the asymptotic behavior of φ.

The method used in the final part of the proof of Theorem 2 can also be used to provide a shorter and different demonstration of previously known results about (1.1) when $H = R$. For details, see [7].

Finally, note that our Corollary 2 is essentially equivalent to the boundedness and asymptotic results of [1]. The key additional assumption made in this corollary is $a(\infty) > 0$ which greatly simplifies all proofs. Also note (compare with (2.15) of [1]) that this additional assumption is a key ingredient in permitting one to deduce that $\varphi_\infty = \lim_{t \to \infty} \varphi(t)$ exists. This last fact could not be deduced under the hypothesis of Theorem 2. If also the assumption (2.2) is made then it follows after using the definition of the subdifferential of a convex function that $\varphi_\infty = \min_{v \in H} \varphi(v)$. As a final point of comparison with [1] we note that the assumption (2.2) and the requirement that $\varphi(u) \to +\infty$ for $|u| \to \infty$ (these hypotheses are both made in Theorem 2 of [1]) together imply (2.24).

COROLLARY 1. *Let* (2.1), (2.7), (2.8), (2.15), (2.16), (2.18), (2.19) *hold. Also suppose*

$$\varphi(t) \geqq -\beta, \qquad t \geqq 0, \tag{2.24}$$

for some constant β. *Then* (2.21), (2.23) *hold.*

COROLLARY 2. *Let* (2.1), (2.7), (2.8), (2.15), (2.16), (2.19), (2.24) *hold. In addition assume*

$$\lim_{t\to\infty} a(t) = a(\infty) > 0, \tag{2.25}$$

$$f' \in L_2(0, \infty; H). \tag{2.26}$$

Then (2.23) *and*

$$w \in L_2(0, \infty; H), \tag{2.27}$$

are satisfied. Moreover, $\varphi' \in L_1(0, \infty)$ *and so* $\lim_{t\to\infty} \varphi(t)$ *exists.*

Our final results, which deal with (2.28), show that a step function, with a finite number of jumps, may be added to the kernel without affecting the results. In (2.28) we take c_k, T_k to be constants, $k = 1, 2, \cdots, N < \infty$, $T_k > 0$.

$$u(t) + \int_0^t a(t-\tau) g(u(\tau))\, d\tau + \sum_{k=1}^{N} c_k \int_{\max(0,t-T_k)}^{t} g(u(\tau))\, d\tau \ni f(t), \qquad t \geqq 0. \tag{2.28}$$

COROLLARY 3. *Let* (2.1)–(2.8) *hold. Then there exists a unique solution of* (2.28) *on* $[0, \infty)$.

COROLLARY 4. *Let the hypothesis of Theorem* 2 *hold, except that u is assumed to be a solution of* (2.28). *In addition let* $c_k > 0$, $k = 1, 2, \cdots, N$. *Then* (2.21)–(2.23) *are satisfied.*

The proofs of Corollaries 3 and 4 are omitted, as being straightforward extensions of those of Theorems 1 and 2.

3. Proof of Theorem 1. We begin by proving existence on $[0, T]$. This will last until (3.72).

Let $u_\lambda(t)$, $\lambda > 0$, be the unique solution of

$$u_\lambda(t) + \int_0^t a(t-\tau) g_\lambda u_\lambda(\tau)\, d\tau = f(t), \qquad 0 \leqq t \leqq T. \tag{3.1}$$

Then

$$u_\lambda'(t) + a(0) g_\lambda u_\lambda(t) + \int_0^t a'(t-\tau) g_\lambda u_\lambda(\tau)\, d\tau = f'(t), \tag{3.2}$$

a.e. on $(0, T)$. Form the scalar product of $g_\lambda u_\lambda$ and (3.2), and integrate over $(0, t)$. This gives

$$\varphi_\lambda(u_\lambda(t)) - \varphi_\lambda(f(0)) + a(0) \int_0^t \|g_\lambda u_\lambda\|^2\, d\tau + \int_0^t \left\langle g_\lambda u_\lambda(\tau), \int_0^\tau a'(\tau - s) g_\lambda u_\lambda(s)\, ds \right\rangle d\tau = \int_0^t \langle g_\lambda u_\lambda, f' \rangle\, d\tau. \tag{3.3}$$

Let T_1 be arbitrary fixed but satisfying

$$0<T_1\leqq T, \qquad 2\int_0^{T_1}|a'(\tau)|\,d\tau\leqq a(0). \tag{3.4}$$

After these preliminaries our goal is to obtain existence on $[0, T_1]$. This will occupy us until (3.60). To accomplish the goal we begin by establishing certain bounds on $g_\lambda u_\lambda$ and u_λ, namely (3.10), (3.11).

Let $A^2=\int_0^{T_1}a^2(\tau)\,d\tau$. Then, from (3.1), after estimating and using the absolute continuity of f, we obtain

$$\|u_\lambda(t)\|\leqq A\left[\int_0^t\|g_\lambda u_\lambda\|^2\,d\tau\right]^{1/2}+F, \tag{3.5}$$

for $0\leqq t\leqq T_1$ if $F=\sup\|f(t)\|$, $0\leqq t\leqq T_1$. Also note that by (2.2)

$$\varphi_\lambda(u_\lambda(t))\geqq-\alpha\|u_\lambda(t)\|-\beta, \tag{3.6}$$

for some constants α, β, and so, combining (3.5), (3.6), we have

$$-\varphi_\lambda(u_\lambda(t))\leqq\alpha A\left[\int_0^t\|g_\lambda u_\lambda\|^2\,d\tau\right]^{1/2}+\alpha F+\beta, \qquad 0\leqq t\leqq T_1. \tag{3.7}$$

Estimating the last term on the left side of (3.3) gives by (3.4)

$$\begin{aligned}&\left|\int_0^t\left\langle g_\lambda u_\lambda(\tau),\int_0^\tau a'(\tau-s)g_\lambda u_\lambda(s)\,ds\right\rangle d\tau\right|\\&\quad\leqq\int_0^t\|g_\lambda u_\lambda\|\int_0^\tau|a'|\|g_\lambda u_\lambda\|\,ds\,d\tau\\&\quad\leqq\left[\int_0^t|a'(s)|\,ds\right]\left[\int_0^t\|g_\lambda u_\lambda\|^2\,d\tau\right]\\&\quad\leqq 2^{-1}a(0)\int_0^t\|g_\lambda u_\lambda\|^2\,d\tau.\end{aligned} \tag{3.8}$$

Using now (2.6), (3.7), (3.8) in (3.3) yields, $0\leqq t\leqq T_1$,

$$\begin{aligned}&2^{-1}a(0)\int_0^t\|g_\lambda u_\lambda(\tau)\|^2\,d\tau\\&\quad\leqq[\alpha A+F_1]\left[\int_0^t\|g_\lambda u_\lambda(\tau)\|^2\,d\tau\right]^{1/2}+\alpha F+\beta+\varphi(f(0)),\end{aligned} \tag{3.9}$$

where $F_1=\left[\int_0^{T_1}\|f'(\tau)\|^2\,d\tau\right]^{1/2}$. But (3.9) clearly implies

$$\sup_{\lambda>0}\left[\int_0^{T_1}\|g_\lambda u_\lambda(\tau)\|^2\,d\tau\right]^{1/2}\overset{\text{def}}{=}c<\infty. \tag{3.10}$$

Combining (3.5), (3.10) gives

$$\sup_{\substack{\lambda>0\\0\leqq t\leqq T_1}}\|u_\lambda(t)\|\leqq cA+F<\infty. \tag{3.11}$$

Observe that T_1 is restricted only by (3.4).

Having the bounds (3.10), (3.11) our next purpose is to show that

$$\lim_{\lambda,\mu\to 0}\left\|\int_0^t [g_\lambda u_\lambda - g_\mu u_\mu]\,d\tau\right\| = 0, \qquad 0\leqq t\leqq T_1. \tag{3.12}$$

Note that once we have (3.12) then it is not hard to show that

$$\lim_{\lambda,\mu\to 0}\sup_{0\leqq t\leqq T_1}\|u_\lambda(t)-\mu_\mu(t)\| = 0, \tag{3.13}$$

which is the crucial fact (in addition to (3.10), (3.11)) needed to get existence on $[0, T_1]$.

In order to obtain (3.12) we prove at first a slightly weaker assertion, namely the following.

Take any fixed γ satisfying $0<\gamma\leqq T_1$ and such that

$$a(s)-16\int_0^s\int_0^\tau |da'(v)|\,d\tau - 20\int_0^s |a'(\tau)|\,d\tau \geqq \eta > 0, \tag{3.14}$$

on $0\leqq s\leqq\gamma$, for some constant η. By (2.1), (2.5) such γ, η exist. We assert that

$$\lim_{\lambda,\mu\to 0}\left\|\int_0^t [g_\lambda u_\lambda - g_\mu u_\mu]\,d\tau\right\| = 0, \qquad 0\leqq t\leqq\gamma. \tag{3.15}$$

The demonstration of (3.15) which follows (and which is the key part of the proof of Theorem 1) will occupy us up to and including the paragraph containing (3.37).

Suppose the assertion (3.15) does not hold. Then there exist $\hat{t}$, δ, $0<\hat{t}\leqq\gamma$, $\delta>0$, $\lambda_n\downarrow 0$, $\mu_n\downarrow 0$, such that

$$\lim_{n\to\infty}\left\|\int_0^{\hat{t}} [g_{\lambda_n}u_{\lambda_n} - g_{\mu_n}u_{\mu_n}]\,d\tau\right\| = 2\delta > 0. \tag{3.16}$$

Take any such $\hat{t}$, δ, $\{\lambda_n\}$, $\{\mu_n\}$. Then note that for s_1, s_2 arbitrary but satisfying

$$0\leqq s_1<s_2\leqq T_1, \qquad 4c^2[s_2-s_1]\leqq\delta^2, \tag{3.17}$$

and for arbitrary λ, μ one has, by (3.10),

$$\begin{aligned}\left\|\int_{s_1}^{s_2} [g_\lambda u_\lambda - g_\mu u_\mu]\,d\tau\right\| &\leqq \left[\int_{s_1}^{s_2}\|g_\lambda u_\lambda - g_\mu u_\mu\|^2\,d\tau\right]^{1/2}[s_2-s_1]^{1/2}\\ &\leqq 2c[s_2-s_1]^{1/2}\leqq\delta.\end{aligned} \tag{3.18}$$

Divide the interval $[0, \hat{t}]$ into subintervals by points $t_0, t_1, \cdots, t_R$ so that

$$0=t_0<t_1<\cdots<t_{R-1}<t_R=\hat{t}, \quad 4c^2[t_{i+1}-t_i]\leqq\delta^2; \quad i=0,1,\cdots,R-1. \tag{3.19}$$

Then take the smallest i-value (call it m) for which there exists an infinite subsequence $n_K\subset n$ (let $n_K=n$) such that

$$\left\|\int_0^{t_i} [g_{\lambda_n}u_{\lambda_n} - g_{\mu_n}u_{\mu_n}]\,d\tau\right\| \geqq \delta, \qquad n=1,2,\cdots. \tag{3.20}$$

Define $h_n = g_{\lambda_n}u_{\lambda_n} - g_{\mu_n}u_{\mu_n}$, and take N sufficiently large so that

$$\left\|\int_0^{t_j} h_n(\tau)\,d\tau\right\| \leqq \delta, \qquad j=0,1,\cdots,m-1; \quad n\geqq N. \tag{3.21}$$

For any $t \in [0, t_m]$, let $\tilde{t} \overset{\text{def}}{=} \max\{t_j | j = 0, 1, \cdots, m-1; t_j \leqq t\}$. Then, by (3.17)–(3.19)

$$\left\| \int_{\tilde{t}}^{t} h_n(\tau)\, d\tau \right\| \leqq \delta. \tag{3.22}$$

Consequently, for any $t \in [0, t_m]$,

$$\left\| \int_{0}^{t} h_n(\tau)\, d\tau \right\| \leqq \delta + \left\| \int_{0}^{\tilde{t}} h_n(\tau)\, d\tau \right\| \leqq 2\delta, \tag{3.23}$$

(the second inequality follows from (3.21) and the way t was defined) and for arbitrary $s_1, s_2 \in [0, t_m]$

$$\left\| \int_{s_1}^{s_2} h_n(\tau)\, d\tau \right\| \leqq \left\| \int_{0}^{s_2} \right\| + \left\| \int_{0}^{s_1} \right\| \leqq 4\delta. \tag{3.24}$$

In the sequel we need the following key

LEMMA. *Let $a(t)$ be absolutely continuous on $[0, T]$ and satisfy $a' \in BV[0, T]$. Assume*

$$z \in L_2(0, T; H), \tag{3.25}$$

and take $t \in [0, T]$. Then

$$\begin{aligned} &\int_0^t \left\langle z(\tau), \int_0^\tau a(\tau - s) z(s)\, ds \right\rangle d\tau \\ &= \frac{a(t)}{2} \left\| \int_0^t z(\tau)\, d\tau \right\|^2 + \frac{1}{2} \int_0^t \int_0^\tau \left\| \int_{\tau - s}^{\tau} z(v)\, dv \right\|^2 da'(s)\, d\tau \\ &\quad - \frac{1}{2} \int_0^t a'(\tau) \left\| \int_0^\tau z(s)\, ds \right\|^2 d\tau - \frac{1}{2} \int_0^t a'(t - \tau) \left\| \int_\tau^t z(s)\, ds \right\|^2 d\tau. \end{aligned} \tag{3.26}$$

Proof of Lemma. We begin by demonstrating

$$\begin{aligned} &\frac{1}{2} \int_0^t a'(t - \tau) \left\| \int_\tau^t z(s)\, ds \right\|^2 d\tau \\ &= \int_0^t \int_0^\tau a'(\tau - s) \left\langle \int_s^\tau z(v)\, dv, z(\tau) \right\rangle ds\, d\tau \\ &\quad + \frac{1}{2} \int_0^t \int_0^\tau \left\| \int_{\tau - s}^{\tau} z(v)\, dv \right\|^2 da'(s)\, d\tau. \end{aligned} \tag{3.27}$$

Let $\{a_n'(t)\}$ be a sequence of functions satisfying (by the assumptions on $a(t)$ such a sequence clearly exists)

$$a_n' \in C'[0, t], \tag{3.28}$$

$$a_n'(s) \to a'(s), \qquad 0 \leqq s \leqq t, \tag{3.29}$$

$$\operatorname{Var}(a_n'; [0, t]) \leqq \operatorname{Var}(a'; [0, t]) < \infty \tag{3.30}$$

Then ((3.31) follows by differentiating both sides)

$$\begin{aligned}&\frac{1}{2}\int_0^t a_n'(t-\tau)\left\|\int_\tau^t z(s)\,ds\right\|^2 d\tau\\ &\quad=\int_0^t\int_0^\tau a_n'(\tau-s)\left\langle\int_s^\tau z(v)\,dv, z(\tau)\right\rangle ds\,d\tau\\ &\quad\quad+\frac{1}{2}\int_0^t\int_0^\tau\left\|\int_{\tau-s}^\tau z(v)\,dv\right\|^2 a_n''(s)\,ds\,d\tau.\end{aligned}\tag{3.31}$$

Using (3.29), (3.30), the fact that (for fixed τ) $\|\int_{\tau-s}^\tau z(v)\,dv\|^2$ is a continuous function of s, and [9, Thm. 16.4, p. 31] gives

$$\lim_{n\to\infty}\int_0^\tau\left\|\int_{\tau-s}^\tau z(v)\,dv\right\|^2 a_n''(s)\,ds=\int_0^\tau\left\|\int_{\tau-s}^\tau z(v)\,dv\right\|^2 da'(s),\tag{3.32}$$

for each $\tau\in[0,t]$. But combining (3.25), (3.30), (3.32) and the dominated convergence theorem yields

$$\lim_{n\to\infty}\int_0^t\int_0^\tau\left\|\int_{\tau-s}^\tau z(v)\,dv\right\|^2 a_n''(s)\,ds\,d\tau=\int_0^t\int_0^\tau\left\|\int_{\tau-s}^\tau z(v)\,dv\right\|^2 da'(s)\,d\tau.\tag{3.33}$$

Taking now limits in (3.31), using (3.33) for the last term and (3.25), (3.29), (3.30) and the dominated convergence theorem for the remaining terms gives (3.27).

Now replace the last term on the right side of (3.26) by the right side of (3.27), then differentiate the relation thus obtained and finally use Fubini's theorem. This provides an identity and our claim (3.26) is hence true.

Apply the Lemma and take $t=t_m$, $z=h_n$ in (3.26). Then use (3.20) (with $i=m$) to estimate the first term on the right side of (3.26), and (3.23), (3.24) for the remaining terms on the right. One obtains

$$\begin{aligned}&\int_0^{t_m}\left\langle h_n(\tau),\int_0^\tau a(\tau-s)h_n(s)\,ds\right\rangle d\tau\\ &\quad\geqq\frac{\delta^2}{2}a(t_m)-8\delta^2\int_0^{t_m}\int_0^\tau|da'(s)|\,d\tau\\ &\quad\quad-10\delta^2\int_0^{t_m}|a'(\tau)|\,d\tau\geqq\frac{\eta\delta^2}{2}\overset{\text{def}}{=}\varepsilon>0,\end{aligned}\tag{3.34}$$

where the second inequality follows by (3.14).

Write (3.1) with $\lambda=\lambda_n$, $\lambda=\mu_n$, take differences and multiply by h_n. Then use (3.34). This yields

$$\int_0^{t_m}\langle u_{\lambda_n}-u_{\mu_n}, g_{\lambda_n}u_{\lambda_n}-g_{\mu_n}u_{\mu_n}\rangle\,d\tau\leqq-\varepsilon.\tag{3.35}$$

But $g_\lambda u_\lambda\in g(J_\lambda u_\lambda)$ and so, by the monotonicity of g,

$$\langle g_{\lambda_n}u_{\lambda_n}-g_{\mu_n}u_{\mu_n}, -J_{\lambda_n}u_{\lambda_n}+J_{\mu_n}u_{\mu_n}\rangle\leqq 0.\tag{3.36}$$

Integrate (3.36) over $(0, t_m)$, add the result to (3.35), and use $u_\lambda - J_\lambda u_\lambda = \lambda g_\lambda u_\lambda$. This gives

$$\int_0^{t_m} \langle g_{\lambda_n} u_{\lambda_n} - g_{\mu_n} u_{\mu_n}, \lambda_n g_{\lambda_n} u_{\lambda_n} - \mu_n g_{\mu_n} u_{\mu_n} \rangle \, d\tau \leqq -\varepsilon < 0. \tag{3.37}$$

But expanding (3.37), using (3.10) and the fact that $\lambda_n \downarrow 0$, $\mu_n \downarrow 0$, shows that (3.37) implies $0 \leqq -\varepsilon$. This is absurd. Hence (3.16) leads to a contradiction, and (3.15) is true. Note that the convergence in (3.15) is uniform on $[0, \gamma]$.

Having (3.15) we proceed to prove (3.12). For this we suppose the equality in (3.12) holds for $0 \leqq t \leqq p\gamma$, p an integer, and show that then it holds on $p\gamma \leqq t \leqq (p+1)\gamma$. Consider

$$u_\lambda(t) + \int_0^t a(t-\tau) g_\lambda u_\lambda(\tau) \, d\tau = f(t), \qquad p\gamma \leqq t \leqq (p+1)\gamma, \tag{3.38}$$

which we rewrite as

$$v_\lambda(s) + \int_0^s a(s-x) g_\lambda v_\lambda(x) \, dx = f(s+p\gamma) - \int_0^{p\gamma} a(s+p\gamma-\tau) g_\lambda u_\lambda(\tau) \, d\tau, \tag{3.39}$$

where $s = t - p\gamma$, $v_\lambda(s) = u_\lambda(s+p\gamma)$. We claim that for $0 \leqq s \leqq \gamma$,

$$\lim_{\lambda,\mu \to 0} \left\| \int_0^s [g_\lambda v_\lambda - g_\mu v_\mu] \, d\tau \right\| = 0. \tag{3.40}$$

To prove this claim we assume it does not hold and show that a contradiction (namely (3.45), (3.48)) follows. Hence suppose there exist $\hat{s}$, δ, $0 < \hat{s} \leqq \gamma$, $\delta > 0$, $\lambda_n \downarrow 0$, $\mu_n \downarrow 0$, such that (compare with (3.16))

$$\lim_{n \to \infty} \left\| \int_0^{\hat{s}} [g_{\lambda_n} v_{\lambda_n} - g_{\mu_n} v_{\mu_n}] \, d\tau \right\| = 2\delta > 0. \tag{3.41}$$

Now repeat the arguments that made up the first half of the proof of (3.15) (including the use of the Lemma), but with $\hat{t}$, h_n replaced by $\hat{s}$, $g_{\lambda_n} v_{\lambda_n} - g_{\mu_n} v_{\mu_n}$ respectively. This gives (compare with (3.34)) that there exists $\varepsilon > 0$ and $0 < s_m \leqq \gamma$ satisfying

$$\int_0^{s_m} \left\langle k_n(\tau), \int_0^\tau a(\tau - s) k_n(s) \, ds \right\rangle d\tau \geqq \varepsilon, \tag{3.42}$$

where $k_n \overset{\text{def}}{=} g_{\lambda_n} v_{\lambda_n} - g_{\mu_n} v_{\mu_n}$. Observe that as $[p\gamma, (p+1)\gamma] \subset [0, T_1]$ we still have (3.10) and so (3.42) can be obtained.

Replace λ in (3.39) by λ_n, then by μ_n, take differences, multiply by k_n, and integrate over $(0, s_m)$. This gives

$$\begin{aligned} &\int_0^{s_m} \langle k_n(\tau), v_{\lambda_n}(\tau) - v_{\mu_n}(\tau) \rangle \, d\tau + \int_0^{s_m} \left\langle k_n(\tau), \int_0^\tau a(\tau - s) k_n(s) \, ds \right\rangle d\tau \\ &\qquad = \int_0^{s_m} \left\langle k_n(\tau), \int_0^{p\gamma} a(\tau + p\gamma - s) h_n(s) \, ds \right\rangle d\tau, \end{aligned} \tag{3.43}$$

where $h_n \overset{\text{def}}{=} g_{\mu_n}u_{\mu_n} - g_{\lambda_n}u_{\lambda_n}$. By the monotonicity of g one has (compare with the arguments running from (3.35) to (3.37))

$$\int_0^{s_m} \langle k_n(\tau),\, v_{\lambda_n}(\tau) - v_{\mu_n}(\tau)\rangle\, d\tau \geqq -\frac{\varepsilon}{2}. \tag{3.44}$$

By (3.42)–(3.44)

$$\int_0^{s_m} \left\langle k_n(\tau),\, \int_0^{p\gamma} a(\tau + p\gamma - s)h_n(s)\, ds \right\rangle d\tau \geqq \varepsilon. \tag{3.45}$$

But

$$\begin{aligned}\int_0^{p\gamma} a(\tau + p\gamma - s)h_n(s)\, ds &= a(\tau + p\gamma)\int_0^{p\gamma} h_n(s)\, ds \\ &\quad - \int_0^{p\gamma} a'(\tau + p\gamma - x)\left\{\int_x^{p\gamma} h_n(s)\, ds\right\} dx,\end{aligned} \tag{3.46}$$

and by assumption

$$\lim_{n\to\infty} \left\| \int_0^t h_n(\tau)\, d\tau \right\| = 0, \quad \text{uniformly on } 0 \leqq t \leqq p\gamma. \tag{3.47}$$

Estimating the left side L_n of (3.45) by the aid of (3.10), (3.46), (3.47) one therefore gets

$$|L_n| \leqq \varepsilon_n \int_0^{s_m} \|k_n(\tau)\| [|a(\tau + p\gamma)| + \int_0^{p\gamma} |a'(\tau + \rho\gamma - s)|\, ds]\, d\tau \leqq \varepsilon_n c_1 \tag{3.48}$$

for some $\varepsilon_n \downarrow 0$ and some constant c_1. But (3.45), (3.48) are in obvious contradiction and so (3.40) holds.

Having (3.40) we recall the lines preceding (3.38) and hence realize that we have proved (3.12).

With (3.12) ready we proceed to show that (3.13) (rewritten below as (3.51)) follows. Take any $\lambda_n \downarrow 0$, $\mu_n \downarrow 0$. By (3.12), (3.18),

$$\lim_{n\to\infty} \left\| \int_{s_1}^{s_2} [g_{\lambda_n}u_{\lambda_n} - g_{\mu_n}u_{\mu_n}]\, d\tau \right\| = 0, \tag{3.49}$$

uniformly for $0 \leqq s_1 < s_2 \leqq T_1$. Let $\lambda = \lambda_n$, μ_n in (3.1), take differences use Fubini's theorem and the absolute continuity of a. This gives

$$\begin{aligned}&u_{\lambda_n}(t) - \mu_{\mu_n}(t) + \int_0^t a(t-\tau)[g_{\lambda_n}u_{\lambda_n}(\tau) - g_{\mu_n}u_{\mu_n}(\tau)]\, d\tau \\ &\quad = u_{\lambda_n}(t) - u_{\mu_n}(t) + a(t)\int_0^t [g_{\lambda_n}u_{\lambda_n} - g_{\mu_n}u_{\mu_n}]\, d\tau \\ &\qquad - \int_0^t a'(t-\tau)\left\{\int_\tau^t [g_{\lambda_n}u_{\lambda_n} - g_{\mu_n}u_{\mu_n}]\, ds\right\} d\tau = 0,\end{aligned} \tag{3.50}$$

and so, by the absolute continuity of a, (3.49), (3.50),

$$\lim_{n\to\infty} \|u_{\lambda_n}(t) - u_{\mu_n}(t)\| = 0, \quad \text{uniformly on } [0, T_1]. \tag{3.51}$$

Once the key result (3.51) is established we may complete the existence proof on $[0, T_1]$. By (3.10) there exists

$$w \in L_2(0, T_1; H), \tag{3.52}$$

such that for some $\lambda_n \downarrow 0$,

$$g_{\lambda_n} u_{\lambda_n} \rightharpoonup w \quad \text{in } L_2(0, T_1; H). \tag{3.53}$$

Take any such $\{\lambda_n\}$, w. By (3.51) there exists u satisfying

$$u_{\lambda_n} \to u \quad \text{in } C([0, T_1], H). \tag{3.54}$$

Let $D\hat{g} \stackrel{\text{def}}{=} \{u | u \in L_2(0, T_1; H), u(t) \in Dg \text{ a.e. on } 0 \leqq t \leqq T_1\}$, and define $\hat{g}$ on $D\hat{g}$ by

$$\hat{g}u = \{v | v \in L_2(0, T_1; H), v(t) \in g(u(t)) \text{ a.e. on } (0, T_1)\}. \tag{3.55}$$

From the maximal monotonicity of g follows that $\hat{g}$ is maximal monotone. Hence $\hat{g}$ is demiclosed. Using this fact together with (3.53), (3.54) gives

$$u \in D\hat{g}, \qquad w \in \hat{g}u, \tag{3.56}$$

and so

$$w(t) \in g(u(t)) \quad \text{a.e. on } (0, T_1). \tag{3.57}$$

Next define $\hat{f}$ by (note that $\hat{f}$ is independent of n)

$$\hat{f}(t) = u(t) - u_{\lambda_n}(t) + \int_0^t a(t-\tau)[w(\tau) - g_{\lambda_n} u_{\lambda_n}(\tau)]\, d\tau, \tag{3.58}$$

take an arbitrary fixed $\hat{t} \in [0, T_1]$, and form the scalar product of (3.58) and $\hat{f}(\hat{t})$. This gives

$$\|\hat{f}(\hat{t})\|^2 = \langle \hat{f}(\hat{t}), u(\hat{t}) - u_{\lambda_n}(\hat{t}) \rangle + \int_0^{T_1} \langle a(\hat{t}-\tau)\hat{f}(\hat{t}), w(\tau) - g_{\lambda_n} u_{\lambda_n}(\tau) \rangle\, d\tau, \tag{3.59}$$

where $a(t) \stackrel{\text{def}}{=} 0$, $t < 0$. But by (3.53), (3.54) the right side of (3.59) $\to 0$, for $n \to \infty$. Hence $\|\hat{f}(\hat{t})\| = 0$. As $\hat{t}$ was arbitrary we conclude from (3.1) and (3.58) that

$$u(t) + \int_0^t a(t-\tau) w(\tau)\, d\tau = f(t), \qquad 0 \leqq t \leqq T_1. \tag{3.60}$$

By (3.52), (3.57), (3.60) we have existence on $[0, T_1]$.

To show that a solution u exists on $[0, T]$ we proceed as follows. We assume existence has been proved on $0 \leqq t \leqq nT_1$ and demonstrate that we then have existence on $[0, (n+1)T_1]$, assuming $(n+1)T_1 \leqq T$. As T_1 is restricted only by (3.4), this will eventually get us to T. Thus suppose there exists

$$w \in L_2(0, nT_1; H), \tag{3.61}$$

and $u \in C([0, nT_1], H)$ such that

$$u(t) + \int_0^t a(t-\tau) w(\tau)\, d\tau = f(t), \qquad 0 \leqq t \leqq nT_1. \tag{3.62}$$

Observe that

$$\int_0^t a(t-\tau)w(\tau)\,d\tau = a(0)\int_0^t w(s)\,ds + \int_0^t \left\{\int_0^\tau w(\tau-s)a'(s)\,ds\right\} d\tau, \tag{3.63}$$

which can be verified by applying Fubini's theorem to the last term. Consequently the convolution term in (3.62) is absolutely continuous, differentiable a.e., and by using (2.6) also, we have

$$u'(t)+a(0)w(t)+\int_0^t a'(t-\tau)w(\tau)\,d\tau = f'(t), \tag{3.64}$$

a.e. on $(0, nT_1)$. From the absolute continuity of a, (2.6), (3.61), (3.64) follows

$$u' \in L_2(0, nT_1; H), \tag{3.65}$$

and so, by (3.57), (3.61), (3.65) and as $u \in L_2(0, nT_1; H)$,

$$\varphi(u(t)) \text{ is absolutely continuous on } [0, nT_1], \tag{3.66}$$

$$\frac{d}{dt}\varphi(u(t)) = \langle w(t), u'(t)\rangle \quad \text{a.e. on } (0, nT_1). \tag{3.67}$$

By (3.66) $\varphi(u(t))$ is bounded on $[0, nT_1]$. Therefore

$$u(t) \in D\varphi, \qquad 0 \leqq t \leqq nT_1. \tag{3.68}$$

We rewrite (1.1) on $nT_1 \leqq t \leqq (n+1)T_1$ as

$$v(s)+\int_0^s a(s-\tau)g(v(\tau))\,d\tau \ni f(s+nT_1)-\int_0^{nT_1} a(s+nT_1-\tau)w(\tau)\,d\tau, \tag{3.69}$$

where $0 \leqq s \leqq T_1$, $s = t-nT_1$, $v(s) = u(s+nT_1)$. Clearly $v(0) = u(nT_1)$, and so, by (3.68), $v(0) \in D\varphi$. Hence, by (2.6), one realizes that to apply the existence result obtained on $[0, T_1]$ to (3.69) one only has to show that the integral term on the right side of (3.69), call it $f_n(s)$, is absolutely continuous and satisfies

$$f_n' \in L_2(0, T_1; H). \tag{3.70}$$

However, by (3.61) and the absolute continuity of a,

$$f_n(s) = f_n(0)+\int_0^s \left\{\int_0^{nT_1} a'(u+nT_1-\tau)w(\tau)\,d\tau\right\} du, \tag{3.71}$$

which can be checked by interchanging the order of integration. Finally, it is not hard to show that

$$\int_0^{T_1} \left\| \int_0^{nT_1} a'(s+nT_1-\tau)w(\tau)\,d\tau \right\|^2 ds < \infty. \tag{3.72}$$

Hence we may apply the existence result obtained on $[0, T_1]$ to (3.69) and consequently, remembering the lines before (3.61), we have existence on $[0, T]$.

This completes the existence part of Theorem 1(a).

To prove the existence part of Theorem 1(b) one proceeds in steps of size T. Thus, one assumes existence has been obtained on $[0, nT]$ and rewrites (1.1) on

$[nT, (n+1)T]$ as

$$v(s)+\int_0^s a(s-\tau)g(v(\tau))\,d\tau \ni f(s+nT)-\int_0^{nT} a(s+nT-\tau)w(\tau)\,d\tau, \qquad 0\leqq s\leqq T, \tag{3.73}$$

where $s=t-nT$, $v(s)=u(s+nT)$. Applying the existence result of part (a) to (3.73), which by (2.7), (2.8), (3.68) (where we replace T_1 by T), and the fact that $w\in L_2(0, nT; H)$ can be done, gives the existence part of Theorem 1(b). Note that (compare with (3.72)) the integral term $(=q_n)$ on the right side of (3.73) satisfies $q_n'\in L_2(0, T; H)$.

To complete the proof of Theorem 1 we demonstrate uniqueness. Let u, v be two solutions of (1.1) on $[0, \gamma]$, γ as in (3.14). Then

$$u(t)-v(t)+\int_0^t a(t-\tau)[w_u(\tau)-w_v(\tau)]\,d\tau=0, \qquad w_u, w_v\in L_2(0, \gamma; H). \tag{3.74}$$

We assert that

$$\int_0^t [w_u-w_v]\,d\tau=0, \qquad 0\leqq t\leqq\gamma. \tag{3.75}$$

Suppose (3.75) does not hold and pick any $\hat{t}$, δ such that

$$\left\|\int_0^{\hat{t}} [w_u-w_v]\,d\tau\right\|=2\delta>0. \tag{3.76}$$

For s_1, s_2 arbitrary $(0\leqq s_1<s_2\leqq\gamma)$ we have, by the second part of (3.74),

$$\left\|\int_{s_1}^{s_2} [w_u-w_v]\,d\tau\right\|\leqq\mu[s_2-s_1]^{1/2}, \tag{3.77}$$

for some constant μ and so

$$\text{if } [s_2-s_1]^{1/2}\leqq\delta\mu^{-1}, \text{ then } \left\|\int_{s_1}^{s_2} [w_u-w_v]\,d\tau\right\|\leqq\delta. \tag{3.78}$$

Divide $[0, \hat{t}\,]$ in subintervals as follows:

$$\begin{aligned} &0=t_0<t_1<\cdots<t_R=\hat{t},\\ &t_{i+1}-t_i\leqq\delta\mu^{-1}, \qquad i=0, 1, \cdots, R-1. \end{aligned} \tag{3.79}$$

Then take the smallest i-value (call it m) for which

$$\left\|\int_0^{t_i} [w_u-w_v]\,d\tau\right\|\geqq\delta. \tag{3.80}$$

For $t\in[0, t_m]$ we may write $\int_0^t=\int_{\tilde{t}}^t+\int_0^{\tilde{t}}$ where $\tilde{t}=\max\{t_j\,|\,j=0, 1, \cdots, m-1;\ t_j\leqq t\}$, and so, by (3.78)–(3.80),

$$\left\|\int_0^t [w_u-w_v]\,d\tau\right\|\leqq 2\delta, \qquad \left\|\int_{s_1}^{s_2} [w_u-w_v]\,d\tau\right\|\leqq 4\delta, \tag{3.81}$$

for $t \in [0, t_m]$, $0 \leqq s_1 < s_2 \leqq t_m$. Take $z = w_u - w_v$ and $t = t_m$ in (3.26). By the second part of (3.74) this can be done. Then use (3.80) (with $i = m$), (3.81), and finally recall (3.14). This gives

$$\int_0^{t_m} \left\langle w_u(\tau) - w_v(\tau), \int_0^\tau a(\tau - s)[w_u(s) - w_v(s)]\, ds \right\rangle d\tau \geqq \varepsilon > 0, \tag{3.82}$$

for some ε, and therefore, after multiplying the equality in (3.74) by $w_u - w_v$, and integrating over $[0, t_m]$,

$$\int_0^{t_m} \langle w_u - w_v, u - v \rangle\, d\tau \leqq -\varepsilon. \tag{3.83}$$

This, however, is absurd as g is monotone. Hence the assertion (3.75) follows. To see that (3.75) implies $u = v$ on $[0, \gamma]$ one performs the usual integration by parts (see (3.50)). By translation and repeating the arguments above one finally obtains $u = v$ on the whole interval of existence. This completes the proof of Theorem 1.

4. Proof of Theorem 2. We begin by proving (2.21). This will be accomplished once we get to (4.9).

Form the scalar product of w and (3.64) (which holds under the present hypothesis), integrate over $(0, t)$ and use (2.19). This gives

$$\begin{aligned} \varphi(t) - \varphi(0) + \int_0^t \left\langle w(\tau), a(0)w(\tau) + \int_0^\tau a'(\tau - s)w(s)\, ds \right\rangle d\tau \\ = \int_0^t \langle w(\tau), f'(\tau) \rangle\, d\tau, \qquad 0 \leqq t < \infty. \end{aligned} \tag{4.1}$$

The equality in (4.2) below can be verified by expanding the right side of this equality. The inequality in (4.2) follows by (2.15), (2.16).

$$\begin{aligned} &\int_0^t \left\langle w(\tau), a(0)w(\tau) + \int_0^\tau a'(\tau - s)w(s)\, ds \right\rangle d\tau \\ &\quad = -\frac{1}{2}\int_0^t \left\{ \int_0^\tau a'(s)\|w(\tau) - w(\tau - s)\|^2\, ds \right\} d\tau \\ &\qquad + \frac{1}{2}\int_0^t a(t - \tau)\|w(\tau)\|^2\, d\tau + \frac{1}{2}\int_0^t a(\tau)\|w(\tau)\|^2\, d\tau \\ &\quad \geqq \frac{1}{2}\int_0^t a(t - \tau)\|w(\tau)\|^2\, d\tau. \end{aligned} \tag{4.2}$$

Pick any $\delta > 0$ such that $a(\delta) > 0$ and such that $\delta M = 1$ for some integer M. By (2.16), $a(t) \geqq a(\delta)$, $0 \leqq t \leqq \delta$. Combining this fact with (2.15), (4.1), (4.2) implies

$$\frac{a(\delta)}{2}\int_{t-\delta}^t \|w(\tau)\|^2 d\tau \leqq \int_0^t \langle w(\tau), f'(\tau) \rangle\, d\tau + \varphi(0) - \varphi(t), \qquad t \geqq \delta. \tag{4.3}$$

But (let $t = \delta N$, where N an integer)

$$\begin{aligned}\int_0^t \langle w, f'\rangle\, d\tau &\leqq \sum_{n=0}^{N} \int_{n\delta}^{(n+1)\delta} |\langle w, f'\rangle|\, d\tau \\ &\leqq \sum_{n=0}^{N} \left\{\left[\int_{n\delta}^{(n+1)\delta} \|w\|^2\, d\tau\right]^{1/2}\left[\int_{n\delta}^{(n+1)\delta} \|f'\|^2\, d\tau\right]^{1/2}\right\}.\end{aligned} \tag{4.4}$$

Assume t_p is such that

$$\int_{t-\delta}^{t} \|w(\tau)\|^2\, d\tau \leqq \int_{t_p-\delta}^{t_p} \|w(\tau)\|^2\, d\tau, \qquad \delta \leqq t \leqq t_p. \tag{4.5}$$

Then, by (4.3)–(4.5), (take $t = t_p$ in (4.3), (4.4)), and using an elementary inequality to to get the second inequality in (4.6), one obtains

$$\begin{aligned}&\frac{a(\delta)}{2}\int_{t_p-\delta}^{t_p} \|w(\tau)\|^2\, d\tau \\ &\leqq \left[\int_{t_p-\delta}^{t_p} \|w(\tau)\|^2\, d\tau\right]^{1/2} \sum_{n=0}^{\infty}\left\{\int_{n\delta}^{(n+1)\delta} \|f'(\tau)\|^2\, d\tau\right\}^{1/2} + \varphi(0) - \varphi(t) \\ &\leqq MK_f\left[\int_{t_p-\delta}^{t_p} \|w\|^2\, d\tau\right]^{1/2} + \varphi(0) - \varphi(t),\end{aligned} \tag{4.6}$$

where $K_f = \sum_{n=0}^{\infty}\{\int_n^{n+1} \|f'\|^2\, d\tau\}^{1/2}$. Take $t = t_p$ in (1.1); estimate and use (4.5). This gives

$$\begin{aligned}\|u(t_p)\| &\leqq \sup_{t\in R_+} \|f(t)\| + \left[\int_{t_p-\delta}^{t_p} \|w\|^2\, d\tau\right]^{1/2} \sum_{n=0}^{\infty}\left[\int_{n\delta}^{(n+1)\delta} a^2(\tau)\, d\tau\right]^{1/2} \\ &\leqq \sup_{t\in R_+} \|f(t)\| + \left[\int_{t_p-\delta}^{t_p} \|w\|^2\, d\tau\right]^{1/2}\left[\delta^{1/2}a(0) + \delta^{-1/2}\int_0^{\infty} a(\tau)\, d\tau\right],\end{aligned} \tag{4.7}$$

where we used (2.15)–(2.17) to obtain the second inequality.

Combining (2.20), (4.6), (4.7) gives

$$\begin{aligned}&\frac{a(\delta)}{2}\int_{t_p-\delta}^{t_p} \|w\|^2\, d\tau \\ &\leqq MK_f\left[\int_{t_p-\delta}^{t_p} \|w\|^2\, d\tau\right]^{1/2} + |\varphi(0)| + \beta \\ &\quad + \alpha\left[\sup_{t\in R^+} \|f\| + \left[\delta^{1/2}a(0) + \delta^{-1/2}\int_0^{\infty} a(\tau)\, d\tau\right]\left[\int_{t_p-\delta}^{t_p} \|w\|^2\, d\tau\right]^{1/2}\right]^q.\end{aligned} \tag{4.8}$$

As $q < 2$ it follows from (4.8) that there exists $K_w < \infty$ such that

$$\int_{t_p-\delta}^{t_p} \|w(\tau)\|^2\, d\tau \leqq K_w. \tag{4.9}$$

Combining (4.5), (4.9) gives (2.21).

Repeating the estimates in (4.7), but with t_p replaced by an arbitrary t, and using (4.9), one obtains (2.22).

We finally prove (2.23). Rewrite (3.64) as follows:

$$u'(t)+\int_0^t a'(\tau)[w(t-\tau)-w(t)]\,d\tau+a(t)w(t)=f'(t). \tag{4.10}$$

Estimating the integral in (4.10) one gets

$$\begin{aligned}&\int_0^t |a'(\tau)|\,\|w(t-\tau)-w(t)\|\,d\tau\\ &\leqq\left[\int_0^t |a'(\tau)|\,d\tau\right]^{1/2}\left[\int_0^t |a'(\tau)|\,\|w(t-\tau)-w(t)\|^2\,dt\right]^{1/2}\\ &\leqq\sqrt{a(0)}\left[\int_0^t |a'(\tau)|\,\|w(t-\tau)-w(t)\|^2\,d\tau\right]^{1/2}.\end{aligned} \tag{4.11}$$

and so, using (4.10), (4.11) and squaring one obtains

$$\|u'(t)\|^2\leqq 3\left[\|f'\|^2+a^2(t)\|w(t)\|^2+a(0)\int_0^t |a'(\tau)|\,\|w(t-\tau)-w(t)\|^2\,dt\right]. \tag{4.12}$$

By (2.18),

$$f'\in L_2(0,\infty;H). \tag{4.13}$$

Invoking (2.18), (2.21), (4.4) one has

$$\sup_{t\in R_+}\left|\int_0^t \langle w(\tau),f'(\tau)\rangle\,d\tau\right|<\infty, \tag{4.14}$$

and from (2.20), (2.22) it immediately follows that

$$\inf_{t\in R^+}\varphi(t)>-\infty. \tag{4.15}$$

Therefore, by (2.15), (2.16), (4.1), the equality in (4.2), (4.14), (4.15) one obtains

$$\sup_{t\in R_+}\int_0^t a(\tau)\|w(\tau)\|^2\,d\tau<\infty, \tag{4.16}$$

$$\sup_{t\in R_+}\int_0^t\int_0^\tau |a'(s)|\,\|w(\tau)-w(\tau-s)\|^2\,ds\,d\tau<\infty. \tag{4.17}$$

But, recalling (2.15), (2.16), one gets

$$\int_0^t a^2(\tau)\|w(\tau)\|^2\,d\tau\leqq a(0)\int_0^t a(\tau)\|w(\tau)\|^2\,d\tau,\qquad t\geqq 0, \tag{4.18}$$

and so, from (4.16), (4.18),

$$a(t)\|w(t)\|\in L_2(0,\infty). \tag{4.19}$$

Integrating (4.12) and using (4.13), (4.17), (4.19), one arrives at (2.23). This completes the proof of Theorem 2.

5. Proofs of Corollaries 1 and 2. The proofs of both Corollaries 1 and 2 closely follow that of Theorem 2 and so we indicate only the necessary changes.

Proof of Corollary 1. The only part of the hypothesis of Theorem 2 that has been dropped is (2.17). On the other hand, when proving (2.21) in Theorem 2 we needed (2.17) only to obtain (4.7) which in turn was needed only to estimate $\varphi(t)$ from below. But as we now a priori assume $\varphi(t)$ bounded below (see (2.24)), then (2.17) is clearly superfluous for the validity of (2.21).

Examining the proof of (2.23) (that is, the arguments running from (4.10) to (4.19)), one realizes that if (4.15) equal to (2.24) is assumed a priori then there is no need for (2.17), (2.22).

Proof of Corollary 2. The relation (4.1) and the equality in (4.2) are still valid. Hence, using also (2.15), (2.16), (2.24), one obtains

$$\frac{1}{2}\int_0^t a(\tau)\|w(\tau)\|^2\,d\tau \leqq \left|\int_0^t \langle f', w\rangle\,d\tau\right|, \qquad t \geqq 0, \tag{5.1}$$

and so, by (2.15), (2.16), (2.25), (2.26), (5.1),

$$\frac{a(\infty)}{2}\int_0^t \|w(\tau)\|^2\,d\tau \leqq \left[\int_0^t \|f'(\tau)\|^2\,d\tau\right]^{1/2}\left[\int_0^t \|w(\tau)\|^2\,d\tau\right]^{1/2}, \tag{5.2}$$

which yields (2.27).

To prove (2.23) we note that (4.10)–(4.13) still hold. The relation (4.14) follows by (2.26), (2.27) and (4.15) is (2.24). Hence we have (4.17). But now, if we integrate (4.12), and use (2.27), (4.13), (4.17), and the fact that $\sup_{t\in R^+}|a(t)|<\infty$, we obtain (2.23).

To show that $\varphi' \in L_1(0,\infty)$ and hence that $\lim_{t\to\infty}\varphi(t)$ exists, it suffices to combine (2.19), (2.23), (2.27).

6. Applications. Let Ω be a bounded open subset of R^N with sufficiently smooth boundary Γ and let H^k, H_0^k stand for the Sobolev spaces. Choose j such that

(6.1) j is a convex, lower semicontinuous map of $R \to (-\infty, \infty]$, $\quad j \not\equiv +\infty$,

and denote $\beta = \partial j$.

Example 1 [2, p. 111]. Define $\varphi(u)$ for $u \in L_2(\Omega)$ by

$$\varphi(u) = \begin{cases} \dfrac{1}{2}\displaystyle\int_\Omega |\operatorname{grad} u|^2\,dx + \int_\Gamma j(u)\,dx, \\ \quad \text{for } u \in D(\varphi) = \{u \mid u \in H^1(\Omega),\ j(u) \in L_1(\Gamma)\}, \\ +\infty, \quad \text{for } u \in L_2(\Omega)\backslash D(\varphi). \end{cases} \tag{6.2}$$

Then $\varphi: L_2(\Omega) \to (-\infty, \infty]$, φ is lower semicontinuous and convex, and clearly $\varphi \not\equiv +\infty$. It also follows that

$$\partial\varphi(u) = -\Delta u, \qquad u \in D(\partial\varphi) = \{u \mid u \in H^2(\Omega),\ -\frac{\partial u}{\partial n} \in \beta(u) \text{ a.e. on } \Gamma\} \tag{6.3}$$

where $\partial/\partial n$ is the outward normal derivative. Let, for any $T>0$, $f(t,x)$ map $[0,T]\times\Omega \to R$ and satisfy

$$f \in AC([0,T]; L_2(\Omega)), \qquad \frac{\partial f}{\partial t} \in L_2(0,T; L_2(\Omega)), \tag{6.4}$$

$$f(0, x) \in D(\varphi). \tag{6.5}$$

Applying Theorem 1 with $H = L_2(\Omega)$ and recalling the absolute continuity of $\varphi(u(t))$ (see the observation following Theorem 1) we immediately have

COROLLARY 5. *Let φ satisfy* (6.1), (6.2) *and suppose f satisfies* (6.4), (6.5) *for every $T > 0$. Moreover assume $a(0) > 0$, $a(t)$ locally absolutely continuous on $[0, \infty)$ and $a' \in BV[0, \hat{T}]$ for some $\hat{T} > 0$. Then there exists a unique function $u(t, x)$, defined on $[0, \infty) \times \Omega$ satisfying*

$$u \in AC([0, T], L_2(\Omega)) \cap L_\infty(0, T; H^1(\Omega)) \cap L_2(0, T; H^2(\Omega)),$$

$$\frac{\partial u}{\partial t} \in L_2(0, T; L_2(\Omega)), \tag{6.6}$$

$$\frac{\partial u}{\partial n} \in -\beta(u) \quad \text{a.e. on } \Gamma \text{ for almost every } t \in (0, T),$$

$$u(t, x) - \int_0^t a(t-\tau)\, \Delta u(\tau, x)\, d\tau = f(t, x) \quad \text{on } [0, T] \times \Omega, \tag{6.7}$$

for every $T > 0$.

An application of Theorem 2 gives:

COROLLARY 6. *Let the hypothesis of Corollary* 1 *hold except that $a' \in BV[0, \hat{T}]$ need not be satisfied. In addition suppose*

$$a(t) \geqq 0, \quad t \geqq 0; \quad a'(t) \leqq 0 \quad \text{a.e. on } t \geqq 0, \quad a \in L_1(0, \infty),$$

$$\sum_{n=0}^{\infty} \left[\int_n^{n+1} \|f'\|^2_{L_2(\Omega)}\, d\tau \right]^{1/2} < \infty.$$

Let $u(t, x)$ be a solution of (6.7) *on $[0, \infty) \times \Omega$. Then*

$$\sup_{t \in R^+} \int_t^{t+1} \|\Delta u(\tau, x)\|_{L_2(\Omega)}\, d\tau < \infty,$$

$$\sup_{t \in R^+} \|u(t, x)\|_{L_2(\Omega)} < \infty,$$

$$u' \in L_2(0, \infty; L_2(\Omega)).$$

Example 2 [2, p. 123]. In addition to (6.1) assume

$$\frac{j(r)}{|r|} \to \infty \quad \text{for } |r| \to \infty. \tag{6.8}$$

For $u \in H^{-1}(\Omega)$ define

$$\varphi(u) = \begin{cases} \int_\Omega j(u)\, dx, & u \in D(\varphi) = \{u \mid u \in L_1(\Omega), j(u) \in L_1(\Omega)\}, \\ +\infty, & u \in H^{-1}(\Omega) \backslash D(\varphi). \end{cases} \tag{6.9}$$

Then $\varphi: H^{-1}(\Omega) \to (-\infty, \infty]$, φ is lower semicontinuous and convex, not identically $+\infty$, and

$$\partial\varphi(u) = \{w \mid w \in H^{-1}(\Omega),\ -\Delta^{-1}w(x) \in \beta(u(x)) \text{ a.e. on } \Omega\}. \tag{6.10}$$

Our assumptions on f are

$$\begin{aligned} &f(t, x) \in AC([0, T]; H^{-1}(\Omega)), \qquad \frac{\partial f}{\partial t} \in L_2(0, T; H^{-1}(\Omega)), \\ &f(0, x) \in D(\varphi), \end{aligned} \tag{6.11}$$

for any $T > 0$ and we suppose a satisfies

$$\begin{aligned} &a(0) > 0, \qquad a(t) \text{ locally absolutely continuous on } [0, \infty), \\ &a' \in BV[0, \hat{T}] \quad \text{for some } \hat{T} > 0, \end{aligned} \tag{6.12}$$

Applying Theorem 1 with $H = H^{-1}(\Omega)$ and using the absolute continuity of $\varphi(u(t))$ gives (assuming β single-valued).

COROLLARY 7. *Let φ satisfy* (6.1), (6.8), (6.9) *and suppose f, a satisfy* (6.11), (6.12) *respectively. Then there exists a unique function $u(t, x)$ such that for any $T > 0$*

$$\begin{aligned} &u \in AC([0, T]; H^{-1}(\Omega)) \cap L_\infty(0, T; L_1(\Omega)), \\ &\frac{\partial u}{\partial t} \in L_2(0, T; H^{-1}(\Omega)), \\ &\beta(u(t, x)) \in H_0^1(\Omega) \quad \textit{on } (0, T), \end{aligned} \tag{6.13}$$

and such that

$$u(t, x) - \int_0^t a(t-\tau)\, \Delta\beta(u(\tau, x))\, d\tau = f(t, x) \quad \textit{on } [0, \infty) \times \Omega. \tag{6.14}$$

Acknowledgment. The present work was done during the academic year 1974–75 when the author was visiting the Mathematics Research Center at the University of Wisconsin, Madison. The author wishes to take this opportunity to thank the MRC for a stimulating and beneficial year.

REFERENCES

[1] V. BARBU, *Nonlinear Volterra equations in a Hilbert space*, this Journal, 6 (1975), pp. 728–741.

[2] H. BRÈZIS, *Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations*, Contributions to Nonlinear Functional Analysis, E. H. Zarantonello, ed., Academic Press, New York, 1971.

[3] ———, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*. North-Holland/American Elsevier, Amsterdam, 1973.

[4] J. J. LEVIN, *A bound on the solutions of a Volterra equation*, Arch. Rational Mech. Anal., 4 (1973), pp. 339–349.

[5] S.-O. LONDEN, *The qualitative behavior of the solutions of a nonlinear Volterra equation*, Michigan Math. J., 18 (1971), pp. 321–330.

[6] ———, *On a nonlinear Volterra integral equation*, J. Differential Equations, 14 (1973), pp. 106–120.

[7] ———, *On an integral equation in a Hilbert space*, MRC Tech. Summary Rep. 1527, Univ. of Wisconsin, Madison, 1975.
[8] R. C. MacCamy, *Nonlinear Volterra equations on a Hilbert space*, J. Differential Equations, 16 (1974), pp. 373–383.
[9] D. V. Widder, *The Laplace Transform*, Princeton University Press, Princeton, NJ, 1946.

SIAM J. MATH. ANAL.
Vol. 8, No. 6, November 1977

HYPOELLIPTIC INFINITESIMAL GENERATORS*

ALBERTO BAIDER AND BARRY CHERKAS†

Abstract. In this paper we study semi-group generation by semi-bounded second order differential operators on a noncompact C^∞ manifold. It is shown that the usual regularity assumptions can be relaxed to include hypoelliptic operators of the Hörmander type. The related question of the identity between weak and strong extensions for such operators is also studied. Sufficient conditions are given in terms of the behavior at infinity of an appropriate exhaustion function. We include examples to illustrate how this function may be chosen in concrete applications.

1. Introduction.

Consider the evolution equation

$$\frac{du}{dt} = Au, \qquad t>0, \tag{1.1}$$

in $L^2(\Omega, \mu)$, where Ω is an n-dimensional C^∞ manifold, μ is a C^∞ measure locally equivalent to Lebesgue measure, and A is a second order hypoelliptic differential operator with real C^∞ coefficients. The purpose of this paper is to discuss growth conditions on the coefficients in A for which there is a unique semi-group of solution operators $\{\exp(tA): t \geqq 0\}$ in $L^2(\Omega, \mu)$ associated with (1.1). At the same time, the techniques used here enable us to study the related problem of when the weak and strong extensions of A in $L^2(\Omega, \mu)$ are identical.

We suppose that in Ω, A can be written in the form

$$A = \sum_1^r X_j^2 + X_0 + c \tag{1.2}$$

where $X_0, \cdots, X_r$ denote first order homogeneous operators in Ω with real C^∞ coefficients and $A1 = c \in C^\infty(\Omega)$. Noting that the termination coefficient for the formal adjoint of A with respect to μ can be written A^*1, we require that

$$\sup\{\tfrac{1}{2}(A1 + A^*1)\} = \lambda_0 < \infty. \tag{1.3}$$

Let $D_1(A) = \{u \in L^2: Au \in L^2\}$, where Au is understood in the distributional sense, and denote by $D_0(A)$ the closure of $C_0^\infty(\Omega)$ in $D_1(A)$ equipped with the graph norm of A. Assuming (1.2) and (1.3), we use elliptic regularization to show that for λ sufficiently large the equation $\lambda u - Au = f$ can be solved in L^2 with a finite Dirichlet integral, $\sum_1^r \|X_j u\|^2$.

In order to insure that the solution u belongs to $D_0(A)$, we need some control over the growth of the coefficients at infinity as well as some kind of regularity for the operator. For the condition at infinity, we assume that there is a real valued proper function $p \in C^\infty(\Omega)$ and constant K such that whenever $|p| \geqq 1$ we have

$$\begin{aligned} &\text{(i)} \quad |X_j p| \leqq K|p|, && j = 0, \cdots, r, \\ &\text{(ii)} \quad |X_j^2 p| \leqq K|p|, && j = 1, \cdots, r. \end{aligned} \tag{1.4}$$

* Received by the editors December 18, 1975, and in revised form May 4, 1976.

† Department of Mathematics, Hunter College of the City University of New York, New York, New York 10021.

In R^n, this condition essentially means that the leading coefficient can grow up to order 2 (see Example 3.5). As far as regularity is concerned, we assume the explicit conditions given by Hörmander [1, Thm. 1.1] for hypoellipticity:

(1.5) The Lie algebra over the reals generated by the vector fields $\{X_0, \cdots, X_r\}$ has rank n at each $x \in \Omega$.

The equality between $D_0(A)$ and $D_1(A)$ will follow if A^* satisfies the same conditions as A. In addition to the conditions above, it is sufficient that

$$|X_j^* 1| \leqq K, \qquad j = 1, \cdots, r. \tag{1.6}$$

Several examples are given at the end.

The identity of weak and strong extensions has been studied by many authors, especially in connection with essential self-adjointness of Schrödinger operators (cf. [4] and the references listed there). Most of the recent papers (see viz. [3], [5], [6], [7]) are improvements upon the pioneering work of Ikebe–Kato [2]. Their main thrust is to find minimal regularity conditions on the coefficients. The present study is different in that we do not restrict our attention to elliptic or symmetric operators. Also, the operators we treat need not be defined in the whole of R^n, although we require the coefficients to be smooth. In addition, since we study semi-group generation, we require a condition (see (1.3)) that insures semi-boundedness of the operators.

Remark. Unlike the strongly elliptic case (see viz. Yosida [8, pp. 419–425]), hypoelliptic operators need not generate holomorphic semi-groups. Indeed, the operator $A = \partial^2/\partial x^2 + \partial/\partial y$ in R^2 generates a semi-group on $D_0(A) = D_1(A)$. The spectrum of A is the entire half plane $\operatorname{Re}(\lambda) \leqq 0$, as one easily sees by taking Fourier transform. However, the resolvent of a holomorphic semi-group generator must be defined in some sector $\pi/2 < |\arg \lambda| < \theta_0 < \pi$, at least when $|\lambda|$ is large (see viz. [8, pp. 256–257]).

2. A necessary condition. The purpose of this section is to prove that any second order differential infinitesimal generator with real C^∞ coefficients must have a positive semi-definite principal part. Although this may have been done before, we include a proof since we know of no reference.

THEOREM 2.1. *Let A be a differential operator with continuous coefficients on an n-dimensional C^∞ manifold Ω and suppose A is locally of order $m \geqq 1$ at $x^0 \in \Omega$. Let $C_0^\infty \subseteq D(A)$. A necessary condition for A to be an infinitesimal generator of a semi-group of class (C_0) in $L^2(\Omega, \mu)$ is that the principal symbol $a_m(x, i\xi)$ of A satisfy the inequality*

$$\operatorname{Re}(a_m(x^0, i\xi)) \leqq 0, \qquad \xi \in T^*_{x^0}\Omega.$$

Proof. Since A generates a semi-group of class (C_0) in $L^2(\Omega, \mu)$, we have the estimate

$$\|\lambda\phi - A\phi\|^2 \geqq M(\operatorname{Re}(\lambda) - \lambda_0)^2 \|\phi\|^2 \tag{2.1}$$

for $\phi \in C_0^\infty(\Omega)$, $\operatorname{Re}(\lambda) > \lambda_0$, and some constants $M > 0$ and λ_0. Consider the expression

$$A(e^{it\langle x,\xi\rangle} t^{-m}\phi) = e^{it\langle x,\xi\rangle}\{a_m(x, i\xi)\phi + \cdots\}$$

where the dots indicate terms involving negative powers of t. Let $\operatorname{Re}(z)>0$ and set $\lambda = zt^m$. Upon replacing ϕ in (2.1) by $e^{it\langle x,\xi\rangle}t^{-m}\phi$ and letting $t\to\infty$, we obtain

$$\int_\Omega |(z-a_m(x,i\xi))\phi|^2\,d\mu \geqq M(\operatorname{Re}(z))^2\int_\Omega |\phi|^2\,d\mu. \tag{2.2}$$

If for some ξ^0 we have $\operatorname{Re}(a_m(x^0,i\xi^0))>0$, set $z=a_m(x^0,i\xi^0)$ so that $\operatorname{Re}(z)>0$ and by the continuity of $a_m(x,\xi^0)$, (2.2) cannot hold for all $\phi\in C_0^\infty(\Omega)$.

COROLLARY 2.2. *If m is even then* $\operatorname{Re}((-1)^{m/2}a_m(x^0,\xi))\leqq 0$. *If m is odd then* $\operatorname{Im}(a_m(x^0,\xi))=0$.

Proof. Let $a_m(x^0,\xi)=\sum_{|\alpha|=m}a^\alpha(x^0)\xi^\alpha$. If m is even then

$$\operatorname{Re}(a_m(x^0,i\xi))=\operatorname{Re}((-1)^{m/2}a_m(x^0,\xi)).$$

If m is odd we have

$$0\geqq\operatorname{Re}(a_m(x^0,-i\xi))=-\operatorname{Re}(a_m(x^0,i\xi)).$$

Thus, $0=\operatorname{Re}(a_m(x^0,i\xi))=\pm\operatorname{Im}(a_m(x^0,\xi))$.

3. Generating semi-groups with hypoelliptic operators. Our starting point is a coordinate free integration by parts. Let A be a second order differential operator which we take to be real (in the sense that $A\phi$ is real whenever ϕ is real.) If $x=(x_1,\cdots,x_n)$ is a local system of coordinates then we may write $A=a_2(x,\partial)+\cdots$ where $a_2(x,\xi)=\sum a_{ij}(x)\xi_i\xi_j$ is the leading term and $\partial=(\partial/\partial x_1,\cdots,\partial/\partial x_n)$. We know that this function is invariantly defined on the cotangent bundle of Ω, for example as the leading term of the polynomial in t

$$e^{-t\phi}A(e^{t\phi})=a_2(x,\xi)t^2+\cdots,\qquad \xi=d\phi(x).$$

We shall use the abbreviation $a_2(d\phi)$. We will also have use for its polarization, the symmetric bilinear form

$$b(d\phi,d\psi)=\tfrac{1}{2}\{a_2(d\phi+d\psi)-a_2(d\phi)-a_2(d\psi)\}.$$

In local coordinates, $b(x;\xi,\eta)=\sum a_{ij}(x)\xi_i\eta_j$.

PROPOSITION 3.1. *For real valued $\phi\in C_0^\infty$ we have*

$$(A\phi,\phi)=-\int_\Omega a_2(d\phi)\,d\mu+\tfrac{1}{2}((A1+A^*1)\phi,\phi). \tag{3.1}$$

Proof. A computation shows that Leibniz' formula for A assumes the form

$$A(\phi\psi)=\phi A\psi+\psi A\phi-\phi\psi A1+2b(d\phi,d\psi).$$

In particular, when $\phi=\psi$ we obtain

$$\phi A\phi=-a_2(d\phi)+\tfrac{1}{2}\{A(\phi^2)+\phi^2A1\}.$$

Integration over Ω gives (3.1), if one notes that

$$\int_\Omega A(\phi^2)\,d\mu=\int_\Omega \phi^2A^*1\,d\mu=((A^*1)\phi,\phi).$$

For the remainder of this paper, we shall take A in the form (1.2). In this case, $a_2(d\phi) = \sum_1^r |X_j\phi|^2$, so that (3.1) may be written

$$(A\phi, \phi) = -\sum_1^r \|X_j\phi\|^2 + \tfrac{1}{2}((A\,1 + A^*1)\phi, \phi). \tag{3.2}$$

THEOREM 3.2. *Suppose A in (1.2) satisfies (1.3). Then for $\lambda > \lambda_0$ and $f \in L^2(\Omega, \mu)$ there is $u \in D_1(A)$ for which $\lambda u - Au = f$ and $X_j u \in L^2(\Omega, \mu)$ for $j = 1, \cdots, r$.*

Proof. We begin by constructing a sequence of elliptic operators that approximate A on $C_0^\infty(\Omega)$ and satisfy certain properties uniformly. Choose $\alpha \in C_0^\infty(R^1)$ so that $0 \leqq \alpha \leqq 1$, $\alpha(t) = 1$ if $|t| \leqq 1$, and $\alpha(t) = 0$ if $|t| \geqq 2$. Let p be a real valued C^∞ proper function on Ω and set

$$\alpha_n = \alpha(p/n) \in C_0^\infty(\Omega). \tag{3.3}$$

Denote by E a real elliptic operator in Ω which is formally self-adjoint with respect to μ, dissipative on C_0^∞, and satisfies $E1 = 0$. That such an operator exists may be seen, for example, by putting a Riemannian metric on Ω and defining $E\phi = \operatorname{div}_\mu(\operatorname{grad}\phi)$, where the divergence is taken with respect to μ and the gradient with respect to the Riemannian metric. If μ is the Riemannian volume element then E is the usual Laplacian on Ω. In general, it will differ from the Laplacian by lower order terms. Consider

$$A_n = \frac{1}{n}E + \alpha_n^2 A - \frac{1}{2}(A^*\alpha_n^2 - \alpha_n^2 A^*1).$$

From (1.3) we have

$$\tfrac{1}{2}(A_n 1 + A_n^* 1) = \tfrac{1}{2}\alpha_n^2(A\,1 + A^*1) \leqq \lambda_0 < \infty. \tag{3.4}$$

Observe that the quadratic form of A_n is nonnegative. Thus, for $\lambda > \lambda_0$ and $\phi \in C_0^\infty(\Omega)$, by Proposition 3.1 we have the estimate

$$\|\lambda\phi - A_n\phi\| = \{\|(\lambda - A_n)\operatorname{Re}\phi\|^2 + \|(\lambda - A_n)\operatorname{Im}\phi\|^2\}^{1/2} \geqq (\lambda - \lambda_0)\|\phi\|.$$

Let H_n be the completion of C_0^∞ in the norm $[\phi]_n^2 = (\lambda\phi - A_n\phi, \phi)$, $\lambda > \lambda_0$. Since A_n is elliptic and its antisymmetric part has compact support, it follows that the Hermitian form $(A_n\phi, \psi)$ is continuous in H_n. Indeed, the antisymmetric part

$$\tfrac{1}{2}(A_n - A_n^*) = \tfrac{3}{4}(\alpha_n^2 A - A^*\alpha_n^2)$$

can be written in the form $\sum_0^r \beta_{ni} X_i + \gamma_n$ where $\beta_{ni}, \gamma_n \in C_0^\infty(\Omega)$. Since $|(\beta_{ni}X_i\phi, \psi)| \leqq \|\beta_{ni}X\phi\|\,\|\psi\|$, it clearly suffices to show that for any compactly supported vector field X one has an estimate of the form $\|X\phi\| \leqq C[\phi]_n$, for ϕ in $C_0^\infty(\Omega)$. Observe that (3.4) together with Proposition (3.1) applied to A_n yields

$$[\phi]_n^2 \geqq (\lambda - \lambda_0)\|\phi\|^2 + \int_\Omega a_{n2}(d\phi)\, d\mu.$$

Noting that $X\phi$ can be expressed as the Riemannian inner product $\langle X, \nabla_n\phi\rangle$, where ∇_n denotes the gradient operator in the Riemannian metric defined by the symbol of A_n, since $\langle \nabla_n\phi, \nabla_n\phi\rangle = a_{n2}(d\phi)$ we have the estimate

$$|X\phi|^2 \leqq \langle X, X\rangle\langle \nabla_n\phi, \nabla_n\phi\rangle \leqq C^2 a_{n2}(d\phi).$$

Thus,

$$\|X\phi\|^2 = \int_\Omega |X\phi|^2 \, d\mu$$

$$\leqq C^2 \int_\Omega a_{n2}(d\phi)\, d\mu$$

$$\leqq C^2[\phi]_n^2.$$

Under these circumstances, the Lax–Milgram theorem (see viz. [8, p. 92]) implies that the resolvent of A_n as an operator on $D(A_n) = D_1(A_n) \cap H_n$ exists for $\lambda > \lambda_0$ and satisfies

$$\|R(\lambda ; A_n)\| \leqq (\lambda - \lambda_0)^{-1}. \tag{3.5}$$

From (3.2) we have

$$\|\alpha_n X_j \phi\|^2 \leqq (\lambda\phi - A_n\phi, \phi) = [\phi]_n^2, \qquad \phi \in C_0^\infty.$$

Thus, the map $\alpha_n X_j$ sends H_n continuously into $L^2(\Omega, \mu)$. It follows that any $u \in D(A_n)$ satisfies

$$\|\alpha_n X_j u\|^2 \leqq \|\lambda u - A_n u\| \|u\|, \tag{3.6}$$

since a continuity argument shows that for such u one has

$$[u]_n^2 \leqq \|\lambda u - A_n u\| \, \|u\|.$$

For $f \in L^2(\Omega, \mu)$ and fixed $\lambda > \lambda_0$, set $u_n = R(\lambda ; A_n)f$. From (3.5) and (3.6) we see that the sequences $\{u_n\}$, $\{A_n u_n\}$, and $\{\alpha_n X_j u_n\}$ are bounded in $L^2(\Omega, \mu)$. Therefore, passing to a subsequence if necessary, we may assume that these sequences converge weakly in $L^2(\Omega, \mu)$ to elements u, v, and w_j, respectively. For $\phi \in C_0^\infty(\Omega)$, we have

$$(A_n u_n, \phi) = \frac{1}{n}(u_n, E^*\phi) + (u_n, \alpha_n^2 A^*\phi) + \frac{1}{2}(u_n, (A^*\alpha_n^2 - \alpha_n^2 A^* 1)\phi).$$

If n is sufficiently large, the right hand side reduces to

$$\frac{1}{n}(u_n, E^*\phi) + (u_n, A^*\phi)$$

since $\alpha_n^2 = 1$ on the support of ϕ and the support of $A^*\alpha_n^2 - \alpha_n^2 A^* 1$ is contained in $\{x \in \Omega : n \leqq |p(x)| \leqq 2n\}$. Thus,

$$(v, \phi) = \lim_{n \to \infty} (A_n u_n, \phi) = (u, A^*\phi)$$

so that $v = Au$ as a distribution. Consequently, $u \in D_1(A)$ and $\lambda u - Au = f$. Moreover, from

$$(\alpha_n X_j u_n, \phi) = -(u_n, \alpha_n(X_j\phi - (X_j^* 1)\phi)) - (u_n, \phi X_j \alpha_n)$$

we have $w_j = X_j u \in L^2(\Omega, \mu)$ for $j = 1, \cdots, r$. This completes the proof.

THEOREM 3.3. *Let A in* (1.2) *satisfy conditions* (1.3), (1.4), *and* (1.5). *Then A with domain $D_0(A)$ is the infinitesimal generator of a semi-group* $\{\exp(tA) : t \geqq 0\}$

of class (C_0) *in* $L^2(\Omega, \mu)$ *satisfying the condition* $\|\exp(tA)\| \leqq \exp(t\lambda_0)$ *for* $t \geqq 0$. *Moreover, for* $u \in D_0(A)$ *we have* $X_j u \in L^2(\Omega, \mu)$, $j = 1, \cdots, r$.

Proof. For $\lambda > \lambda_0$, Proposition 3.1 and (1.3) imply that the estimate

$$\|\lambda\phi - A\phi\| \geqq (\lambda - \lambda_0)\|\phi\|$$

holds in $C_0^\infty(\Omega)$ and therefore in $D_0(A)$. Thus, by the Hille–Yosida–Phillips theorem (see viz. [8, p. 248]), it suffices to prove that for $f \in C_0^\infty(\Omega)$ and $\lambda > \lambda_0$ the solution u of the equation $\lambda u - Au = f$ given in Theorem 3.2 belongs to $D_0(A)$.

Consider the sequence $\{\alpha_n u\}$, where α_n is defined in (3.3) and the function p used there satisfies (1.4). From (1.5) we see that $\lambda - A$ is hypoelliptic. Thus, since $f \in C_0^\infty(\Omega)$, by hypoellipticity we have $u \in C^\infty(\Omega)$ so that $\alpha_n u \in C_0^\infty(\Omega)$. Clearly, s-$\lim_{n\to\infty} \alpha_n u = u$ and s-$\lim_{n\to\infty} \alpha_n Au = Au$. Observe that we can write

$$A(\alpha_n u) = \alpha_n Au + 2\sum_1^r (X_j\alpha_n)X_j u + \sum_1^r u X_j^2 \alpha_n + u X_0 \alpha_n. \tag{3.7}$$

From

$$X_j\alpha_n = \frac{1}{n}\alpha'(p/n)X_j p,$$

$$X_j^2\alpha_n = \frac{1}{n}\alpha'(p/n)X_j^2 p + \frac{1}{n^2}\alpha''(p/n)(X_j p)^2,$$

and (1.4) we see that the terms $X_j\alpha_n$, $j = 0, \cdots, r$, and $X_j^2\alpha_n$, $j = 1, \cdots, r$, are uniformly bounded with support in $\{x \in \Omega: n \leqq |p(x)| \leqq 2n\}$. Therefore, since $X_j u \in L^2(\Omega, \mu)$, $j = 1, \cdots, r$, the last three terms in (3.7) converge to zero in $L^2(\Omega, \mu)$. Thus, s-$\lim_{n\to\infty} A(\alpha_n u) = Au$ so that $u \in D_0(A)$.

THEOREM 3.4. *Suppose* A *in* (1.2) *satisfies conditions* (1.3), (1.4), (1.5) *and* (1.6). *Then* $D_0(A) = D_1(A)$.

Proof. Observe that

$$A^* = \sum_1^r X_j^2 - \left\{2\sum_1^r (X_j^* 1)X_j + X_0\right\} + A^* 1$$

is of the form (1.2) and satisfies (1.3). Using (1.6) it follows that the vector fields $2\sum_1^r (X_j^* 1)X_j - X_0, X_1, \cdots, X_r$ satisfy condition (1.4). Let $\mathscr{L}(A)$ denote the Lie algebra over the reals generated by $\{X_0, \cdots, X_r\}$ and let $\mathscr{M}(A)$ denote the Lie algebra over $C^\infty(\Omega)$ generated by $\{X_0, \cdots, X_r\}$. Without assuming (1.5), it is easy to check that the rank of the natural evaluation maps ev_{x^0} from either $\mathscr{L}(A)$ or $\mathscr{M}(A)$ to the tangent space at x^0, $T_{x^0}(\Omega)$, are the same. Since $\mathscr{M}(A^*) = \mathscr{M}(A)$, using (1.5) we have

$$\mathrm{ev}_{x^0}(\mathscr{L}(A^*)) = \mathrm{ev}_{x^0}(\mathscr{M}(A^*)) = \mathrm{ev}_{x^0}(\mathscr{M}(A)) = \mathrm{ev}_{x^0}(\mathscr{L}(A)) = T_{x^0}(\Omega).$$

Thus, A^* satisfies all the conditions in Theorem 3.3 so that for $\lambda > \lambda_0$, $(\lambda - A^*)D_0(A) = L^2(\Omega, \mu)$.

Given $u \in D_1(A)$ there is $v \in D_0(A)$ for which $\lambda v - Av = \lambda u - Au$. Therefore, $w = v - u \in \mathrm{Ker}\,(A)$. But, $\mathrm{Ker}\,(A) = 0$ since the range of $\lambda - A^*$ on $D_0(A^*)$ is $L^2(\Omega, \mu)$. Hence, $w = 0$ so that $u = v \in D_0(A)$.

In order to illustrate the flexibility of condition (1.4) over a fixed set of growth conditions, several examples are given for which the conclusions of Theorems 3.3 and 3.4 hold. Note that in each case we choose a different proper function.

Example 3.5. For any positive integer n, let $A = \partial^2/\partial x^2 + x^n(\partial/\partial y)$ on $(R^2, dx\,dy)$. The bracket between $\partial/\partial x$ and $x^n(\partial/\partial y)$ reduces the exponent of x in the second of these vectors by one. It follows from taking brackets n times that Hörmander's condition (1.5) is satisfied. Now take $p = x^{2n} + y^2$. A computation shows that (1.4) and (1.6) hold.

Example 3.6. Let $A = (e^{u(y)}(\partial/\partial x))^2 + \partial/\partial y$, where $|u'(y)| \leqq M$. This is slightly more general than the above as far as growth at infinity is concerned. Take $p = e^{2u} + x^2 + y^2$.

Example 3.7. In R^{n+1}, let $A = \sum_1^n (a_i(\partial/\partial x_i))^2 + a_0(\partial/\partial x_0)$ where for some constant K we have

$$\text{(i)} \qquad |a_i(x)| \leqq K|x|, \qquad i = 0, \cdots, n,$$

$$\text{(ii)} \qquad \left|\frac{\partial a_i(x)}{\partial x_i}\right| \leqq K, \qquad i = 1, \cdots, n,$$

$$\text{(iii)} \qquad \sum_1^n \frac{\partial}{\partial x_i}\left\{a_i \frac{\partial a_i}{\partial x_i}\right\} - \frac{\partial a_0}{\partial x_0} \leqq K.$$

Assuming condition (1.5), it is enough to take $p = \sum_0^n x_i^2$.

Example 3.8. We include this example to illustrate the kind of pathology that may occur if the coefficients are chosen appropriately. For arbitrary $\beta \in R^1$, let

$$A_\beta = \frac{\partial}{\partial x^2} + (1+x^2)^\beta (2 + \sin(ye^x))\frac{\partial}{\partial y} + \frac{1}{2}(1+x^2)^\beta e^x \cos(e^x y).$$

Here, we take $p = (1+x^2)^{2\beta} + y^2$. Note the wild oscillation of the termination coefficient.

REFERENCES

[1] L. HÖRMANDER, *Hypoelliptic second order differential equations*, Acta Math., 119 (1968), pp. 147–171.

[2] T. IKEBE AND T. KATO, *Uniqueness of the self-adjoint extension of singular elliptic differential operators*, Arch. Rational Mech. Anal., 9 (1962), pp. 77–92.

[3] T. KATO, *Schrödinger operators with singular potentials*, Israel J. Math., 13 (1972), pp. 135–148.

[4] M. SCHECHTER, *Spectra of Partial Differential Operators*, Elsevier–North Holland, Amsterdam, 1971.

[5] ———, *Essential self-adjointness of the Schrödinger operator with magnetic vector potential*, preprint.

[6] B. SIMON, *Essential self-adjointness of Schrödinger operators with positive potentials*, Math. Ann., 201 (1973), pp. 211–220.

[7] ———, *Schrödinger operators with singular magnetic vector potentials*, Math. Z., 131 (1973), pp. 361–370.

[8] K. YOSIDA, *Functional Analysis*, 3rd ed., Springer-Verlag, New York, 1971.

SIAM J. MATH. ANAL.
Vol. 8, No. 6, November 1977

NEW RELATIONS BETWEEN TWO TYPES OF BESSEL FUNCTION INTEGRALS*

HENRY E. FETTIS†

Abstract. It is shown that, in certain special cases, the incomplete Lipschitz–Hankel integrals are related to incomplete integrals of Hankel–Nicholson type.

Integrals of the form

$$F_{p,q}(x,z)=\int^{x} e^{-at}t^{q}Z_{p}(zt)\,dt \tag{1}$$

where $Z_p(t)$ may be any one of the Bessel functions, fall under the general category of incomplete Lipschitz–Hankel integrals (see [1]). These integrals were studied in some detail by Luke (see [4]), where recursive relations were developed connecting integrals in which p and q differed by integers. Later, Ng [6] extended Luke's results to the more general case

$$\int^{x} e^{-at^{n}}t^{q}Z_{p}(zt)\,dt. \tag{2}$$

In the present paper, we show that the integrals defined by (1) are related to incomplete integrals of Hankel–Nicholson (or Sonine) type, i.e., to integrals having the general form

$$\int^{z} \frac{t^{\alpha}J_{p}(xt)\,dt}{(a^{2}+t^{2})^{\beta}}. \tag{3}$$

To obtain the desired connection between these two types of integrals, we use the following relations which are essentially those given in [6] and [4]:

$$F_{p,q}(x,z)=-\frac{1}{a}e^{-ax}x^{q}Z_{p}(xz)+\frac{p+q}{a}F_{p,q-1}(x,z)-\frac{\beta z}{a}F_{p+1,q}(x,z), \tag{4a}$$

$$F_{p,q}(x,z)=\frac{\alpha}{z}e^{-ax}x^{q}Z_{p+1}(xz)+\alpha\frac{a}{z}F_{p+1,q}(x,z)+\alpha\frac{p-q+1}{z}F_{p+1,q-1}(x,z), \tag{4b}$$

where the parameters α and β are defined as follows:

$$\alpha=\begin{cases}1 & \text{for } J_p,\ Y_p \text{ and } I_p,\\ -1 & \text{for } K_p;\end{cases}$$
$$\beta=\begin{cases}1 & \text{for } J_p,\ Y_p \text{ and } K_p,\\ -1 & \text{for } I_p.\end{cases} \tag{5}$$

* Received by the editors April 9, 1976, and in revised form June 1, 1976.
† 1885 California, Apartment 62, Mountain View, California 94041.

By eliminating $F_{p+1,q}(x, z)$ between (4a) and (4b), we obtain the following result:

$$(a^2+\alpha\beta z^2)F_{p,q}(x,z)=e^{-ax}x^q[\beta z Z_{p+1}(xz)-aZ_p(x,z)] \\ +a(p+q)F_{p,q-1}(x,z)+\beta z(p-q+1)F_{p+1,q-1}(x,z) \tag{6}$$

which, in the special case $p=-q$ reduces to

$$(a^2+\alpha\beta z^2)F_{p,-p}(x,z)=e^{-ax}x^{-p}[\beta z Z_{p+1}(xz)-aZ_p(xz)] \\ +\beta z(2p+1)F_{p+1,-p-1}(x,z). \tag{7}$$

If, now, the above equation is multiplied by z^p, and use is made of the easily verified relation

$$\frac{\partial}{\partial z}[z^pF_{p,q}(x,z)]=\alpha z^pF_{p-1,q+1}(x,z) \tag{8}$$

we arrive at the following

$$(a^2+\alpha\beta z^2)\frac{\partial}{\partial z}[z^pF_{p,-p}(x,z)]+\alpha\beta z(1-2p)[z^pF_{p,-p}(x,z)] \\ =\alpha e^{-ax}x^{-p+1}[\beta z^{p+1}Z_p(xz)-az^pZ_{p-1}(xz)] \tag{9}$$

or, equivalently,

$$(a^2+\alpha\beta z^2)^{1/2+p}\left[\frac{\partial}{\partial z}z^p(a^2+\alpha\beta z^2)^{1/2-p}F_{p,-p}(x,z)\right] \\ =\alpha e^{-ax}x^{-p+1}[\beta z^{p+1}Z_p(xz)-az^pZ_{p-1}(xz)]. \tag{10}$$

In an entirely analogous manner, we find

$$(a^2+\alpha\beta z^2)^{1/2-p}\frac{\partial}{\partial z}[z^{-p}(a^2+\alpha\beta z^2)^{1/2+p}F_{p,p}(x,z)] \\ =\alpha\beta e^{-ax}x^{p+1}[z^{-p+1}Z_p(xz)+a\alpha z^{-p}Z_{p+1}(xz)]. \tag{11}$$

By assigning definite limits to the integrals $F_{p,-p}$ and $F_{p,p}$ and specifying the type of Bessel function, integration of (10) and (11) between the limits 0 and z yields relations between these integrals and those of Hankel–Nicholson type. For example, if $Z_p \equiv J_p$, and Re $p>0$, the limits in the former ones may be taken to be 0 and x. Integration of (10) then gives

$$z^p(a^2+z^2)^{1/2-p}\int_0^x e^{-at}t^{-p}J_p(zt)\,dt \\ =\frac{a}{2^{p-1}\Gamma(p)}\int_0^z\frac{t^{2p-1}\,dt}{(a^2+t^2)^{1/2+p}} \\ +e^{-ax}x^{1-p}\left[\int_0^z\frac{t^{p+1}J_p(xt)\,dt}{(a^2+t^2)^{1/2+p}}-a\int_0^z\frac{t^pJ_{p-1}(xt)\,dt}{(a^2+t^2)^{1/2+p}}\right], \tag{12}$$

Re $(p)>0$.

Similarly, from (11),

$$z^{-p}(a^2+z^2)^{1/2+p}\int_0^x e^{-at}t^pJ_p(zt)\,dt$$

$$=\frac{|a|^{2p+1}\int_0^x e^{-at}t^{2p}\,dt}{2^p\Gamma(1+p)} \tag{13}$$

$$+e^{-ax}x^{1+p}\left[\int_0^z \frac{t^{1-p}J_p(xt)\,dt}{(a^2+t^2)^{1/2-p}}+a\int_0^z \frac{t^{-p}J_{p+1}(xt)\,dt}{(a^2+t^2)^{1/2-p}}\right],$$

$\mathrm{Re}\,(p)\geqq 0$.

By replacing a with $-a$ in (12) and combining the result with (12), the incomplete Hankel–Nicholson integrals can be expressed in terms of the Lipschitz–Hankel integrals:

$$x^{1-p}\int_0^z \frac{t^{p+1}J_p(xt)\,dt}{(a^2+t^2)^{1/2+p}}=z^p(a^2+z^2)^{1/2-p}\int_0^x t^{-p}J_p(zt)\cosh a(x-t)\,dt$$

$$-\frac{a}{2^{p-1}\Gamma(p)}\sinh ax\int_0^z \frac{t^{2p-1}\,dt}{(a^2+t^2)^{1/2+p}}, \tag{14}$$

$$\mathrm{Re}\,(p)>0,$$

$$ax^{1-p}\int_0^z \frac{t^pJ_{p-1}(xt)\,dt}{(a^2+t^2)^{1/2+p}}=-z^p(a^2+z^2)^{1/2-p}\int_0^x t^{-p}J_p(zt)\sinh a(x-t)\,dt$$

$$+\frac{a}{2^{p-1}\Gamma(p)}\cosh ax\int_0^z \frac{t^{2p-1}\,dt}{(a^2+t^2)^{1/2+p}}, \tag{15}$$

$$\mathrm{Re}\,(p)>0.$$

Similar relations may be obtained by applying the same procedure to (13).

Analogous relations involving the modified functions or functions of second kind can be derived in a similar manner. In particular for $p=0$ and $a>0$, the following typical set of relations can be found:

$$(a^2+z^2)^{1/2}\int_0^x e^{-at}J_0(zt)\,dt=1-e^{-ax}+x\,e^{-ax}\left\{\int_0^z \frac{J_0(xt)t\,dt}{(a^2+t^2)^{1/2}}\right.$$

$$\left.+a\int_0^z \frac{J_1(xt)\,dt}{(a^2+t^2)^{1/2}}\right\}, \tag{16}$$

$$(a^2-z^2)^{1/2}\int_0^x e^{-at}I_0(zt)\,dt=1-e^{-ax}-x\,e^{-ax}\left\{\int_0^z \frac{I_0(xt)t\,dt}{(a^2-t^2)^{1/2}}\right.$$

$$\left.+a\int_0^z \frac{I_1(xt)\,dt}{(a^2-t^2)^{1/2}}\right\}, \qquad z<a, \tag{17}$$

$$(z^2-a^2)^{1/2}\int_0^x e^{-at}K_0(zt)\,dt=x\,e^{-ax}\left\{a\int_z^\infty \frac{K_1(xt)\,dt}{(t^2-a^2)^{1/2}}-\int_z^\infty \frac{K_0(xt)t\,dt}{(t^2-a^2)^{1/2}}\right\},$$

$$z>a, \tag{18}$$

$$(19)\quad (a^2-z^2)^{1/2}\int_0^x e^{-at}K_0(zt)\,dt = x\,e^{-ax}\left\{\int_z^a \frac{K_0(xt)t\,dt}{(a^2-t^2)^{1/2}} - a\int_z^a \frac{K_1(xt)\,dt}{(a^2-t^2)^{1/2}}\right\},$$

$$z<a,$$

$$(20)\quad (z^2-a^2)^{1/2}\int_0^x e^{-at}I_0(zt)\,dt = x\,e^{-ax}\left\{\int_a^z \frac{I_0(xt)t\,dt}{(t^2-a^2)^{1/2}} + a\int_a^z \frac{I_1(xt)\,dt}{(t^2-a^2)^{1/2}}\right\},$$

$$z>a.$$

The incomplete Lipschitz–Hankel integrals of order zero are also known as "Schwarz integrals" and find application in the theory of unsteady aerodynamics [7]. Those appearing on the right side of the (16–20) are of a type encountered in problems in radiation theory [2].

Appendix. The first integral appearing on the right side of (12) and in subsequent relations can be expressed in terms of the hypergeometric function:

$$\int_0^z \frac{t^{2p+1}dt}{(a^2+t^2)^{1/2+p}} = \frac{1}{2a^2p}z^{2p}(a^2+z^2)^{1/2-p}\,{}_2F_1\left(\frac{1}{2}, 1; 1+p; -\frac{z^2}{a^2}\right).$$

Letting $x\to\infty$, we obtain the known result [5]:

$$\int_0^\infty e^{-at}t^{-p}J_p(zt)\,dt = \left(\frac{z}{2}\right)^p \frac{1}{a\Gamma(1+p)}\,{}_2F_1\left(\frac{1}{2}, 1; 1+p; -\frac{z^2}{a^2}\right).$$

In addition, the following special results are obtained by letting $z\to\infty$ in (12):

$$\int_0^\infty \frac{t^{p+1}J_p(xt)\,dt}{(a^2+t^2)^{1/2+p}} = a\int_0^\infty \frac{t^pJ_{p-1}(xt)\,dt}{(a^2+t^2)^{1/2+p}}, \qquad a>0,$$

$$\int_0^\infty \frac{t^{p+1}J_p(xt)\,dt}{(a^2+t^2)^{1/2+p}} = \frac{\sqrt{\pi}}{2^p}\frac{x^{p-1}e^{ax}}{\Gamma(\frac{1}{2}+p)}, \qquad a<0.$$

See, e.g. [3, art. 6.565; (3)].

The integrals on the right side of (13) are convergent for $z\to\infty$, provided

$$0\leqq \mathrm{Re}\,(p)<\tfrac{1}{2}.$$

This leads to the following, perhaps less familiar, result:

$$e^{-ax}x^{1+p}\left\{\int_0^\infty \frac{t^{1-p}J_p(xt)\,dt}{(a^2+t^2)^{1/2-p}} + a\int_0^\infty \frac{t^{-p}J_{p+1}(xt)\,dt}{(a^2+t^2)^{1/2-p}}\right\}$$

$$= \frac{1}{2^p\Gamma(1+p)}\Gamma[(1+2p), ax], \qquad a>0,$$

relating the complete Hankel–Nicholson integrals to the incomplete gamma function.

REFERENCES

[1] M. M. AGREST AND M. S. MAKSIMOV, *Theory of Incomplete Cylindrical Functions and Their Applications*, Springer, New York/Berlin, 1971.

[2] CARL-ERIK FROBERG AND HANS WILHELMSSON, *Table of the function* $F(a, b) = \int_0^\alpha J_1(x)(x^2+b^2)^{-1/2}\,dx$, Kungl. Fysiografiska Sallskapets I Lund Forhandlingar, 27 (1957), pp. 202–215.

[3] I. S. GRADSTEIN AND I. S. RYZIK, *Table of Integrals, Series and Products*, Academic Press, New York, 1965.

[4] Y. L. LUKE, *Integrals of Bessel Functions*, McGraw-Hill, New York, 1962.

[5] W. MAGNUS AND F. OBERHETTINGER, *Formeln und satze fur die speziellen Funktionen der mathematischen Physik*, Springer, Berlin, 1948.

[6] E. W. NG, *Recursive formulae for the computation of certain integrals of Bessel functions*, J. Math. and Phys., 46 (1967), pp. 223–224.

[7] L. SCHWARZ, *Untersuchen Einiger mit der Zylinderfunktionen nuller ordung verwandter funktionen*, Luftfahrtforshung, 20 (1944), pp. 341–372.

SIAM J. MATH. ANAL.
Vol. 8, No. 6, November 1977

GENERALIZATIONS OF FARKAS' THEOREM*

B. D. CRAVEN† AND J. J. KOLIHA†

Abstract. A unified treatment is given of generalizations of Farkas' theorem on linear inequalities to arbitrary convex cones and to dual pairs of real vector spaces of arbitrary dimension. Various theorems for locally convex spaces readily follow. The results are applied to duality and converse duality theory for linear programming and to a generalization of the Kuhn–Tucker theorem, both of these in spaces of arbitrary dimension and with inequalities involving arbitrary convex cones.

1. Introduction. Farkas' theorem on systems of linear inequalities [11] (see also [10, Thm. 4]) has been generalized to systems involving polyhedral cones [1], and to arbitrary cones in locally convex spaces [12]. A unified theory is presented here, based on arbitrary cones and dual pairs of real vector spaces. A necessary and sufficient condition is obtained (Theorem 2) for the solvability of a linear equation $Ax = b$ by a vector x in a given convex cone. A related necessary and sufficient condition (Theorem 1) is obtained as an inclusion relation between two cones. Although these results are closely related to the Hahn–Banach separation theorem, the proofs of Theorems 1 and 2 (based on the approach of Bourgin [6]) require only a finite-dimensional separation result, of which a simple proof is given in Lemma 1.

When these results are applied to continuous maps and real locally convex spaces, it is found that each theorem appears in two versions, standing in a duality relation, and not generally deducible from each other. Theorems 5 and 6 are a typical such pair; Theorem 6 was previously known, but apparently not Theorem 5. Such a theorem holds if and only if a certain cone is closed, in an appropriate topology. This fact was stated in [2], for finite dimensions only, and is here extended to any dimension. The equivalence of this hypothesis to an alternative hypothesis given in [1] for finite dimensions only is proved in Theorem 7 for Fréchet spaces.

The main application of Farkas' theorem and its various generalizations is to Lagrangian and duality theory in mathematical programming. The present results are applied in Theorems 8 and 9, to duality and converse duality theory for linear programming, in spaces of arbitrary dimension. In Theorem 10, the Kuhn–Tucker theorem is generalized to arbitrary spaces; an appropriate generalization of the Kuhn–Tucker constraint qualification is formulated.

Some further results are given for finite-dimensional problems in Theorems 11 to 14. These concern the solvability of linear operator equations by operators which map one given cone into another.

2. Notation and preliminaries. For any real vector space X, X^* denotes the *algebraic dual* of X, i.e., the set of all real-valued linear functionals on X, and $\langle x, x^*\rangle$ denotes the evaluation of $x^* \in X^*$ at the point $x \in X$. Let $A: X \to Y$ be a linear map between real vector spaces X, Y. The *algebraic adjoint* A^* of A is the (linear) map from Y^* to X^* satisfying the equation

$$\langle x, A^* y^*\rangle = \langle Ax, y^*\rangle, \qquad \forall x \in X, \quad \forall y^* \in Y^*.$$

* Received by the editors July 17, 1975, and in revised form May 3, 1976.

† Mathematics Department, University of Melbourne, Parkville, Victoria 3052, Australia.

Let X^+ be a subspace of $X^\#$ separating points of X. (This means that for any two distinct points x_1, x_2 in X there is an $x^\# \in X^+$ satisfying $\langle x_1, x^\# \rangle \neq \langle x_2, x^\# \rangle$.) Then X can be regarded as a subspace of $(X^+)^\#$ when we identify each $x \in X$ with the linear functional $x^+ \mapsto \langle x, x^+ \rangle$, $\forall x^+ \in X^+$, and observe that X separates points of X^+. The pair $\langle X, X^+ \rangle$ is called a *dual pair* [15, p. 32].

Given a dual pair $\langle X, X^+ \rangle$, we denote by $\sigma(X, X^+)$ the *weak topology* on X, i.e., the coarsest topology on X for which the functionals $x^+ \in X^+$ are continuous. The weak topology $\sigma(X^+, X)$ on X^+ is defined symmetrically. It is known that X with the weak topology $\sigma(X, X^+)$ is a locally convex space (= real locally convex Hausdorff topological vector space), and the weak topology is generated by any family of seminorms $\|x\|_\alpha = |\langle x, x_\alpha^+ \rangle|$, where $\{x_\alpha^+\}$ is a Hamel basis of X^+ [16, p. 124]. In the following we often abbreviate "weak topology", "closed in the weak topology", etc., to "w-topology", "w-closed", etc.

Let $\langle X, X^+ \rangle$ and $\langle Y, Y^+ \rangle$ be two dual pairs. A map $f: X \to Y$ is *w-continuous* (with respect to $\langle X, X^+ \rangle$ and $\langle Y, Y^+ \rangle$) if it is a continuous map between topological spaces $(X, \sigma(X, X^+))$ and $(Y, \sigma(Y, Y^+))$. A necessary and sufficient condition that a linear map $A: X \to Y$ is w-continuous is that $A^\#(Y^+) \subseteq X^+$ ([16], p. 128). In this case the restriction A^+ of $A^\#$ to Y^+ is w-continuous (with respect to $\langle Y^+, Y \rangle$ and $\langle X^+, X \rangle$). The map A^+ is called the *adjoint* of A with respect to the dual pairs $\langle X, X^+ \rangle$ and $\langle Y, Y^+ \rangle$.

A nonempty set K in the real vector space X is called a *convex cone* if

$$K + K \subseteq K \quad \text{and} \quad \alpha K \subseteq K, \qquad \alpha \geqq 0.$$

If $\langle X, X^+ \rangle$ is a dual pair, the *anticone* K^+ of the convex cone $K \subseteq X$ (with respect to $\langle X, X^+ \rangle$) is defined as the set

$$K^+ = \{z \in X^+ : \langle x, z \rangle \geqq 0, \forall x \in K\}.$$

It is easily seen that K^+ is a w-closed (with respect to $\langle X^+, X \rangle$) convex cone in X^+. We prove the following characterization of anticone:

$$K^+ = \left\{ z \in X^+ : \inf_{x \in K} \langle x, z \rangle > -\infty \right\}. \tag{2.1}$$

Let $\mu(z) = \inf \{\langle x, z \rangle : x \in K\} > -\infty$. If $x \in K$, then $\alpha x \in K$ for all $\alpha > 0$, and $\langle x, z \rangle = \alpha^{-1} \langle \alpha x, z \rangle \geqq \alpha^{-1} \mu(z)$. Passing to the limit as $\alpha \to \infty$, we get $\langle x, z \rangle \geqq 0$ for all $x \in K$.

The *second anticone* K^{++} of a convex cone $K \subseteq X$ (with respect to the dual pair $\langle X, X^+ \rangle$) is defined to be the anticone $(K^+)^+$ of K^+ (with respect to $\langle X^+, X \rangle$).

If X is a locally convex space and X' its topological dual, then the anticone K^+ of a convex cone K (with respect to $\langle X, X^+ \rangle$) coincides with the dual cone K^* of K [16, p. 218]. If X' is regarded as a locally convex space under the strong topology [16, p. 140], then K^* is a strongly closed convex cone. The second dual cone $K^{**} = (K^*)^*$ is in the second dual X'' of X. Under the natural embedding of X into X'' we have

$$K^{++} = (K^*)^+ \subseteq K^{**}.$$

It can be shown that if K is a strongly closed convex cone in X, then

$$K^{**} \cap X = K.$$

The concept of anticone is introduced to provide symmetry which is missing in the case of dual cones. In fact, it will be shown presently that

$$K^{++} = K$$

for any w-closed convex cone K in X (cf. Lemma 3).

In the rest of the section, $\langle X, X^+\rangle$ is a certain real dual pair, and K a convex cone in X. The main result of this section is Lemma 4, the main tool for its proof is the separation theorem for convex sets given in Lemma 1. We give a completely elementary proof of the separation theorem which does not invoke the full force of Mazur's theorem for topological vector spaces [16, p. 46], and which is essentially an adaptation of Bourgin's argument employed in [6, p. 643].

LEMMA 1: *Let S be a w-closed convex subset of X, and let $a \in X\backslash S$. Then there is an $a^+ \in X^+$ such that*

$$\inf_{x\in S} \langle x-a, a^+\rangle > 0. \tag{2.2}$$

Proof. There exist $z_1, \cdots, z_n \in X^+$ and an $\varepsilon > 0$ such that

$$|\langle x-a, z_i\rangle| \geqq \varepsilon, \qquad \forall x \in S, \quad i = 1, \cdots, n.$$

Define a map $\psi: X \to \mathbb{R}^n$ *by*

$$\psi(x) = (\langle x, z_1\rangle, \cdots, \langle x, z_n\rangle), \qquad \forall x \in X.$$

Let F denote the $\mathbb{R}^n$ closure of $\psi(S)$, let $p = \psi(a)$, and let q be the point in F with minimum (Euclidean) distance from p. It is easily checked that F is convex, that the distance from p to F is not less than ε, and that

$$(u-q, q-p) \geqq 0, \qquad \forall u \in F,$$

where round brackets refer to the standard inner product in $\mathbb{R}^n$. Define $a^+ = \sum_{i=1}^n \alpha_i z_i$, where $(\alpha_1, \cdots, \alpha_n) = q - p$. Then

$$\langle x-a, a^+\rangle = (\psi(x)-p, q-p) \geqq (q-p, q-p) \geqq \varepsilon^2$$

for all $x \in S$. □

The following three lemmas will be needed in the next section.

LEMMA 2. *Let K be a w-closed convex cone in X, and let $a \in X\backslash K$. Then there is an $a^+ \in K^+$ such that $\langle a, a^+\rangle < 0$.*

Proof. Let $a^+ \in X^+$ be a functional satisfying (2.2); from (2.1) it follows that $a^+ \in K^+$. Inequality (2.2) then implies that $\langle a, a^+\rangle < \langle 0, a^+\rangle = 0$ since $0 \in$ K. □

LEMMA 3. *Let K be a convex cone in X with w-closure $\bar{K}$. Then*

$$K^{++} = \bar{K}.$$

Proof. $\bar{K}$ is a w-closed convex cone in X with $\bar{K}^+ = K^+$. From the definition of anticone we deduce $\bar{K} \subseteq K^{++}$. *If* $a \in X\backslash\bar{K}$, Lemma 2 guarantees the existence of an $a^+ \in K^+$ with $\langle a, a^+\rangle < 0$; this shows that $a \notin K^{++}$. □

LEMMA 4. *Let P and Q be convex cones in X with P w-closed. Then*

$$P^+ \subseteq Q^+ \Leftrightarrow P \supseteq Q.$$

Proof. From Lemma 3 and the definition of anticone,

$$P \supseteq Q \Rightarrow P^+ \subseteq Q^+ \Rightarrow P^{++} \supseteq Q^{++}$$
$$\Rightarrow \bar{P} \supseteq \bar{Q} \Rightarrow P \supseteq \bar{Q} \supseteq Q$$

since P is w-closed. □

3. Farkas' theorems for dual pairs. The original Farkas' theorem [11] dating back to 1902 gives a necessary and sufficient condition for the solvability of the real linear system

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \qquad x_j \geqq 0 \quad \text{for all } j, \quad 1 \leqq i \leqq m.$$

The result has been since generalized by many authors (e.g. [1], [2], [3], [7], [12]) to give necessary and sufficient conditions for the solvability of the linear system

$$Ax = b, \qquad x \in S, \tag{3.1}$$

where $A: X \to Y$ is a linear map, and S a convex cone in X. If the cone S is such that $S \cap (-S) = \{0\}$, the relation

$$x \leqq y \Leftrightarrow y - x \in S$$

defines a partial order on X. In this case, any x satisfying (3.1) is a positive solution ($x \geqq 0$) of the linear equation $Ax = b$.

In the present paper we adopt a new, more encompassing approach to the theorem: Instead of regarding Farkas' theorem as a criterion for the *positive solvability* of a linear equation, we view it as a criterion for a *cone inclusion*. The formulation in the setting of dual pairs instead of locally convex spaces is a vital part of the generalization: It yields two important results for locally convex spaces, neither of which can be deduced from the other (Theorems 3 and 4).

THEOREM 1 (Two cones theorem). *Let* $\langle X, X^+\rangle$, $\langle Y, Y^+\rangle$ *and* $\langle Z, Z^+\rangle$ *be real dual pairs, let* $S \subseteq X$ *and* $T \subseteq Z$ *be convex cones, and let* $A: X \to Y$, $B: Z \to Y$ *be* w*-continuous linear maps. If* $A(S)$ *is* w*-closed, the following conditions on* B *are equivalent*:

(a) $B(T) \subseteq A(S)$.

(b) $A^+y^+ \in S^+ \Rightarrow B^+y^+ \in T^+$.

Proof. First we observe that $A(S)$ and $B(T)$ are convex cones in Y. Let $y^+ \in Y^+$. Then

$$A^+y^+ \in S^+ \Leftrightarrow (\forall y \in S)\, 0 \leqq \langle y, A^+y^+\rangle = \langle Ay, y^+\rangle$$
$$\Leftrightarrow y^+ \in A(S)^+,$$

where $A(S)^+$ denotes $(A(S))^+$. Similarly

$$B^+y^+ \in T^+ \Leftrightarrow y^+ \in B(T)^+.$$

Therefore

$$\text{(b)} \Leftrightarrow A(S)^+ \subseteq B(T)^+ \Leftrightarrow \text{(a)}$$

by Lemma 4 since $A(S)$ is w-closed. □

Let us consider the special case of the preceding theorem when $Z = Z^+ = \mathbb{R}$, and T is the cone $\mathbb{R}_+$ of nonnegative real numbers. In this case each linear map $B: \mathbb{R} \to Y$ is uniquely represented in the form $B\xi = \xi b$ $(\forall \xi \in \mathbb{R})$, where $b \in Y$. Since $\mathbb{R}_+$ is its own anticone with respect to the dual pair $\langle \mathbb{R}, \mathbb{R} \rangle$, the cone $B(T)$ is the ray $\{\lambda b : \lambda \geqq 0\}$, and

$$B(T) \subseteq A(S) \Leftrightarrow b \in A(S). \tag{3.2}$$

Next,

$$\langle \xi, B^+ y^+ \rangle = \xi \langle b, y^+ \rangle \qquad \forall \xi \in \mathbb{R}, \quad \forall y^+ \in Y^+.$$

Hence

$$B^+ y^+ \in T^+ \Leftrightarrow \langle b, y^+ \rangle \geqq 0. \tag{3.3}$$

This leads to the following result.

THEOREM 2 (Generalized Farkas theorem). *Let $\langle X, X^+ \rangle$, $\langle Y, Y^+ \rangle$ be real dual pairs, let S be a convex cone in X, and let $A : X \to Y$ be a w-continuous linear map. If $A(S)$ is w-closed, the following are equivalent conditions on $b \in Y$:*

(a) The equation $Ax = b$ has a solution $x \in S$.

(b) $A^+ y^+ \in S^+ \Rightarrow \langle b, y^+ \rangle \geqq 0$.

Conversely, the equivalence of (a) *and* (b) *implies that $A(S)$ is w-closed.*

Proof. Let $A(S)$ be w-closed. The equivalence of (a) and (b) follows from Theorem 1, (3.2) and (3.3) on setting $Z = Z^+ = \mathbb{R}$, $T = \mathbb{R}_+$ and $B\xi = \xi b$ $(\forall \xi \in \mathbb{R})$ in Theorem 1.

Assume that the conditions (a) and (b) are equivalent. If $\{x_\alpha\}$ is a net in S such that $\{Ax_\alpha\}$ is convergent to b in the w-topology, then

$$\langle b, y^+ \rangle = \lim_\alpha \langle Ax_\alpha, y^+ \rangle = \lim_\alpha \langle x_\alpha, A^+ y^+ \rangle \geqq 0$$

whenever $A^+ y^+ \in S^+$. The implication (b) $\Rightarrow$ (a) now guarantees that $b \in A(S)$. □

4. Farkas' theorems for locally convex spaces. Let X be a locally convex Hausdorff space with the topological dual X' [16, p. 48]. The original topology on X is called the *strong topology*, $\sigma(X, X')$ is called the *weak topology* on X, and $\sigma(X', X)$ is called the *weak* topology* on X'. We recall that every continuous linear map $A: X \to Y$ between locally convex spaces is also weakly continuous. Let $x_\alpha \to 0$ strongly in X; then $\langle x_\alpha, A^\# y' \rangle = \langle Ax_\alpha, y' \rangle \to 0$ for each $y' \in Y'$, i.e., the necessary and sufficient condition $A^\#(Y') \subseteq X'$ for the weak continuity of A is satisfied. The *adjoint* of A with respect to the dual pairs $\langle X, X' \rangle$ and $\langle Y, Y' \rangle$ will be denoted by A'; it is a weak* continuous map. If S is a convex cone in a locally convex space X, S* denotes the *dual cone* of S, that is, the anticone of S with respect to $\langle X, X' \rangle$.

One of the deeper properties of locally convex spaces used in this section is the fact that the strong and the weak closure of any convex set are identical. Unlike the results of § 3 of this paper, the proof of equality of the two closures requires a separation theorem [16, p. 65] based on Mazur's theorem [16, p. 46], and hence on transfinite induction.

The two cones theorem for dual pairs gives rise to two independent results for locally convex spaces. For readers' convenience we include diagrams illustrating the respective theorems.

THEOREM 3. *Let X, Y, Z be locally convex spaces, let $S \subseteq X$, $T \subseteq Z$ be convex cones, and let $A: X \to Y$, $B: Z \to Y$ be strongly continuous linear maps. If $A(S)$ is strongly closed, the following conditions are equivalent*:

(a) $B(T) \subseteq A(S)$.

(b) $A'y' \in S^* \Rightarrow B'y' \in T^*$.

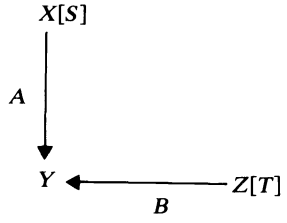


FIG. 1

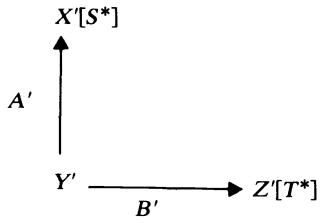


FIG. 1'

THEOREM 4. *Let X, Y, Z be locally convex spaces, let $S \subseteq X$, $T \subseteq Z$ be strongly closed convex cones, and let $C: Y \to X$, $D: Y \to Z$ be strongly continuous linear maps. If $C'(S^*)$ is weak* closed, the following conditions are equivalent*:

(a) $D'(T^*) \subseteq C'(S^*)$.

(b) $Cy \in S \Rightarrow Dy \in T$.

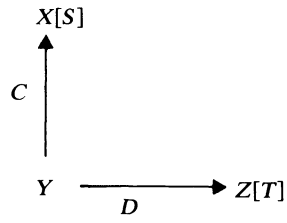


FIG. 2

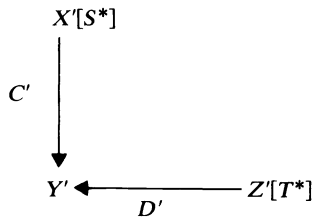


FIG. 2'

Proofs. Theorem 3 is a straightforward application of Theorem 1.

To prove Theorem 4, we make the following replacement in Theorem 1: X, Y, Z, S, T, A, B are replaced by X', Y', Z', S^*, T^*, C', D' in this order. The maps C' and D' are weak* continuous, i.e., w-continuous with respect to the dual pairs $\langle X', X\rangle$, $\langle Y', Y\rangle$ and $\langle Z', Z\rangle$. We have to show that $(C')^+ = C$, $(D')^+ = D$, $(S^*)^+ = S$ and $(T^*)^+ = T$ with respect to the mentioned dual pairs. Now $(C')^+$ is the restriction of $(C')^\#$ to X, i.e.,

$$\langle (C')^+ x, y'\rangle = \langle x, C'y'\rangle = \langle Cx, y'\rangle$$

for $x \in X$ and $y' \in Y'$; hence $(C')^+ = C$. Next, S is strongly closed, and consequently also w-closed with respect to $\langle X, X'\rangle$. According to Lemma 3, $(S^*)^+ = S^{++} = S$. The remaining two equalities are proved similarly. □

It should be observed that in Theorem 3 we do not make the assumption that the cones S, T are closed; this hypothesis, however, is indispensable in Theorem 4 as we need the relations $(S^*)^+ = S$ and $(T^*)^+ = T$.

Specializing the space Z in Theorems 3 and 4 to $\mathbb{R}$ and the cone T to $\mathbb{R}_+$ and applying (3.2), we obtain the following results.

THEOREM 5. *Let X, Y be locally convex spaces, let $S \subseteq X$ be a convex cone in X, and let $A: X \to Y$ be a strongly continuous linear map. If $A(S)$ is strongly closed, the following conditions on $b \in Y$ are equivalent*:

(a) $Ax = b$ *has a solution* $x \in S$.

(b) $A'y' \in S^* \Rightarrow \langle b, y' \rangle \geqq 0$.

Conversely, the equivalence of (a) *and* (b) *implies that $A(S)$ is strongly closed.*

THEOREM 6. *Let X, Y be locally convex spaces, let $S \subseteq X$ be a strongly closed convex cone, and let C: $Y \to X$ be a strongly continuous linear map. If $C'(S^*)$ is weak* closed, the following conditions on $b' \in Y'$ are equivalent*:

(a) $C'x' = b'$ *has a solution* $x' \in S^*$.

(b) $Cy \in S \Rightarrow \langle y, b' \rangle \geqq 0$.

Conversely, the equivalence of (a) *and* (b) *implies that $C'(S^*)$ is weak* closed.*

Berman and Ben-Israel proved the direct part of Theorem 5 in the special case that the spaces X, Y are finite dimensional under the additional hypothesis that the cone S is also closed (cf. [2, Thm. 1]). The direct part of a finite dimensional version of Theorem 2 is given in Theorem 2.4 of [1] with the assumption "$A(S)$ is closed in Y" replaced by "$S + A^{-1}(0)$ is closed in X"; R. A. Abrams (see [5, Lemma 3.1]) proved that these two hypotheses are equivalent in finite dimensional spaces. We prove, in Theorem 7, that this equivalence holds more generally. Let us remark that the results of [1], [2], and [3] are presented for complex vector spaces.

Our Theorem 6 is essentially Hurwicz's generalization of Farkas' theorem (cf. [12, Thm. III.4]); it is hoped that our formulation is more transparent. However, Hurwicz uses the Hahn–Banach separation theorem, and thus relies on transfinite induction, whereas our Theorems 1 and 2 do not. It should be noted that Hurwicz's result does not lead to our Theorem 5. In fact, neither of the Theorems 5 and 6 implies the other (witness the hypotheses about the cone S). Finally, Hurwicz's version of Farkas' theorem requires only that X be locally convex, while Y can be any linear topological vector space.

If $C(Y) = X$ in Theorem 4, and no hypotheses are made about the closure of cones, then Theorem 2.2 of [7] shows that the conditions (a) and (b) are equivalent when X and Y are Fréchet spaces. This statement remains true when X and Y are any real locally convex spaces—see [8, Thm. 1].

The next theorem provides an alternative to the hypothesis made in Theorems 3 and 5, namely that $A(S)$ is closed, in a certain special case.

THEOREM 7: *Let $A: X \to Y$ be a strongly continuous linear map with closed range between Fréchet spaces* [16, p. 49] *X and Y, and let S be a convex cone in X. Then*

$$A(S) \text{ is closed in } Y \Leftrightarrow S + A^{-1}(0) \text{ is closed in } X.$$

Proof. Assume first that $A(X) = Y$. Denote $N = A^{-1}(0)$ and $M = (S+N)^{\sim}$ *where* $^{\sim}$ denotes set complements (in X or Y). Let $y = As$ for some $s \in S$. For any $x \in A^{-1}y$, $A(x - s) = 0$, so $x \in S + N$; hence

$$A^{-1}(A(S)) \subseteq S + N, \tag{4.1}$$

and $x \notin M$, so $y \notin A(M)$. Consequently,

$$A(S) \cap A(M) = \varnothing.$$

Since also

$$Y = A(X) = A((S+N) \cup M) \subseteq A(S) \cup A(M),$$

we have

$$A(S) = A(M)^{\sim} \tag{4.2}$$

Also, from $A(S+N) = A(S)$, $S+N \subseteq A^{-1}(A(S))$. With (4.1) this shows that

$$A^{-1}(A(S)) = S+N. \tag{4.3}$$

Since A is surjective and X and Y Fréchet spaces, the open mapping theorem [16, p. 77] shows that the image $A(G)$ of each (strongly) open set $G \subseteq X$ is (strongly) open in Y. If $S+N$ is closed, then M is open, so $A(S)$ is closed in view of (4.2). Conversely, if $A(S)$ is closed in Y, then $S+N$ is closed in view of (4.3) as the inverse image of a closed set by a continuous map.

Suppose now that the range $A(X)$ is closed in Y, but $A(X) \neq Y$. Then $A(X)$ is a Fréchet space, and the above argument shows that $S+N$ is closed in X iff $A(S)$ is closed in the relative topology of $A(X)$; the latter is true iff $A(S)$ is closed in Y. □

5. Duality and converse duality in linear programming. Let $\langle X, X^+ \rangle$ and $\langle Y, Y^+ \rangle$ be dual pairs; let $A: X \to Y$ be a w-continuous linear map; let $K \subseteq Y$ be a convex cone; let $c \in X^+$, and $b \in Y$. Consider the pair of linear programming problems:

$$\text{(P):} \quad \underset{x \in X}{\text{minimize}} \{\langle x, c \rangle : Ax - b \in K\};$$

$$\text{(D):} \quad \underset{v \in Y^+}{\text{maximize}} \{\langle b, v \rangle : A^+ v = c, v \in K^+\}.$$

If (P) attains a minimum at $x = a$, denote $d = Aa - b$, and define the map $\hat{d}: \mathbb{R} \to Y$ by $(\forall r \in \mathbb{R}) r \mapsto rd$.

The mathematical program

$$\text{(B):} \quad \text{maximize } \{\phi(z) : z \in U\}$$

is called a *dual program* to

$$\text{(A):} \quad \text{minimize } \{f(x) : x \in S\}$$

if

(i) $(\forall x \in S, \forall z \in U)\ \phi(z) \leqq f(x)$;

and

(ii) if (A) attains a maximum at $x^* \in S$, then (B) attains a minimum at some $z^* \in U$, and $f(x^*) = \phi(z^*)$.

THEOREM 8. *Let $A: X \to Y$ be a w-continuous linear map, and let $K \subseteq Y$ be a convex cone. If* (P) *attains a minimum at $x = a$, $d = Aa - b$, and the convex cone*

$$\left\{ \begin{bmatrix} A^+ s \\ \langle d, s \rangle \end{bmatrix} \in X^+ \times \mathbb{R} : s \in K^+ \right\} \tag{5.1}$$

is w-closed, then (D) *is a dual program to* (P).

Proof. Let $Ax - b \in K$, $A^+v = c$, $v \in K^+$. Then, for some $k \in K$,

$$\langle b, v\rangle = \langle Ax - k, v\rangle = \langle x, A^+v\rangle - \langle k, v\rangle \leqq \langle x, c\rangle$$

since $A^+v = c$, and $k \in K$, $v \in K^+ \Rightarrow \langle k, v\rangle \geqq 0$. So requirement (i) holds for (D) to be a dual to (P). It remains only to find a solution v to (D) satisfying $\langle a, c\rangle = \langle b, v\rangle$.

Let $z \in X$ and $r \in \mathbb{R}$ satisfy $Az + rd \in K$; then $Az = k - rd$ for some $k \in K$. If this implies that $z = 0$, then $\langle z, c\rangle = 0$; if $z \neq 0$ then, for all sufficiently small $t \in \mathbb{R}_+$,

$$A(a + tz) - b = d + t(k - rd) = tk + (1 - tr)\, d \in K.$$

Since (P) attains a minimum at $x = a$, it follows that $\langle a + tz, c\rangle \geqq \langle a, c\rangle$. Therefore

$$[A \quad \hat{d}]\begin{bmatrix} z \\ r \end{bmatrix} \in K \Rightarrow \left\langle \begin{bmatrix} z \\ r \end{bmatrix}, \begin{bmatrix} c \\ 0 \end{bmatrix} \right\rangle \geqq 0. \tag{5.2}$$

If $z = 0$, then (5.2) holds trivially.

By hypothesis, the cone

$$[A \quad \hat{d}]^+(K^+) = \left\{ \begin{bmatrix} A^+s \\ \langle d, s\rangle \end{bmatrix} \in X^+ \times \mathbb{R} : s \in K^+ \right\}$$

is w-closed. Then Theorem 2 shows that there exists an $h \in K^+$ such that

$$[A \quad \hat{d}]^+h = \begin{bmatrix} c \\ 0 \end{bmatrix},$$

and thus $A^+h = c$ and $\langle d, h\rangle = 0$. Then $v = h$ is a solution to (D), satisfying

$$\langle b, h\rangle = \langle Aa, h\rangle = \langle a, A^+h\rangle = \langle a, c\rangle$$

as required. □

Remarks. The hypothesis that the cone (5.1) is w-closed cannot be omitted, since an infinite-dimensional linear program can have a *duality gap*, i.e. min (P) > max (D) can occur. See Duffin [9] for relevant earlier results.

Kretschmer [13] has given an analogous result, with the hypothesis that (5.1) is w-closed replaced by the hypothesis that the cone

$$\left\{ \begin{bmatrix} A^+s \\ r - \langle b, s\rangle \end{bmatrix} : s \in K^+, r \in \mathbb{R}_+ \right\}$$

is w-closed. These hypotheses are nontrivial, even in finite dimensions, for convex cones which are not polyhedral.

Now let U and V be real locally convex spaces; let $L: U \to V$ be a strongly continuous linear map; let $S \subseteq U$ and $T \subseteq V$ be convex cones; let $m \in U'$ and $p \in V$. Consider the two linear programming problems:

$$\text{(P1):} \quad \underset{u \in U}{\text{minimize}} \{\langle u, m\rangle : Lu - p \in S\};$$

$$\text{(P2):} \quad \underset{v' \in V'}{\text{minimize}} \{\langle p, v'\rangle : L'v' - m \in T^*\}.$$

In finite dimensions, each problem can be put in the form of the other; but this does not necessarily hold in infinite dimensions. Define the cones

$$Q_1 = \left\{ \begin{bmatrix} L's \\ \langle La - p, s \rangle \end{bmatrix} \in U' \times \mathbb{R} : s \in S^+ \right\},$$

$$Q_2 = \left\{ \begin{bmatrix} Lu + s \\ \left\langle \begin{bmatrix} u \\ s \end{bmatrix}, \begin{bmatrix} 0 \\ a' \end{bmatrix} \right\rangle \end{bmatrix} : u \in U, s \in S \right\} = \begin{bmatrix} L(U) \\ \{0\} \end{bmatrix} + \left\{ \begin{bmatrix} s \\ \langle s, a' \rangle \end{bmatrix} : s \in S \right\};$$

where $a \in U$ and $a' \in V'$.

THEOREM 9. *If* (P1) *attains a minimum at* $u = a$, *and if the cone* Q_1 *is weakly closed in* $U' \times \mathbb{R}$, *then* (P1) *has a dual program*:

$$\text{(D1)}: \quad \underset{v' \in V'}{\text{maximize}} \{\langle p, v' \rangle : L'v' = m, v' \in S^*\},$$

which is equivalent to a minimization problem

$$\text{(D1')}: \underset{v' \in V'}{\text{minimize}} \left\{ \langle -p, v' \rangle : [L \quad I]'v' - \begin{bmatrix} m \\ 0 \end{bmatrix} \in \begin{bmatrix} U \\ S \end{bmatrix}^* \right\}$$

of the form of (P2). (I denotes the identity map.)

If (D1) *attains a maximum at* $v' = a'$, *and if the cone* Q_2 *is weak* closed in* $V \times \mathbb{R}$, *then* (D1) *has a dual program* (P1).

Proof. Apply Theorem 8.

Remarks. The second part of Theorem 9 is a *converse duality* result.

Analogous results hold for the dual of (P2).

6. Kuhn–Tucker conditions in dual pairs. Let $\langle X, X^+ \rangle$ and $\langle Y, Y^+ \rangle$ be dual pairs; let $f : X \to \mathbb{R}$ and $k : X \to Y$ be Fréchet-differentiable maps; let $K \subseteq Y$ be a convex cone such that $(-K) \cap (\bar{K}) = \{0\}$, where $\bar{K}$ denotes the w-closure of K. Let $Q = \{x \in X : k(x) \in K\}$. Fix $a \in X$, such that $d = k(a) \in K$.

Suppose $d \neq 0$; since $(-K) \cap (\bar{K}) = \{0\}$, $-d \notin \bar{K}$. Then by Lemma 2, there exists a $\psi \in \bar{K}^+ = K^+$ such that $\langle -d, \psi \rangle = -1$, so $\langle d, \psi \rangle = 1$. If $d = 0$, set $\psi = 0$. Define a projection P: $Y \to Y$ by $(\forall y \in Y)$ $Py = y - \langle y, \psi \rangle d$; then $Pd = 0$. Let $H = \{\xi \in X : Pk'(a)\xi \in P(K)\}$. Define the map $\hat{d} : \mathbb{R} \to Y$ by $(\forall r \in \mathbb{R})$ $\hat{d}r = rd$.

DEFINITION. The minimization problem

$$\text{(MP)}: \underset{x \in X}{\text{minimize}} \{f(x) : k(x) \in K\}$$

satisfies the *extended Kuhn–Tucker constraint qualification* (EKTCQ) at the point $a \in X$ (such that $d = k(a) \in K$) if to each $w \in H$ there exists a continuous arc $x = \beta(t)$ $(0 \leqq t \leqq \delta)$ in Q, for which $\beta(0) = a$ and the initial slope $\beta'(0) = \xi$.

Remark. An elementary calculation shows that the EKTCQ reduces to the classical Kuhn–Tucker constraint qualification in the case in which $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$, K is the positive orthant in $\mathbb{R}_m$, and P is the orthogonal projection orthogonal to d. The present definition is applicable to any dimension, and to arbitrary cones.

THEOREM 10 (Extended Kuhn–Tucker conditions). *Let* (MP) *attain a local minimum at* $x = a$; *let* $d = k(a)$; *let the* EKTCQ *hold at* a; *let the cone*

$$[k'(a) \quad \hat{d}]^+(K^+) \tag{6.1}$$

be w-closed in $X^+ \times \mathbb{R}$ (*where* $k'(a)$ *is the Fréchet derivative of* k *at* a). *Then there exists a Lagrange multiplier* $\lambda \in K^+$ *such that*

$$f'(a) = \lambda \circ k'(a) \quad \textit{and} \quad \langle k(a), \lambda \rangle = 0.$$

Proof. If $w \in H$, thus if

$$(\exists z \in K, \exists r \in \mathbb{R}) \quad k'(a)w - z = -rd,$$

then

$$f'(a)\xi = \frac{d}{dt} f(\beta(t))|_{t=0} \geqq 0$$

since (MP) attains a minimum at a, and the arc $\{\beta(t)\} \subseteq Q$. Thus

$$[k'(a) \quad \hat{d}]\begin{bmatrix} w \\ r \end{bmatrix} \in K^+ \Rightarrow [f'(a) \quad 0]\begin{bmatrix} w \\ r \end{bmatrix} \geqq 0.$$

Since the cone (6.1) is w-closed, Theorem 2 shows there exists $\lambda \in K^+$ such that

$$[f'(a) \quad 0] = \lambda \circ [k'(a) \quad \hat{d}],$$

and the theorem follows. □

7. Finite-dimensional applications. In this section we assume that X, Y, Z are finite-dimensional real vector spaces, and that $S \subseteq X$, $T \subseteq Z$ are convex cones. We identify X with its dual by choosing any inner product which induces the unique locally convex topology in X; the same convention applies to Y and Z. We are concerned with the following two problems:

PROBLEM I. Given linear maps $A: X \to Y$, $B: Z \to Y$, find a linear map $U: Z \to X$ such that

$$B = AU \quad \text{and} \quad U(T) \subseteq S$$

(cf. Fig. 3).

PROBLEM II. Given linear maps $C: Y \to X$, $D: Y \to Z$, find a linear map $V: X \to Z$ such that

$$D = VC \quad \text{and} \quad V(S) \subseteq T$$

(cf. Fig. 4).

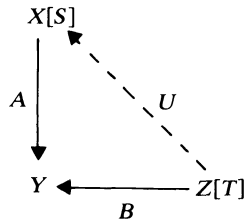


FIG. 3

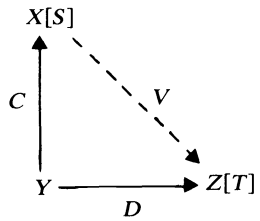


FIG. 4

If the cones S and T are closed, so that $S^{**} = S$ and $T^{**} = T$, we can transform Problem II to Problem I according to the following diagram:

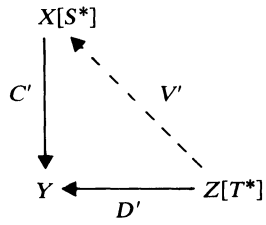


FIG. 5

The following lemma, the proof of which we omit, provides a link between Problem I and Theorem 3, and between Problem II and Theorem 4.

LEMMA 5. *The conditions*

$$B(Z) \subseteq A(X) \quad \textit{and} \quad B(T) \subseteq A(S) \tag{7.1}$$

$$(\textit{resp. } D'(Z) \subseteq C'(X) \quad \textit{and} \quad D'(T^*) \subseteq C'(S^*)) \tag{7.2}$$

are necessary for the solvability of Problem I (*resp. Problem* II). *They are also sufficient if A is injective* (*resp. C is surjective*).

We note that (7.2) is equivalent to

$$C^{-1}(0) \subseteq D^{-1}(0) \quad \text{and} \quad C^{-1}(S) \subseteq D^{-1}(T).$$

If Z is specialized to $\mathbb{R}$ and the cone T to $\mathbb{R}_+$, Problem I is answered by Theorem 5, and Problem II by Theorem 6. Mangasarian [14, Thm. 3.1] gave a solution to Problem I in the case that $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$, $Z = \mathbb{R}^k$, and that the cones S, T are the positive orthants $\mathbb{R}^n_+$, $\mathbb{R}^k_+$. He proved that under these assumptions Problem I has a solution iff

$$A'y \in \mathbb{R}^n_+ \Rightarrow B'y \in \mathbb{R}^k_+.$$

Our first theorem in the present section is devoted to a generalization of Mangasarian's result. We show that S can be any convex cone, while T is restricted to a special polyhedral cone (sometimes called simplicial) which in $\mathbb{R}^k$ reduces to an isomorphic image of the positive orthant. Necessary information about polyhedral cones can be found in [17].

THEOREM 11. *Let $A(S)$ be closed and let T be a polyhedral cone whose generators form a basis of Z. Then the following are equivalent*:

(a) *Problem* I *has a solution.*

(b) $B(T) \subseteq A(S)$.

(c) $A'y \in S^* \Rightarrow B'y \in T^*$.

Proof. The conditions (b) and (c) are equivalent by Theorem 3, and (a) implies (b) by Lemma 5. Only (b)$\Rightarrow$(a) remains to be proved. Let $t_1, t_2, \cdots, t_k$ be generators of T that form a basis of Z, so that

$$T = \left\{ \sum_{i=1}^{k} \lambda_i t_i : \lambda_i \geqq 0 \right\}. \tag{7.3}$$

By (b), there are vectors $s_i \in S$ such that $Bt_i = As_i$, $i = 1, \cdots, k$. Let U be the linear map from Z to X that carries t_i to s_i, i $= 1, \cdots, k$. Then

$$AUz = AU\Big(\sum_{i=1}^{k} \xi_i t_i\Big) = \sum_{i=1}^{k} \xi_i AUt_i = \sum_{i=1}^{k} \xi_i As_i = \sum_{i=1}^{k} \xi_i Bt_i = B\Big(\sum_{i=1}^{k} \xi_i t_i\Big) = Bz$$

for each $z \in Z$, so that $AU = B$. The relation (7.3) shows that $U(T) \subseteq S$. □

Transforming Problem II to Problem I via the diagram in Fig. 5 and recalling that a polyhedral cone is closed, we obtain the following result.

THEOREM 12. *Let S and $C'(S^*)$ be closed and let T be a polyhedral cone whose generators form a basis of Z. Then the following are equivalent*:

(a) *Problem* II *has a solution.*

(b) $D'(T^*) \subseteq C'(S^*)$.

(c) $Cy \in S \Rightarrow Dy \in T$.

When we attempt to replace T in Theorems 11 and 12 by an arbitrary polyhedral cone, we may discover that the condition (b) is no longer sufficient for the solvability of the appropriate problem. We are forced to restrict the cone S also to a polyhedral cone, and to resort to necessary and sufficient conditions of a different type. Let us assume that $S \subseteq X$ and $T \subseteq Z$ are polyhedral cones, and let us introduce the following notation:

$\mathscr{X} = L(Z, X) = \mathscr{X}'$, the set of all linear maps from Z to X;

$\mathscr{Y} = L(Z, Y) = \mathscr{Y}'$, the set of all linear maps from Z to Y;

$(U, V) = \text{tr}\,(V'U)$, the inner product in $\mathscr{X}$ (resp. $\mathscr{Y}$);

$\mathscr{A}: \mathscr{X} \to \mathscr{Y}$, the map $\mathscr{A}(U) = AU$ for all $U \in \mathscr{X}$;

$\mathscr{S} = \{U \in \mathscr{X} : U(T) \subseteq S\}$;

$\mathscr{S}^* = \{V \in \mathscr{X} : (U, V) \geqq 0 \text{ for all } U \in \mathscr{S}^*\}$.

Recall that the trace tr M of a linear map $M: Z \to Z$ is defined by tr $M = \sum_{k=1}^{n} (Mz_k, z_k)$, where $(z_1, \cdots, z_n)$ is any orthonormal basis of Z.

We observe that $\mathscr{A}$ is a linear map from $\mathscr{X}$ to $\mathscr{Y}$ with adjoint $\mathscr{A}': \mathscr{Y} \to \mathscr{X}$ defined by $\mathscr{A}'(W) = A'W$ for all $W \in \mathscr{Y}$. It is easily seen that $\mathscr{S}$ is a closed convex cone. We prove that $\mathscr{S}$ is in fact a polyhedral cone in $\mathscr{H}$. Then $\mathscr{A}(\mathscr{S})$ is closed being a polyhedral cone in $\mathscr{Y}$, and Theorem 5 is applicable to $\mathscr{H}$, $\mathscr{Y}$, $\mathscr{A}$ and $\mathscr{S}$.

In the next lemma we give a description of the dual cone $\mathscr{Y}^*$ motivated by Theorem 3.1 of Berman and Gaiha [4] for cones of matrices. We assume that

$$t_1, \cdots, t_p \quad \text{and} \quad s_1^*, \cdots, s_q^*$$

are generators of T and S^*, respectively.

LEMMA 6. *The cones $\mathscr{S}$ and $\mathscr{S}^*$ are polyhedral, with $\mathscr{S}^*$ generated by $P_{ij} \in \mathscr{X}$, $1 \leqq i \leqq p$, $1 \leqq j \leqq q$, where*

$$P_{ij}z = (z, t_i)s_j^* \quad \textit{for all } z \in Z. \tag{7.4}$$

Proof. First we show that for each $U \in \mathscr{X}$

$$\text{tr}\,(P_{ij}'U) = (Ut_i, s_j^*). \tag{7.5}$$

Pick an orthonormal basis $(z_1, \cdots, z_n)$ of Z. Then

$$(Ut_i, s_j^*) = \Big(\sum_{k=1}^{n} (t_i, z_k) Uz_k, s_j^* \Big) = \sum_{k=1}^{n} (Uz_k, (z_k, t_i) s_j^*)$$

$$= \sum_{k=1}^{n} (Uz_k, P_{ij} z_k) = \sum_{k=1}^{n} (P_{ij}' Uz_k, z_k).$$

Let $\mathscr{P}$ be the polyhedral cone in $\mathscr{X}$ with generators P_{ij}. Since $U \in \mathscr{S}$ iff $(Ut_i, s_j^*) \geqq 0$ for all i and j, we deduce from (7.5) that

$$\mathscr{S} = \mathscr{P}^*, \quad \text{and so} \quad \mathscr{S}^* = \mathscr{P}.$$

Since $\mathscr{P}$ is a polyhedral cone, so are $\mathscr{S}$ and $\mathscr{S}^*$. □

Write $\mathscr{T} = \{V \in \mathscr{X} : V(T^*) \subseteq S^*\}$. It is proved in [4] that $\mathscr{S}^* \subseteq \mathscr{T}$ even if S and T are merely closed convex cones. This inclusion follows directly from (7.4) as

$$(P_{ij} t^*, s) = (t^*, t_i)(s_j^*, s) \geqq 0$$

for all $t^* \in T^*$ and all $s \in S$.

Theorem 5 and Lemma 6 combine to give the following result:

THEOREM 13. *Let $S \subseteq X$ and $T \subseteq Z$ be polyhedral cones. Then Problem* I *has a solution iff*

$$A'V \in \mathscr{S}^* \Rightarrow \operatorname{tr}(B'V) \geqq 0 \quad (V \in L(Z, Y)). \tag{7.6}$$

Since $\mathscr{S}^* \subseteq \mathscr{T}$, Theorem 13 implies the following corollary:

COROLLARY. *Let $S \subseteq X$ and $T \subseteq Y$ be polyhedral cones. If*

$$(A'V)(T^*) \subseteq S^* \Rightarrow \operatorname{tr}(B'V) \geqq 0 \qquad (V \in L(Z, Y)), \tag{7.7}$$

then Problem I *has a solution.*

Applying the transformation of Fig. 5 and recalling that S^*, T^* are polyhedral when S, T are polyhedral, we obtain the following analogue of Theorem 13, in which

$$\mathscr{T} = \{V \in \mathscr{X} : V(T^*) \subseteq S^*\} = \{V \in \mathscr{X} : V'(S) \subseteq T\}.$$

THEOREM 14. *Let $S \subseteq X$ and $T \subseteq Z$ be polyhedral cones. Then Problem* II *has a solution iff*

$$CW \in \mathscr{T}^* \Rightarrow \operatorname{tr}(DW) \geqq 0 \qquad (W \in L(Z, Y)). \tag{7.8}$$

More generally, (7.6) (resp. (7.8)) is a necessary and sufficient condition for the solvability of Problem I (resp. Problem II) if S and T are closed convex cones and if the cone $\mathscr{A}(\mathscr{S})$ (resp. $C'(\mathscr{T})$) is closed. In this case we have the following characterization of $\mathscr{S}^*$: For $t \in T$ and $s^* \in S^*$ define $P_{ts^*} \in \mathscr{X}$ by $P_{ts^*} z = (z, t) s^*$ for all $z \in Z$, and put

$$\mathscr{P} = \{P_{ts^*} \in \mathscr{X} : t \in T, s^* \in S^*\}.$$

Then

$$\mathscr{S} = \mathscr{P}^* \quad \text{and} \quad \mathscr{S}^* = \overline{\text{co}}\, \mathscr{P},$$

where $\overline{\text{co}}$ denotes the closed convex hull in $\mathscr{X}$. An analogous representation can be given for the cone $\mathscr{T}^*$. Also in this more general case, $\mathscr{S}^* \subseteq \mathscr{T}$. This representation is given in Theorem 3.1 of [4] for matrices.

For other finite-dimensional applications of Farkas' theorem (Theorem 5) including a theorem of Bellman and Fan on positive definite matrices, and theorems of the Lyapunov and Stein type, as well as further generalizations, see Ben-Israel and Berman [3].

REFERENCES

[1] A. BEN-ISRAEL, *Linear equations and inequalities on finite dimensional, real or complex, vector spaces: A unified theory*, J. Math. Anal. Appl., 27 (1969), pp. 367–389.

[2] A. BERMAN AND A. BEN-ISRAEL, *Linear inequalities, mathematical programming and matrix theory*, Math. Programming, 1 (1971), pp. 291–300.

[3] ———, *More on linear inequalities with applications to matrix theory*, J. Math. Anal. Appl., 33 (1971), pp. 482–496.

[4] A. BERMAN AND P. GAIHA, *A generalization of irreducible monotonicity*, Linear Algebra Appl., 5 (1972), pp. 29–38.

[5] A. BERMAN, *Cones, Matrices and Mathematical Programming*, Lecture Notes in Economics and Math. Systems 79, Springer-Verlag, Berlin, 1973.

[6] D. G. BOURGIN, *Linear topological spaces*, Amer. J. Math., 65 (1943), pp. 637–659.

[7] B. D. CRAVEN, *Nonlinear programming in locally convex spaces*, J. Optimization Theory Appl., 10 (1972), pp. 197–210.

[8] B. D. CRAVEN, *A generalized Motzkin alternative theorem*, Melbourne University School of Mathematical Sciences Res. Rep. 13, 1975.

[9] R. J. DUFFIN, *Infinite programs*, Linear Inequalities and Related Systems, H. W. Kuhn and A. W. Tucker, eds., Annals of Math. Studies 38, Princeton University Press, Princeton, NJ, 1956, pp. 157–170.

[10] KY FAN, *On systems of linear inequalities*, Linear Inequalities and Related Systems, H. W. Kuhn and A. W. Tucker, eds., Annals of Math. Studies 38, Princeton University Press, Princeton, NJ, 1956, pp. 99–156.

[11] J. FARKAS, *Über die Theorie der einfachen Ungleichungen*, J. Reine Angew. Math., 124 (1902), pp. 1–24.

[12] L. HURWICZ, *Programming in linear spaces*, Studies in Linear and Nonlinear Programming, K. J. Arrow, L. Hurwicz and H. Uzawa, eds., Stanford University Press, Stanford, CA., 1958, pp. 38–102.

[13] K. KRETSCHMER, *Programming in paired spaces*, Canad. J. Math., 13 (1961), pp. 221–238.

[14] O. L. MANGASARIAN, *Perron–Frobenius properties of $Ax - \lambda Bx$*, J. Math. Anal. Appl., 36 (1971), pp. 86–102.

[15] A. P. ROBERTSON AND W. ROBERTSON, *Topological Vector Spaces*, Cambridge University Press, Cambridge, 1964.

[16] H. H. SCHAEFFER, *Topological Vector Spaces*, Springer-Verlag, New York, 1971.

[17] H. UZAWA, *A theorem on convex polyhedral cones*, Studies in Linear and Nonlinear Programming, K. J. Arrow, L. Hurwicz and H. Uzawa, eds., Stanford University Press, Stanford, CA, 1958, pp. 23–31.

SIAM J. MATH. ANAL.
Vol. 8, No. 6, November 1977

NONEXISTENCE OF NONTRIVIAL SOLUTIONS OF SCHRÖDINGER TYPE SYSTEMS*

G. B. KHOSROVSHAHI†

Abstract. This paper is primarily concerned with the question of nonexistence of nontrivial solutions of the Schrödinger equation $\Delta u + p(x)u = 0$ in an exterior domain Ω, which is a connected open region in $\mathbb{E}^n$, containing $\{x \mid \|x\| \geqq R_0\}$. The function $p(x)$, $(p(x) = p_0(x) + p_1(x))$, is assumed to be a continuous function, p_0 satisfies $r\,\partial p_0/\partial r + \delta_1 p_0 \geqq \delta_2$ where $\delta_1, \delta_2 > 0$, and p_1 is a complex-valued function such that $\sup_{\|x\|=r} |rp(x)| \leqq K$. It is shown that any C^2 solution which satisfies $\int_{R_0}^{\infty} \rho^{\beta} \oint_{S_\rho} p_0 |u|^2\, ds\, d\rho \leqq M$ for $\beta > -1$ is identically zero. The theorem is generalized to a system of equations of the type $\Delta u_i + \sum_{j=1}^{m} p_{ij} u = 0$, $i = 1, \cdots, m$, and a theorem concerning upper bounds for positive eigenvalues of the Schrödinger operator $-\Delta + V(x)$ is given. The main theorem has features in common with Agmon's result [J. Analyse Math., 23 (1970), pp. 1–25]. However somewhat different conditions on $p(x)$ are assumed and a different line of proof is given. This method will be extended to more general elliptic differential equations and nonlinear inequalities in subsequent papers.

Introduction. In this paper we are concerned with questions of growth and of nonexistence of solutions of Schrödinger type equations in exterior domains. In particular we consider solutions of the equation

$$\Delta u + p(x)u = 0 \quad \text{in } \Omega,$$

where Ω is an open unbounded connected region in n-dimensional Euclidean space ($n \geqq 3$) and Δ denotes the Laplace operator in n-dimension. Throughout we shall assume $u \in C^2(\Omega)$ and that the potential p is a prescribed complex-valued function in Ω.

This problem has been investigated by F. Rellich [11], T. Kato [6], and S. Agmon [1], [2] and [3]. Successive authors have in general considered an over widening class of potentials thus in some way extending the result of previous authors.

The primary objective in this paper is to take potentials similar to those of Agmon's [1] and by imposing somewhat different conditions on $p(x)$ and on the class of admissible solutions give a variation of his theorem using a somewhat different method of proof. This method can be applied to classes of solutions not treated by Agmon, and it can be extended to more general elliptic operators and to classes of nonlinear inequalities. These latter questions will be dealt with in subsequent papers. We indicate in § 2 a reasonably straightforward extension to elliptic systems of the form $\Delta u_i + \sum_{j=1}^{m} p_{ij}(x) u_j = 0$, $i = 1, \cdots, m$.

In § 2 we also specialize our results to an eigenvalue problem and state a theorem concerning the positive eigenvalues of the operator $H = -\Delta + V(x)$. This theorem handles a wider class of potentials than those cited in the literature; in particular, it includes Agmon's Theorem 4[1] and it also provides upper bounds which are at least as sharp as Kato's, for the positive eigenvalues of the operator H with $V(x) = O(1/r)$.

* Received by the editors August 26, 1974, and in revised form April 20, 1976.

† Mathematics Department, University of Tehran, Tehran, Iran, and Department of Mathematics, Cornell University, Ithaca, New York, 14850. This research was supported by the National Science Foundation under Grant GP 33031X.

1. Single Schrödinger equation with complex potentials. In this section we shall deal with the Schrödinger equation for a specific class of potentials, studying in particular the behavior of solutions as $\|x\| \to \infty$. The main purpose of the section is to demonstrate a method of proving nonexistence of solution in a certain class of functions. Although the method has some features in common with those of Agmon [1] and Kato [6], it will nevertheless yield sharper results in some cases.

Preliminaries. (i) The following lemma will be used in subsequent sections.

MAIN LEMMA. *Let F be a nonnegative, real valued function of t, continuous in the half open interval $(0, t_0]$ and twice continuously differentiable in $(0, t_0)$, where t_0 is an arbitrary and finite number.*

Suppose for all $t \in (0, t_0]$,

$$\int_0^t \eta^{-\alpha_2} F(\eta)\, d\eta < \infty, \tag{1.1}$$

and that

$$FF'' - (F')^2 \geqq -C_1 t^{-1} FF' + \varepsilon t^{-\alpha_1} F \int_0^t \eta^{-\alpha_2} F(\eta)\, d\eta, \tag{1.2}$$

where C_1 and ε are strictly positive, $\alpha_2 > 1$ and $-2C_1 + \alpha_1 + \alpha_2 > 1$. Under these conditions $F(t) \equiv 0$, $\forall t \in [0, t_0]$.

The proof of this lemma is given in the Appendix.

Remark 1. One may generalize the above lemma by replacing (1.1) and (1.2) by

$$\int_0^t h_2(\eta) F(\eta)\, d\eta < \infty, \tag{1.1}'$$

and

$$FF'' - (F')^2 \geqq -C_1 t^{-1} FF' - C_2(t) F^2 + \varepsilon h_1(t) F \int_0^t h_2(\eta) F(\eta)\, d\eta \tag{1.2}'$$

provided C_2, h_1 and h_2 are appropriately chosen.

Remark 2. For a certain range of values of $\alpha_1 > 1$ this lemma can be obtained from Agmon's Lemma 1 [1].

(ii) Throughout this paper we let Ω be an open connected region in $\mathbb{R}^n$ for $n \geqq 3$ and use the notation throughout $D_{R_0} = \{x \mid \|x\| \geqq R_0\}$. It will be assumed that $D_{R_0} \subset \Omega$.

We shall be concerned with solutions $u \in C^2(\Omega)$ which satisfy

$$\Delta u + p(x) u = 0 \quad \text{in } \Omega, \tag{1.3}$$

where $p(x)$ is a complex-valued function in Ω which can be decomposed as

$$p(x) = p_0(x) + p_1(x) \quad \text{in } \|x\| \geqq R_0. \tag{1.4}$$

In (1.4) we assume that $p_0(x)$ is real and positive and possesses a continuous radial derivative while $p_1(x)$ is a complex valued function satisfying for sufficiently large R_0,

$$\sup_{\|x\| = r} |r p_1(x)| \leqq K \quad \text{for } r \geqq R_0. \tag{1.5}$$

We shall subsequently be imposing a further assumption on $p_0(x)$ of the following type:

$$r\frac{\partial p_0}{\partial r}+\delta_1 p_0 \geqq \delta_2 \quad \text{for } \delta_1, \delta_2 > 0, \tag{1.6}$$

assumed to hold along every ray for $r \geqq R_0$ (sufficiently large).

From (1.6) it follows that at every point outside a sufficiently large sphere one can compute a constant K_2 such that

$$p_0(r, \xi) \geqq K_2, \tag{1.7}$$

(we use the symbol $p_0(r, \xi)$ to indicate that p_0 may depend on the angle variable).

Before establishing our main theorem, let us investigate the asymptotic behavior of solutions of (1.3).

LEMMA 1.1. *Let u and p_0 be as in* (1.3) *and* (1.6). *Then*

$$\int_{R_0-1}^{\infty} \rho^{\beta} \oint_{S_\rho} p_0|u|^2 \, ds \, d\rho \leqq M_1 \Rightarrow \int_{R_0}^{\infty} \rho^{\beta} \oint_{S_\rho} |\text{grad } u|^2 \, ds \, d\rho \leqq M_2 \tag{1.8}$$

for some positive M_1 and M_2.

COROLLARY 1.1. *Suppose u and p_0 are as in Lemma* 1.1 *and*

$$\int_{R_0}^{\infty} \rho^{\beta} \oint_{S_\rho} p_0|u|^2 \, ds \, d\rho \leqq M. \tag{1.9}$$

Then

$$\text{(i)} \quad \liminf_{r\to\infty} r^{\beta+1} \oint_{S_r} p_0|u|^2 \, ds = 0,$$

$$\text{(ii)} \quad \liminf_{r\to\infty} r^{\beta+1} \oint_{S_r} p|\text{grad } u|^2 \, ds = 0. \tag{1.10}$$

$$\text{(iii)} \quad \liminf_{r\to\infty} r^{\beta+1} \oint_{S_r} u_r \bar{u} \, ds \leqq 0.$$

We do not give the proof of Lemma 1.1 and Corollary 1.1 here. They may be found in [8] and are obtained in a reasonably standard manner. One introduces a function $\delta(\rho)$ defined by

$$\delta = \begin{cases} (R_0-1-\rho)^2, & R_0-1<\rho<R_0, \\ 1, & R_0 \leqq \rho \leqq r, \\ (r-\rho+1)^2, & r<\rho<r+1, \\ 0, & \text{otherwise}, \end{cases} \tag{1.11}$$

notes that

$$\int_{R_0}^{r} \rho^{\beta} \oint_{S_\rho} |\text{grad } u|^2 \, ds \, d\rho \leqq \int_{R_0-1}^{r+1} \rho_{\beta} \oint_{S_\rho} \delta(\rho)|\text{grad } u|^2 \, ds \, d\rho, \tag{1.12}$$

and by appropriate use on the right of integration by parts, the arithmetic-geometric mean inequality and the differential equation arrives at (1.8). The proof of Corollary 1.1 is straightforward.

We are now ready to prove the main theorem of this section.

THEOREM 1.1. *Let $u \in C^2(\Omega)$ be a solution of* (1.3) *for $p(x)$ satisfying* (1.4) *and* (1.5). *If for some constant β,*

$$\int_{R_0}^{\infty} \rho^{\beta} \oint_{S_\rho} p_0 |u|^2 \, ds \, d\rho \leqq M$$

and if for R_0 sufficiently large p_0 satisfies either

$$r\frac{\partial p_0}{\partial r} + 2(\beta + 1 - \gamma)p_0 \geqq \frac{2K^2}{\gamma} + \varepsilon, \qquad r \geqq R_0, \tag{1.13(i)}$$

when $-1 < \beta \leqq 0$ (here γ is an arbitrary positive constant which satisfies $|\beta - \gamma| < 1$) or

$$r\frac{\partial p_0}{\partial r} + (2 - \varepsilon)p_0 \geqq \frac{K^2}{\beta} + \varepsilon_1, \qquad r \geqq R_0, \tag{1.13(ii)}$$

where $\beta > 0$, then u must vanish identically in Ω. (Hereafter, ε will stand for an arbitrary small positive constant.)

The proof of this theorem will utilize three inequalities which follow from generalized Green's identities and asymptotic properties just derived. These are stated as lemmas.

LEMMA 1.2. *Let u and p be as in Theorem* 1.1 *and let*

$$\int_{r}^{\infty} \rho^{\beta} \oint_{S_\rho} p_0 |u|^2 \, ds \, d\rho \leqq M \quad \text{for some } \beta > -1;$$

then for $r \geqq R_0$ (sufficiently large),

$$\begin{aligned} I_1 \equiv & -2 \oint_{S_r} |u_\rho|^2 \, ds + \oint_{S_r} |\operatorname{grad} u|^2 \, ds - \oint_{S_r} p_0 |u|^2 \, ds \\ \geqq & (2\beta - \gamma) r^{-(\beta+1)} \int_{D_r} \rho^{\beta} |u_\rho|^2 \, dx - (n + \beta - 2) r^{-(\beta+1)} \int_{D_r} \rho^{\beta} |\operatorname{grad} u|^2 \, dx \\ & - \frac{K^2}{\gamma} r^{-(\beta+1)} \int_{D_r} \rho^{\beta} |u|^2 \, dx + r^{-(\beta+1)} \int_{D_r} \rho^{\beta} [\rho p_{0\rho} + (n + \beta) p_0] |u|^2 \, dx. \end{aligned} \tag{1.14}$$

To establish this inequality we consider the identity

$$\operatorname{Re} \int_{D_r - D_{r_1}} \rho^{\beta} x_i \bar{u}_{,i} (\Delta u + pu) \, dx = 0 \quad \text{for } r_1 > r. \tag{R}^1$$

We apply the divergence theorem to integrals involving Δu and $p_0 u$ and the arithmetic-geometric mean inequality to the term involving $p_1 u$. We also make use of (1.10).

LEMMA 1.3. *Let u and p be as in Theorem* 1.1. *If the condition of Corollary* 1.1 *is satisfied, then*

$$\begin{aligned} \int_{D_r} \rho^{\beta} |\operatorname{grad} u|^2 \, dx \leqq & -\frac{1}{1 - \varepsilon_1} r^{\beta} \operatorname{Re} \oint_{S_r} u_\rho \bar{u} \, ds + \varepsilon_2 \int_{D_r} \rho^{\beta} |u|^2 \, dx \\ & + \frac{1}{1 - \varepsilon_1} \int_{D_r} \rho^{\beta} p_0 |u|^2 \, dx \end{aligned} \tag{1.15}$$

[1] Throughout this paper we have adopted the summation convention over repeated indices and a comma will denote differentiation.

for any β, $r \geqq R_0$ (sufficiently large) and for ε_i positive and arbitrarily small.

The proof utilizes the following identity, integration by parts and divergence theorem:

$$(1.16)^2 \quad \oint_{D_r - D_{r_1}} \rho^\beta |\operatorname{grad} u|^2 \, dx = \operatorname{Re} \oint_{S_\rho} \rho^\beta \frac{x_i}{\rho} u_{,i} \bar{u} \, ds \Big|_{\rho = r_1}^{\rho = r}$$
$$-\beta \operatorname{Re} \int_{D_r - D_{r_1}} \rho^{\beta - 1} \rho_{,i} u_{,i} \bar{u} \, dx$$
$$-\operatorname{Re} \int_{D_r - D_{r_1}} \rho^\beta u_{,ii} \bar{u} \, dx.$$

Note 1. From (1.16), by taking lim inf as $r_1 \to \infty$, making use of arithmetic-geometric mean inequality and (1.3), the following inequality follows:

$$(1.17) \quad \int_{D_r} \rho^\beta |\operatorname{grad} u|^2 \, dx \leqq -r^\beta \operatorname{Re} \oint_{S_r} u_\rho \bar{u} \, ds + \varepsilon_1 \int_{D_r} \rho^\beta |u|^2 \, dx$$
$$+\varepsilon_2 \int_{D_r} \rho^\beta |u_\rho|^2 \, dx + \int_{D_r} \rho^\beta p_0 |u|^2 \, dx \qquad \forall r \geqq R_0.$$

In (1.17) the quantities ε_1 and ε_2 may be chosen to be proportional to R_0^{-1}. Thus by choosing R_0 sufficiently large ε_1 and ε_2 may be made arbitrarily small. In what follows we shall in any given equation put different indices on the ε's whenever they may be chosen independently. However, an ε with the same index in two different equations will not in general denote the same number.

Note 2. By choosing R_0 sufficiently large in (1.17), it follows that

$$(1.18) \quad \int_{D_r} \rho^\beta |\operatorname{grad} u|^2 \, dx \leqq \varepsilon_1 r^{\beta+1} \oint_{S_r} |u|^2 \, ds + \varepsilon_2 r^{\beta+1} \oint_{S_r} |u_\rho|^2 \, ds$$
$$+\varepsilon_3 \int_{D_r} \rho^\beta |u|^2 \, dx + \varepsilon_4 \int_{D_r} \rho^\beta |u_\rho|^2 \, dx + \int_{D_r} \rho^\beta p_0 |u|^2 \, dx$$

for arbitrarily small ε_i's.

LEMMA 1.4. *Let u and p be as in Theorem* 1.1, *and satisfy* (1.9). *Then for $\beta > 0$ it is possible to choose R_0 so large that*

$$(1.19) \quad \int_{D_r} \rho^\beta |u_\rho|^2 \, dx \leqq -\frac{r^\beta}{2 - \varepsilon_1} \operatorname{Re} \oint_{S_r} u_\rho \bar{u} \, ds + \varepsilon_2 \int_{D_r} \rho^\beta p_0 |u|^2 \, dx$$
$$+\varepsilon_1 \int_{D_r} \rho^\beta |u|^2 \, dx$$

for ε's arbitrarily small and positive and for all $r \geqq R_0$.

---

[2] $\rho_{,i} = \partial \rho / \partial x_i$.

To establish Lemma 1.4, we let α_1 and α_2 be two, as yet arbitrary, positive numbers, and set

$$J(r) = -(2-\alpha_1)\int_{D_r} \rho^{\beta}|u_\rho|^2\,dx + \int_{D_r} \rho^{\beta}|\operatorname{grad} u|^2\,dx - (1-\alpha_2)\int_{D_r} \rho^{\beta}p_0|u|^2\,dx \tag{1.20}$$

for any $r \geqq r_0$. Then by differentiating with respect to r and using Lemma 1.2 we obtain

$$rJ_r \leqq 0 \Rightarrow \int_r^{r_1} J_\rho\,d\rho \leqq 0 \Rightarrow J(r_1) - J(r) \leqq 0. \tag{1.21}$$

Since $\lim_{r_1\to\infty} J(r_1) = 0$, (1.21) implies $J(r) \geqq 0$. By Note 1, this leads to the desired result. (For detailed proof of Lemma 1.4 see Appendix B.)

Proof of Theorem 1.1. Let us assume that u is the solution of (1.3) and does not vanish identically in Ω. We define

$$F(r) = \int_\Sigma |u(r,\xi)|^2\,d\xi, \tag{1.22}$$

where Σ is the solid angle and $d\xi$ is the element of the surface on the unit sphere. We will show that F satisfies a second order differential inequality.

By successive differentiation we obtain

$$F_r(r) = 2\,\mathrm{Re}\int_\Sigma u_r\bar{u}\,d\xi, \tag{1.23}$$

and

$$F_{rr}(r) = 2\int_\Sigma |u_r|^2\,d\xi + 2\,\mathrm{Re}\int_\Sigma u_{rr}\bar{u}\,d\xi, \tag{1.24}$$

where $F_r = \partial F/\partial r$ and $F_{rr} = \partial^2 F/\partial r^2$.

We introduce the new variable $t = r^{-(n-2)}$ and compute F_t, F_{tt} as follows:

$$F_t = -\frac{2}{n-2}r^{n-1}\int_\Sigma u_r\bar{u}\,d\xi \tag{1.25}$$

and

$$F_{tt} = \frac{2}{(n-2)^2}r^{2(n-1)}\left[\int_\Sigma |\operatorname{grad} u|^2\,d\xi - \mathrm{Re}\int_\Sigma p|u|^2\,d\xi\right]. \tag{1.26}$$

We now form $FF_{tt} - (F_t)^2$, and use Schwarz's inequality to obtain

$$\begin{aligned} FF_{tt} - (F_t)^2 &= \frac{2}{(n-2)^2}r^{2(n-1)}\int_\Sigma |u|^2\,d\xi\left[\int_\Sigma |\operatorname{grad} u|^2\,d\xi - \mathrm{Re}\int_\Sigma p|u|^2\,d\xi\right] \\ &\quad - \frac{4}{(n-2)^2}r^{n-1}\left(\mathrm{Re}\int_\Sigma u_r\bar{u}\,d\xi\right)^2 \\ &\geqq \frac{2}{(n-2)^2}r^{2(n-1)}F\left[\int_\Sigma |\operatorname{grad} u|^2\,d\xi - \mathrm{Re}\int_\Sigma p|u|^2\,d\xi - \int_\Sigma |u_r|^2\,d\xi\right], \end{aligned} \tag{1.27}$$

or

(1.28)
$$FF_{tt}-(F_t)^2 \geqq \frac{2}{(n-2)^2} r^{n-1} F\left[\oint_{S_r} |\operatorname{grad} u|^2\, ds - \operatorname{Re}\oint_{S_r} p|u|^2\, ds - 2\oint_{S_r} |u_\rho|^2\, ds\right],$$

where $ds = r^{n-1}\, d\xi$.

Suppose $u \not\equiv 0$ in D_{R_0}. Then it follows from unique continuation theorems for the solution of (1.3), Hormander [4], that u does not vanish identically in any open set. This in turn implies that the function $F(r)$ does not vanish in any interval. To draw a contradiction to this assertion, we proceed as follows:

Let

(1.29)
$$I_2 = -2\oint_{S_r} |u_\rho|^2\, ds + \oint_{S_r} |\operatorname{grad} u|^2\, ds - \operatorname{Re}\oint_{S_r} p|u|^2\, ds;$$

then by Lemma 1.2 we have

$$\begin{aligned} I_2 &= I_1 - \operatorname{Re}\oint_{S_r} p_1|u|^2\, ds \\ &\geqq (2\beta-\gamma) r^{-(\beta+1)} \int_{D_r} |u_\rho|^2\, dx - (n+\beta-2) r^{-(\beta+1)} \int_{D_r} \rho^\beta |\operatorname{grad} u|^2\, dx \\ &\quad + r^{-(\beta+1)} \int_{D_r} \rho^\beta [\rho p_{0\rho} + (n+\beta) p_0] |u|^2\, dx - \frac{K^2}{\gamma} r^{-(\beta+1)} \int_{D_r} \rho^\beta |u|^2\, dx \\ &\quad - \operatorname{Re}\oint_{S_r} p_1|u|^2\, ds + \left[-2Kr^{-(\beta+1)} \operatorname{Re}\int_{D_r - D_{r_1}} \rho^{\beta-1} x_i u_{,i} \bar{u}\, dx \right. \\ &\quad \left. + 2Kr^{-(\beta+1)} \int_{D_r - D_{r_1}} \rho^{\beta-1} x_i u_{,i} \bar{u}\, dx\right]. \end{aligned} \tag{1.30}$$

Using the divergence theorem and the arithmetic-geometric mean inequality on the last two terms of (1.30), as in Lemma 1.2, gives

$$-2Kr^{-(\beta+1)} \operatorname{Re}\int_{D_r} \rho^{\beta-1} x_i u_{,i} \bar{u}\, dx = Kr^{-1}\oint_{S_r} |u|^2\, ds + \varepsilon r^{-(\beta+1)} \int_{D_r} \rho^\beta |u|^2\, dx, \tag{1.31}$$

and

$$\begin{aligned} 2Kr^{-(\beta+1)} \operatorname{Re}\int_{D_r} \rho^{\beta-1} x_i u_{,i} \bar{u}\, dx &\geqq -(r)^{-(\beta+1)} \frac{K^2}{\gamma} \int_{D_r} \rho^\beta |u|^2\, dx \\ &\quad - (r)^{-(\beta+1)} \gamma \int_{D_r} \rho^\beta |u_\rho|^2\, dx. \end{aligned} \tag{1.32}$$

By substituting (1.31) and (1.32) into (1.30) and noting that

$$\frac{K}{r}\oint_{S_r} |u|^2\, ds - \operatorname{Re}\oint_{S_r} p_1|u|^2\, ds \geqq 0$$

we obtain

$$
\begin{aligned}
I_2 \geqq\ & 2(\beta-\gamma) r^{-(\beta+1)} \int_{D_r} \rho^{\beta} |u_\rho|^2 \, dx - (n+\beta-2) r^{-(\beta+1)} \int_{D_r} \rho^{\beta} |\operatorname{grad} u|^2 \, dx \\
& + r^{-(\beta+1)} \int_{D_r} \rho^{\beta} [\rho p_{0\rho} + (n+\beta) p_0] |u|^2 \, dx \\
& - \left(\frac{2K^2}{\gamma} + \varepsilon \right) r^{-(\beta+1)} \int_{D_r} \rho^{\beta} |u|^2 \, dx.
\end{aligned} \tag{1.33}
$$

We are now in a position to apply Lemma 1.3 and 1.4. For convenience we treat the cases $-1 < \beta \leqq 0$ and $\beta \geqq 0$ separately.

Case 1. $-1 < \beta \leqq 0$. In this case we replace $|u_\rho|$ by $|\operatorname{grad} u|$ and use Lemma 1.3. Thus

$$
\begin{aligned}
I_2 \geqq\ & \frac{n-\beta-2+2\gamma}{1-\varepsilon_1} r^{-1} \operatorname{Re} \oint_{S_r} u_\rho \bar{u} \, ds \\
& + r^{-(\beta+1)} \int_{D_r} \rho^{\beta} \left[\rho p_{0\rho} + \frac{2+2\beta-2\gamma+\varepsilon_2}{1-\varepsilon_1} p_0 - \frac{2K^2}{\gamma} - \varepsilon_3 \right] |u|^2 \, dx.
\end{aligned} \tag{1.34}
$$

We now replace I_2 in (1.27) by (1.34) and write everything in terms of t to obtain

$$
\begin{aligned}
FF_{tt} - (F_t)^2 \geqq\ & -\frac{n-\beta-2+2\gamma}{(n-2)(1-\varepsilon_1)} t^{-1} F F_t \\
& + \varepsilon_2 t^{-(n-\beta-2)/(n-2)} \int_0^t \eta^{-(2n-2+\beta)/(n-2)} F(\eta) \, d\eta.
\end{aligned} \tag{1.35}
$$

We are prepared to use the main lemma with $\alpha_1 = (n-\beta-2)/(n-2)$, $\alpha_2 = (2n-2+\beta)/(n-2)$ and $C_1 = (n-\beta-2+2\gamma)/(n-2)$. Thus from the main lemma, it follows that $F(t) \equiv 0$ for $0 \leqq t \leqq t_0$ which in turn implies that $u \equiv 0$ for $0 \leqq T \leqq t_0$ which is contrary to our assumption. Thus $u \equiv 0$ in D_{R_0}. By the previously mentioned unique continuation theorem it follows that $u \equiv 0$ in Ω for $-1 < \beta \leqq 0$.

Case 2. $\beta > 0$. In this case we first recall Note 1 and write

$$
\begin{aligned}
I_2 \geqq\ & (n+\beta-2) r^{-1} \operatorname{Re} \oint_{S_r} u_\rho \bar{u} \, ds + [2(\beta-\gamma) - \varepsilon_1] r^{-(\beta+1)} \int_{D_r} \rho^{\beta} |u_\rho|^2 \, dx \\
& + r^{-(\beta+1)} \int_{D_r} \rho^{\beta} \left[\rho p_{0\rho} + 2p_0 - \frac{2K^2}{\gamma} - \varepsilon_2 \right] |u|^2 \, dx.
\end{aligned} \tag{1.36}
$$

We now apply Lemma 1.4 and using (1.13(ii)) we obtain

$$
I_2 \geqq \left[\frac{2n-4+\varepsilon}{(n-2)(2-\varepsilon)} \right] t^{-1} F F_t + \varepsilon t^{-(n-\beta-2)/(n-2)} \int_0^t \eta^{-(2n-2+\beta)/(n-2)} F(\eta) \, d\eta. \tag{1.37}
$$

Again as in Case 1 we have

$$
\begin{aligned}
FF_{tt} - (F_t)^2 \geqq\ & \frac{2n-4+\varepsilon}{(n-2)(2-\varepsilon)} t^{-1} F F_t \\
& + \varepsilon t^{-(n-\beta-2)/(n-2)} \int_0^t \eta^{-(2n-2+\beta)/(n-2)} F(\eta) \, d\eta
\end{aligned} \tag{1.38}
$$

and the rest of the proof of Case 2 follows as in Case 1.

This completes the proof of Theorem 1.1. Q.E.D.

The following corollary follows immediately from Theorem 1.1.

COROLLARY 1.2. *Let u be a nontrivial C^2 solution of* (1.3) *in* Ω. *Suppose* $p(x)$ *satisfies* (1.4) *and* (1.5) *together with* (1.13(i)) *or* (1.13(ii)); *then*

$$\lim_{R\to\infty} \int_{R_0}^{R} r^{\beta} \oint_{S_r} p_0|u|^2 \, ds\, dr = \infty.$$

It should be noted that although our results have features in common with those of Agmon [1] and Kato [6], they nevertheless differ in certain respects. In the first place we are able to deal with a wider range of β's in Theorem 1.1. The conditions imposed on p_0 and p_1 are slightly different from those imposed by Agmon [1]. This allows us to treat some cases which his methods does not handle. Our method easily generalizes to systems as we shall see in the next section, and it will be shown in subsequent papers that there is a reasonably straightforward generalization to more general elliptic operators and nonlinear elliptic inequalities.

2. Applications. In this section we shall first utilize the proof of Theorem 1.1 to generalize the results of § 1 to systems of Schrödinger type equations and then state a theorem concerning the positive eigenvalue problem for Schrödinger type operators.

A. Systems of Schrödinger type equations. We consider systems of the following form:

$$\Delta u_i + p_{ij} u_j = 0 \quad \text{in } \Omega \quad \text{for } i = 1, \cdots, m, \tag{2.1}$$

where Ω is again an open connected region in $\mathbb{R}^n (n \geqq 3)$ and contains $D_{R_0} = \{x \mid \|x\| \geqq R_0\}$, $u(x) = \{u_1(x), \cdots, u_m(x)\}$ with $u_i(x)$ a complex-valued function in $C^2(\Omega)$ and $(p_{ij}(x))$ a complex-valued matrix which will be described later.

We shall study the asymptotic behavior of the solution of (2.1) as $\|x\| \to \infty$.

For $m = 1$ equation (2.1) clearly has the general form of (1.3). Thus the method of proving nonexistence results for (2.1) is basically the same as (1.3).

By a simple application of the result of (2.1) one can establish nonexistence of solutions (in an appropriate class) of equations of the form $\sum_{i=0}^{n} a_i \Delta^{n-i} u = 0$ in exterior domain.

2.1. Preliminaries and hypotheses. We shall consider $(p_{ij}(x))$ to have the following decomposition:

$$p_{ij}(x) = s_{ij}(x) + t_{ij}(x), \qquad \|x\| \geqq R_0. \tag{2.2}$$

We shall assume that $(s_{ij}(x))$ is a real-valued, strongly positive definite $m \times m$-matrix, the matrix $(s_{ij}(x))$ is symmetric and there exists a positive constant K_0 such that for all vectors η_i the inequality

$$s_{ij} \eta_i \bar{\eta}_j \geqq K_0 \eta_i \bar{\eta}_i \tag{2.3}$$

holds at every point in D_{R_0}.

The matrix $(t_{ij}(x))$ is assumed to be complex-valued and for sufficiently large R_0 satisfies the inequality

$$\sup_{\|x\|=r} |t_{ij}(x)\eta_i\bar{\eta}_j| \leqq \frac{K}{r}\eta_i\bar{\eta}_i, \qquad r \geqq R_0, \tag{2.4}$$

for positive constant K.

Furthermore, the components of the matrix $(s_{ij}(x))$ are assumed to possess continuous radial derivatives and to satisfy for any complex vector η_i the inequality

$$\left[r\frac{\partial}{\partial r}s_{ij} + \delta_1 s_{ij}\right]\eta_i\bar{\eta}_j \geqq \delta_2\eta_i\bar{\eta}_i \quad \forall x \in D_{R_0}, \tag{2.5}$$

where δ_1 and δ_2 are two positive numbers given explicitly by

$$\text{for } -1 < \beta \leqq 0, \begin{cases} \delta_1 = 2(\beta + 1 - \gamma), \\ \delta_2 = \dfrac{2K^2}{\gamma} + \varepsilon, \end{cases} \tag{2.5}'$$

$$\text{for } 0 < \beta, \begin{cases} \delta_1 = 2 - \varepsilon, \\ \delta_2 = \dfrac{K^2}{\beta} + \varepsilon_1. \end{cases} \tag{2.5}''$$

In (2.5), γ is an arbitrary positive constant and the ε's are positive arbitrary small numbers.

Remark. Assumption (2.5) together with (1.7) implies that $s_{ii} > 0$ for $x \in D_{R_0}$. Thus a sufficient condition for (2.3) would be

$$s_{ii}(x) \geqq \sum_{\substack{j=i \\ j\neq i}} |s_{ij}(x)| \quad \forall x \in D_{R_0}. \tag{2.6}$$

2.2. The main result. We are now ready to state the main theorem.

THEOREM 2.1. *Let $u_i \in C^2(\Omega)$ be a solution of* (2.1) *for $(p_{ij}(x))$ satisfying* (2.2), (2.3), (2.4), *and* (2.5). *If for some constant β, $\int_{R_0}^{\infty} \rho^{\beta} \oint_{S_\rho} s_{ij}u_i\bar{u}_j\, ds\, d\rho < \infty$ and if for R_0 sufficiently large, $(s_{ij}(x))$ satisfies either* (2.5)′ *or* (2.5)″, *then u must vanish identically.*

The proof of this theorem utilizes lemmas similar to Lemmas 1.1 through 1.4. For instance the equivalent of Lemma 1.1 is the following:

LEMMA 2.1. *Let u_i and $(s_{ij}(x))$ satisfy* (2.1) *and* (2.3). *Then*

$$\int_{R_0-1}^{\infty} \rho^{\beta} \oint_{S_\rho} S_{ij}u_i\bar{u}_j\, ds\, d\rho \leqq M_1 \Rightarrow \int_{R_0}^{\infty} \rho^{\beta} \oint_{S_\rho} |\text{grad } u|^2\, ds\, d\rho \leqq M_2 \tag{2.7}$$

for some positive M_1 and M_2. (Notations: $|grad\, u|^2 = u_{i,j}\bar{u}_{i,j}$).

Finally we obtain an inequality similar to (1.38) and use the main lemma to show that u vanishes identically in Ω.

COROLLARY 2.1. *Let u be a nontrivial C^2 solution of* (2.1) *in Ω for $(p_{ij}(x))$, $(s_{ij}(x))$ and β as in Theorem* 2.1. *Then, for R_0 sufficiently large,*

$$\lim_{R\to\infty}\int_{R_0}^{R}\rho^{\beta}\oint_{S_\rho} s_{ij}u_i\bar{u}_j\,ds\,d\rho=\infty.$$

B. On positive eigenvalues of $-\Delta+V$.

THEOREM 2.2. *Suppose $V(x)$ is a real valued function, satisfying the following conditions:*

(i) *$V(x)$ is locally $L_2(\Omega)$, (Ω as in Theorem* 1.1).

(ii) *$V(x)$ is locally Hölder continuous in a connected open set $\Omega_0\subset\Omega$ where $\Omega_0\supset\{x\,|\,\|x\|\geqq R_0\}$ and $\Omega-\Omega_0$ is of measure zero.*

(iii) *$V(x)$ has, for $\|x\|\geqq R_0$, the decomposition*

$$V(x)=V_0(x)+V_1(x), \tag{2.8}$$

where V_0 is a real continuous function, possessing a continuous radial derivative and $V_0(x)=o(1)$,

$$\limsup_{r\to\infty} r\frac{\partial V_0}{\partial r}=\Lambda_0 \quad\text{and}\quad \limsup_{r\to\infty}|rV_1(x)|\leqq K. \tag{2.9}$$

If H is the self-adjoint extension of $-\Delta+V$ in $L_2(\Omega)$, where $V(x)$ satisfies the above hypothesis, then H has no eigenvalue λ which satisfies

$$\lambda>\frac{K^2/\gamma+\Lambda_0/2}{1-\gamma},\qquad 0<\gamma<1.$$

Choosing the optimal value of γ, we find that there are no eigenvalues λ in the interval

$$\left(\left\{K+\sqrt{K^2+\frac{\Lambda_0}{2}}\right\}^2,\infty\right).$$

Remark 1. Under the conditions (i) and (ii) imposed on $V(x)$, $-\Delta+V$ is lower semi-bounded and essentially self-adjoint in $L_2(\Omega)$. Furthermore, if H is the unique self-adjoint extension of $-\Delta+V$ and H_0 the corresponding operator for the case $V(x)\equiv 0$, then $D(H)=D(H_0)$. (See Ikebe [5] or Kato [7].)

Remark 2. By self-adjoint extension we mean if $u\in D(H)$, then we have $\Delta u\in L_2(\Omega)$ in the distribution sense, $Vu\in L_2(\Omega)$ and $Hu=-\Delta u+Vu$.

Remark 3. If $V_1(x)=O(r^{-1-\varepsilon})$ for $\varepsilon>0$, then Theorem 2.2 is exactly Agmon's Theorem 4[1].

Proof of Theorem 2.2. Let $p_0(x)=\lambda-V_0(x)$, $p_1(x)=-V_1(x)$ and $\beta=0$. Then p_0 satisfies condition (1.13(i)) and p_1 satisfies (1.5). Thus Theorem 2.2 follows immediately from Theorem 1.1.

Example. Let

$$V(r)=-\frac{|\sin 2r|}{r}-\frac{7\sin 2r}{r}+O\left(\frac{1}{r^2}\right).$$

This is similar to the von Neumann–Wigner [9] potential. By Theorem 2.2, if we let

$$V_0(r) = -\frac{7\sin 2r}{r} \quad \text{and} \quad V_1(r) = -\frac{|\sin 2r|}{r} + O\left(\frac{1}{r^2}\right),$$

then $\lambda > 14.656$, meaning that there exists no eigenvalue for the operator $-\Delta + V$ larger than 14.656. Kato's result gives a lower bound of 64.

Appendix.

A. Proof of main Lemma. We first show that if F vanishes at any point in $(0, t_0]$, then it vanishes identically in $(0, t_0]$. To demonstrate this we suppose that $F(t_1) = 0$ for some $t_1 \in (0, t_0]$ and $F(t) > 0$ in the interval $[t_1 + \varepsilon, t_2 - \varepsilon]$ for some arbitrarily small $\varepsilon > 0$ and some $t_2 \in (t_1, t_0]$. Then from (1.2) we have

$$FF'' - (F')^2 > -C_1 t^{-1} FF'. \tag{A.1}$$

Setting σt^{1-C_1}, denoting $\tilde{F}(\sigma(t)) = F(t)$, we find that

$$\frac{d^2}{d\sigma^2}[\log \tilde{F}(\sigma)] > 0, \tag{A.2}$$

from which it follows, for $\sigma \in [\sigma_1, \sigma_2]$ where $\sigma_1 = (t_1 + \varepsilon)^{1-C_1}$ and $\sigma_2 = (t_2 - \varepsilon)^{1-C}$, that

$$\tilde{F}(\sigma) < [\tilde{F}(\sigma_1)]^{(\sigma_2-\sigma)/(\sigma_2-\sigma_1)}[\tilde{F}(\sigma_2)]^{(\sigma-\sigma_1)/(\sigma_2-\sigma_1)}, \tag{A.3}$$

or

$$\begin{aligned} F(t) < {} & [F(t_1+\varepsilon)]^{[(t_2-\varepsilon)^{1-C_1}-(t)^{1-C_1}]/[(t_2-\varepsilon)^{1-C_1}-(t_1+\varepsilon)^{1-C_1}]} \\ & \cdot [F(t_2-\varepsilon)]^{[(t)^{1-C_1}-(t_1+\varepsilon)^{1-C_1}]/[(t_2-\varepsilon)^{1-C_1}-(t_1+\varepsilon)^{1-C_1}]}. \end{aligned} \tag{A.4}$$

Letting $\varepsilon \to 0$ in (A.4) and F being a continuous function of t implies that $F(t) \equiv 0$ in $[t_1, t_2]$. By the same argument we can conclude that $F(t) \equiv 0$ in $(0, t_0]$. Since $F(0) = 0$, therefore $F(t) \equiv 0$ in $[0, t_0]$.

Suppose, now, that $F(t)$ is positive for every point in $(0, t_0]$ and satisfies (1.1) and (1.2). We shall establish a contradiction.

We, first, assume $F'(t) > 0$ $\forall t \in (0, t_0]$. Then from (1.2) we have

$$\left[\left(\frac{t_0}{t}\right)^{-C_1}\frac{F'}{F}\right]' \geqq \varepsilon t^{-\alpha_1}\left(\frac{t_0}{t}\right)^{-C_1}\frac{1}{F}\int_0^t \eta^{-\alpha_2}F(\eta)\,d\eta. \tag{A.5}$$

Multiplying (A.5) by $(t_0/t)^{-C_1}F'/F$ and integrating over the interval $[t, t_0]$, we obtain

$$\begin{aligned} \left[\frac{F'(t_0)}{F(t_0)}\right]^2 - \left[\left(\frac{t_0}{t}\right)^{-C_1}\frac{F'(t)}{F(t)}\right]^2 \geqq {} & 2\varepsilon t_0^{-2C_1}\int_t^{t_0} \xi^{-(-2C_1+\alpha_1)}\frac{F'(\xi)}{F^2(\xi)} \\ & \cdot \left(\int_0^{\xi} \eta^{-\alpha_2}F(\eta)\,d\eta\right)d\xi \\ \geqq {} & C_2\int_t^{t_0}\frac{F'(\xi)}{F(\xi)}\left(\int_0^{\xi}\eta^{-\tilde{\alpha}_2}F(\eta)\,d\eta\right)d\xi, \end{aligned} \tag{A.6}$$

where

$$C_2=\begin{cases}2\varepsilon/t_0^{\alpha_1}, & -2C_1+\alpha_1\geqq 0,\\ 2\varepsilon t_0^{-2C_1}, & -2C_1+\alpha_1<0,\end{cases}\qquad \tilde{\alpha}_2=\begin{cases}\alpha_2, & -2C_1+\alpha_1\geqq 0,\\ -2C_1+\alpha_1+\alpha_2, & -2C_1+\alpha_1<0.\end{cases}$$

Integrating the right hand side of (A.6) by parts we obtain

$$\left[\frac{F'(t_0)}{F(t_0)}\right]^2-\left[\left(\frac{t_0}{t}\right)^{-C_1}\frac{F'(t)}{F(t)}\right] \geqq C_2\left[-\frac{1}{F(\xi)}\int_0^{\xi}\eta^{-\tilde{\alpha}_2}F(\eta)\,d\eta\Big|_t^{t_0}\right]+C_2\int_t^{t_0}\xi^{-\tilde{\alpha}_2}\,d\xi \tag{A.7}$$

or

$$\left[\frac{F'(t_0)}{F(t_0)}\right]^2+\frac{C_2}{F(t_0)}\int_0^{t_0}\eta^{-\tilde{\alpha}_2}F(\eta)\,d\eta+(\tilde{\alpha}_2-1)C_2t_0^{1-\tilde{\alpha}_2} \geqq\left[\left(\frac{t_0}{t}\right)^{-C_1}\frac{F'(t)}{F(t)}\right]^2+(\tilde{\alpha}_2-1)C_2t^{(1-\tilde{\alpha}_2)}+\frac{C_2}{F(t)}\int_0^{t}\eta^{-\tilde{\alpha}_2}F(\eta)\,d\eta. \tag{A.8}$$

We note that the left hand side of (A.8) is positive and constant while the right hand side is nonnegative and a function of t. Clearly the right hand side tends to $+\infty$ as $t\to 0$ and we are led to a contradiction.

Now, suppose $F'(\hat{t})\leqq 0$ for some $\hat{t}\in(0, t_0]$. To draw a contradiction in this case, we consider (A.1) from which, for $\hat{t}\in(0, t_0]$, we have

$$\left[\left(\frac{\hat{t}}{t}\right)^{-C_1}\frac{F'(t)}{F(t)}\right]'>0. \tag{A.9}$$

Upon integration over the interval $(t, \hat{t})$ we obtain

$$\frac{F'(\hat{t})}{F(\hat{t})}-\left(\frac{\hat{t}}{t}\right)^{-C_1}\frac{F'(t)}{F(t)}>0, \tag{A.10}$$

or

$$\frac{F'(\hat{t})}{F(\hat{t})}\left(\frac{t}{\hat{t}}\right)^{-C_1}F(t)-F'(t)>0. \tag{A.11}$$

Integrating (A.11) again we have

$$\int_t^{\hat{t}}\left[F(\xi)\exp\left[\frac{F'(\hat{t})}{F(\hat{t})}\int_{\xi}^{\hat{t}}\left(\frac{\eta}{\hat{t}}\right)^{-C_1}d\eta\right]\right]'d\xi<0, \tag{A.12}$$

from which it follows that

$$F(\hat{t})-\exp\left[\frac{F'(\hat{t})}{F(\hat{t})}\int_t^{\hat{t}}\left(\frac{\eta}{\hat{t}}\right)^{-C_1}d\eta\right]F(t)<0; \tag{A.13}$$

thus

$$F(t)>F(\hat{t})\exp\left[\frac{F'(\hat{t})}{F(\hat{t})}\int_t^{\hat{t}}\left(\frac{\eta}{\hat{t}}\right)^{-C_1}d\eta\right], \tag{A.14}$$

which implies that $F(t)>F(\hat{t})$, $\forall t\leqq\hat{t}$ which in turn is in contradiction to (1.1). This completes the proof. Q.E.D.

B. Proof of Lemma 1.4. Let α_1 and α_2 be two, as yet arbitrary, positive numbers, and set

$$J(r) = -(2-\alpha_1)\int_{D_r} \rho^\beta |u_\rho|^2\,dx + \int_{D_r} \rho^\beta |\operatorname{grad} u|^2\,dx - (1-\alpha_2)\int_{D_r} \rho^\beta p_0 |u|^2\,dx \tag{B.1}$$

for any $r \geqq r_0$. Then differentiating with respect to r we get

$$J_r(r) = (2-\alpha_1) r^\beta \oint_{S_r} |u_\rho|^2\,ds - r^\beta \oint_{S_r} |\operatorname{grad} u|^2\,ds + (1-\alpha_2) r^\beta \oint_{S_r} p_0 |u|^2\,ds. \tag{B.2}$$

By Lemma 1.2 we have

$$\begin{aligned} rJ_r &= -r^{\beta+1} I_1 - \alpha_1 r^{\beta+1} \oint_{S_r} |u_\rho|^2\,ds - \alpha_2 r^{\beta+1} \oint_{S_r} p_0 |u|^2\,ds \\ &\leqq -(2\beta-\gamma)\int_{D_r} \rho^\beta |u_\rho|^2\,dx + (n+\beta-2)\int_{D_r} \rho^\beta |\operatorname{grad} u|^2\,dx \\ &\quad -\int_{D_r} \rho^\beta \left[\rho p_{0\rho} + (n+\beta) p_0 - \frac{K^2}{\gamma}\right] |u|^2\,dx - \alpha_1 r^{\beta+1} \oint_{S_r} |u_\rho|^2\,ds \\ &\quad -\alpha_2 r^{\beta+1} \oint_{S_r} p_0 |u|^2\,ds. \end{aligned} \tag{B.3}$$

Using (1.18) we now find

$$\begin{aligned} rJ_r &\leqq -(2\beta-\gamma)\int_{D_r} \rho^\beta |u_\rho|^2\,dx + (n+\beta-2)\left[\varepsilon_5 r^{\beta+1} \oint_{S_r} |u|^2\,ds \right. \\ &\quad + \varepsilon_5 r^{\beta+1} \oint_{S_r} |u_\rho|^2\,ds + \varepsilon_3 \int_{D_r} \rho^\beta |u|^2\,dx + \varepsilon_4 \int_{D_r} \rho^\beta |u_\rho|^2\,dx \\ &\quad \left. + \int_{D_r} \rho^\beta p_0 |u|^2\,dx\right] - \int_{D_r} \rho^\beta \left[\rho p_{0\rho} + (n+\beta) p_0 - \frac{K^2}{\gamma}\right] |u|^2\,dx \\ &\quad -\alpha_1 r^{\beta+1} \oint_{S_r} |u_\rho|^2\,ds - \alpha_2 r^{\beta+1} \oint_{S_r} p_0 |u|^2\,ds, \end{aligned} \tag{B.4}$$

or

$$\begin{aligned} rJ_r &\leqq -(2\beta-\gamma-\hat{\varepsilon}_4)\int_{D_r} \rho^\beta |u_\rho|^2\,dx \\ &\quad -\int_{D_r} \rho^\beta \left[\rho p_{0\rho} + 2p_0 - \frac{K^2}{\gamma} - \hat{\varepsilon}_3\right] |u|^2\,dx \\ &\quad -(\alpha_1 - \hat{\varepsilon}_5) r^{\beta+1} \oint_{S_r} |u_\rho|^2\,ds - \oint_{S_r} \rho^{\beta+1} [\alpha_2 p_0 - \hat{\varepsilon}_5] |u|^2\,ds, \end{aligned} \tag{B.5}$$

where $\hat{\varepsilon}_i = (n+\beta-2)\varepsilon_i$.

In order to insure that $rJ_r \leqq 0$, we impose the conditions

$$\text{(B.6)} \qquad \begin{cases} \text{(a)} \quad 2\beta - \gamma - \hat{\varepsilon}_4 \geqq 0 \quad \text{or} \quad 2\beta - \hat{\varepsilon}_4 \geqq \gamma, \\ \text{(b)} \quad \alpha_1 = \hat{\varepsilon}_5, \qquad \alpha_2 = \hat{\varepsilon}_5 / p_0 = \varepsilon_6, \\ \text{(c)} \quad \rho p_{0\rho} + 2p_0 - \dfrac{K^2}{\gamma} - \hat{\varepsilon}_3 \geqq 0. \end{cases}$$

Condition (c) follows from (1.13(ii)) and (B.6(a)). Now

$$\text{(B.7)} \qquad \begin{aligned} rJ_r \leqq 0 &\Rightarrow \int_r^{r_1} J_\rho \, d\rho \leqq 0, \\ &\Rightarrow J(r_1) - J(r) \leqq 0. \end{aligned}$$

Since $\lim_{r_1 \to \infty} J(r_1) = 0$, (B.7) implies $J(r) \geqq 0$, i.e.,

$$\text{(B.8)} \qquad \int_{D_r} \rho^\beta |u_\rho|^2 \, dx \leqq \frac{1}{2-\varepsilon_6} \int_{D_r} \rho^\beta |\operatorname{grad} u|^2 \, dx - \frac{1-\hat{\varepsilon}_5}{2-\varepsilon_6} \int_{D_r} \rho^\beta p_0 |u|^2 \, dx.$$

By (1.17), Note 1, this leads to

$$\text{(B.9)} \qquad \begin{aligned} \int_{D_r} \rho^\beta |u_\rho|^2 \, dx \leqq \frac{1}{2-\varepsilon_6} \Big[& -r^\beta \operatorname{Re} \oint_{S_r} \bar{u}_\rho u \, ds + \varepsilon_3 \int_{D_r} \rho^\beta |u|^2 \, dx \\ & + \varepsilon_4 \int_{D_r} \rho^\beta |u_\rho|^2 \, dx + \int_{D_r} \rho^\beta p_0 |u|^2 \, dx \Big] \\ & - \frac{1-\hat{\varepsilon}_5}{2-\varepsilon_6} \int_{D_r} \rho^\beta p_0 |u|^2 \, dx, \end{aligned}$$

or

$$\text{(B.10)} \qquad \begin{aligned} (2-\varepsilon_6-\varepsilon_4) \int_{D_r} \rho^\beta |u_\rho|^2 \, dx \leqq & -r^\beta \operatorname{Re} \oint_{S_r} \bar{u}_\rho u \, ds + \varepsilon_3 \int_{D_r} \rho^\beta |u|^2 \, dx \\ & + \hat{\varepsilon}_5 \int_{D_r} \rho^\beta p_0 |u|^2 \, dx, \end{aligned}$$

which may be written as

$$\text{(B.11)} \qquad \begin{aligned} \int_{D_r} \rho^\beta |u_\rho|^2 \, dx \leqq & -\frac{1}{2-\varepsilon_7} r^\beta \operatorname{Re} \oint_{S_r} \bar{u}_\rho u \, ds \\ & + \frac{\varepsilon_3}{2-\varepsilon_7} \int_{D_r} \rho^\beta |u|^2 \, dx + \frac{\hat{\varepsilon}_5}{2-\varepsilon_7} \int \rho^\beta p_0 |u|^2 \, dx, \end{aligned}$$

where we have set $\varepsilon_7 = \varepsilon_6 + \varepsilon_4$. Letting $\varepsilon_8 = \varepsilon_3/(2-\varepsilon_7)$ and $\varepsilon_9 = \hat{\varepsilon}_5/(2-\varepsilon_7)$ we establish (1.19). Q.E.D.

REFERENCES

[1] S. AGMON, *Lower bounds for solutions of Schrödinger equations*, J. Analyse Math., 23 (1970), pp. 1–25.

[2] ———, *Lower bounds for solutions of Schrödinger-type equations in unbounded domain*, Proc. Tokyo International Conference on Functional Analysis and Related Topics, (1969), pp. 216–224.

[3] ———, *Spectral properties of Schrödinger operators*, Actes, Congrés International Mathematiques, 1970, vol. 2, pp. 679–683.

[4] L. HORMANDER, *Linear Partial Differential Operators*, Springer-Verlag, Berlin, 1963.

[5] T. IKEBE, *Eigenfunction expansion associated with the Schrödinger operators and their applications to scattering theory*, Arch. Rational Mech. Anal., 5 (1960), pp. 1–34.

[6] T. KATO, *Growth properties of solutions of reduced wave equation with a variable coefficient*, Comm. Pure Appl. Math., 12 (1959), pp. 403–425.

[7] ———, *Some Mathematical Problems in Quantum Mechanics*, Progr. Theoret. Phys. Suppl., 40 (1967), pp. 3–14.

[8] G. B. KHOSROVSHAHI, *Growth properties of solutions of Schrödinger type systems*, Doctoral thesis, Cornell Univ., Ithaca, N.Y., 1972.

[9] J. VON NEUMANN AND E. P. WIGNER, *Über Merkwürdige Diskrete Eigenwerte*, Phys. Z., 50 (1929), pp. 465–467.

[10] L. E. PAYNE, *On a priori bounds in the Cauchy problem for elliptic equations*, this Journal, 1 (1970), pp. 82–89.

[11] F. RELLICH, *Uber das Asymptotische Verhalten der Lösungen von* $\Delta u + \lambda u = 0$ *in unendlichen Gebieten*, Jber. Deutsch. Math.-Verein, 53 (1943), pp. 57–65.

[12] B. SIMON, *On positive eigenvalues of one-body Schrödinger operators*, Comm. Pure Appl. Math., 22 (1967), pp. 531–538.

SIAM J. MATH. ANAL.
Vol. 8, No. 6, November 1977

ON A GENERALIZATION OF THE POISSON KERNEL FOR JACOBI POLYNOMIALS*

MIZAN RAHMAN†

Abstract. A symmetric, square-integrable, continuous and positive kernel, considered as a generalization of Bailey's Poisson kernel, is shown to have the Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ as eigenfunctions with eigenvalues expressed as infinite series with arbitrary sequence-coefficients. The consequent bilinear formula, via Mercer's theorem, includes, as special cases, Bailey's sum, Bateman's and Gegenbauer's degenerate addition formulas for Bessel functions and Feldheim's projection formula for Jacobi polynomials. In another special case the formula leads to the integrated version of a well-known addition theorem for the sum of a product of three Jacobi polynomials.

1. Introduction. In a recent series of papers [17]–[19] we made extensive use of Mercer's theorem [23] for square-integrable, continuous, positive symmetric kernels to derive some bilinear sums for Jacobi, Laguerre, Hahn and Meixner polynomials. With a given set of complete, orthonormal square-integrable functions $\{f_n(x)\}_{n=0}^{\infty}$ over some interval (a, b), finite or infinite, the method essentially depends on constructing a kernel $K(x, y)$ such that

$$\int_a^b K(x, y) f_n(y)\, dy = \lambda_n f_n(x), \qquad n = 0, 1, \cdots, \tag{1.1}$$

where the eigenvalues λ_n are positive with $\sum_{n=0}^{\infty} \lambda_n^2 < \infty$.

Apart from the fact that equation (1.1) may be read as a transformation of $f_n(x)$ to a multiple of itself and that such transformations belong to a whole class of what Al-Salam and Verma [1] call "orthogonality preserving transformations" there seems little hope of setting up a general method for constructing such kernels. One approach is to exploit the numerous differential and integral relations of the hypergeometric function since most classical orthogonal systems are expressible in terms of ordinary or confluent hypergeometric functions. This is what is essentially involved in the so-called shift-operator or ladder-operator method [16], [5]. A slightly more general approach is to make use of the known fractional derivatives of the $f_n(x)$ [14]. Both these methods usually lead to fairly simple expressions for the eigenvalues λ_n, often a ratio of products of Pochhammer functions like $(a)_n = a(a+1)\cdots\cdots(a+n-1)$. However, the eigenvalues for the Jacobi and Hahn polynomials in [17]–[19] were found to be balanced ${}_4F_3(1)$ series (commonly known in the literature as Saalschutzian) which, of course, reduce to ratios of Pochhammer functions in special limiting cases. Our method was, in that sense, a generalization of the ladder-operator method for integer-valued parameters, and that of the fractional-derivatives method for nonintegral complex-valued parameters.

The positive kernels and the corresponding bilinear sums obtained in [17], [18] might be considered as generalizations of the Bateman sum [3] since the latter sum could be obtained as a special case of our general result. The presence of five

* Received by the editors October 3, 1975, and in revised form June 29, 1976.

† Department of Mathematics, Carleton University, Ottawa, Canada, K1S 5B6. This work was supported by The National Research Council of Canada under Grant A6197.

parameters in our formula enabled us to derive a number of other known results as special cases. However, in spite of the freedom of five parameters, it was quite clear that we could not choose them in any way to reduce our result to the well-known Poisson kernel (see Bailey [4, p. 102]):

$$\sum_{n=0}^{\infty} \frac{t^n P_n^{(\alpha,\beta)}(x) P_n^{(\alpha,\beta)}(y)}{h_n^{(\alpha,\beta)}}$$

$$= \frac{(1-t)\Gamma(\alpha+\beta+2)}{2^{\alpha+\beta+1}\Gamma(\alpha+1)\Gamma(\beta+1)(1+t)^{\alpha+\beta+2}} \cdot F_4\left(\frac{\alpha+\beta+2}{2}, \frac{\alpha+\beta+3}{2}; \quad \alpha+1, \beta+1; \frac{a^2}{r^2}, \frac{b^2}{r^2}\right), \tag{1.2}$$

where $x = \cos 2\varphi$, $y = \cos 2\psi$, $a = \sin\varphi \sin\psi$, $b = \cos\varphi \cos\psi$, $r = \frac{1}{2}(t^{1/2} + t^{-1/2})$ and

$$F_4(a_1, a_2; b_1, b_2; u, v) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(a_1)_{m+n}(a_2)_{m+n}}{(b_1)_m (b_2)_n m! n!} u^m v^n \tag{1.3}$$

is an Appell function , $P_n^{(\alpha,\beta)}(x)$ being the standard Jacobi polynomial, with $h_n^{(\alpha,\beta)}$ the normalizing constant for $P_n^{(\alpha,\beta)}(x)$. The positivity of the kernel on the right hand side of (1.2) is obvious for $\alpha > -1$, $\beta > -1$ and $0 \leqq t < 1$. It is also known to be positive in the range $-1 < t < 0$ for certain sets of values of α, β (see Askey [2]).

The F_4 function does have a double integral representation [4], but not the type we considered in [17]–[18]. It is obvious that to generalize (1.2) we need a different approach. An immediate generalization of (1.2) is $F_4(\alpha_1, \beta_1; \alpha+1, \beta+1; \rho xy, \rho(1-x)(1-y))$ where α_1, β_1, ρ are arbitrary. A further generalization, however, would be to consider the double series [13]

$$F\left[\begin{matrix}\alpha_1, \cdots, \alpha_p; & — ; & — \\ \beta_1, \cdots, \beta_q; & \alpha+1; & \beta+1\end{matrix}; \rho xy, \rho(1-x)(1-y)\right]$$

$$= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(\alpha_1)_{m+n} \cdots (\alpha_p)_{m+n}}{(\beta_1)_{m+n} \cdots (\beta_q)_{m+n}} \frac{\rho^{m+n}}{(\alpha+1)_m (\beta+1)_n} \frac{(xy)^m}{m!} \frac{\{(1-x)(1-y)\}^n}{n!}. \tag{1.4}$$

We shall assume that this series converges absolutely and uniformly on the unit square $0 \leqq x \leqq 1, 0 \leqq y \leqq 1$ and $|\rho| < 1$. (Note that x, y in (1.4) are not the same as in (1.2)). This property will be ensured if we require

$$p \leqq q+2,$$

with Re $\alpha > -1$, Re $\beta > -1$, and that none of the denominator parameters is a negative integer. Note that the kernel defined by the double series (1.4) is positive if $\alpha_i > 0$, $i = 1, \cdots, p$; $\beta_j > 0$, $j = 1, \cdots, q$ and $0 \leqq \rho < 1$. Later we shall see that certain results will also hold for $\rho = 1$.

An even more general kernel, as pointed out by Ismail [15], can be defined by the double series

$$F(x, y) \equiv \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} a_{m+n} \frac{\rho^{m+n}}{(\alpha+1)_m (\beta+1)_n} \frac{(xy)^m}{m!} \frac{\{(1-x)(1-y)\}^n}{n!} \tag{1.5}$$

where $\{a_k\}$ is an arbitrary sequence of complex numbers. In fact, we shall first

show that the Jacobi polynomials $P_n^{(\alpha,\beta)}(1-2x)$ are eigenfunctions of the kernel (1.5) and then specialize to hypergeometric coefficients

$$a_k = \frac{(\alpha_1)_k \cdots (\alpha_p)_k}{(\beta_1)_k \cdots (\beta_p)_k}. \tag{1.6}$$

The basic purpose of the above generalizations is to identify a family of Poisson-type kernels as distinct from the Bateman-type kernels considered in [17], [18] which will, hopefully, bring out the connection between some known but apparently disconnected results in the literature as well as produce some new ones. In § 3 we will show, for example, that Bailey's sum (1.2), Feldheim's projection formulas [10], Bateman's addition formula [24] involving Jacobi polynomials and Bessel functions, and Bateman's generating function for Jacobi polynomials [20] are, in fact, special cases of the same general formula. As an application of our results we will show in § 4 how our kernel leads naturally to the trilinear sum of the Jacobi polynomials obtained previously by Gasper [11], [12] by an entirely different method.

2. Derivation of the connection relation. The connection relation (1.1) between the Jacobi polynomials with kernel (1.5) follows in a very straightforward manner.

Using the integral

$$\begin{aligned}&\int_0^1 dy\, y^{m+\alpha}(1-y)^{n+\beta}\, {}_2F_1(-k, k+\alpha+\beta+1; \alpha+1; y)\\ &\quad = B(m+\alpha+1, n+\beta+1)\, {}_3F_2\left[\begin{matrix}-k, k+\alpha+\beta+1, m+\alpha+1\\ \alpha+1, m+n+\alpha+\beta+2\end{matrix}; 1\right],\end{aligned} \tag{2.1}$$

$k, m, n = 0, 1, 2, \cdots$; $\operatorname{Re}\alpha > -1$, $\operatorname{Re}\beta > -1$, and the identity (see, for example, Slater [21, p. 76, (2.5.11)]),

$$\begin{aligned}&{}_3F_2\left[\begin{matrix}-k, k+\alpha+\beta+1, l-n+\alpha+1\\ \alpha+1, l+\alpha+\beta+2\end{matrix}; 1\right]\\ &\quad = \frac{(-l)_k(\beta+1)_k}{(l+\alpha+\beta+2)_k(\alpha+1)_k}\, {}_3F_2\left[\begin{matrix}-k, k+\alpha+\beta+1, -n\\ -l, \beta+1\end{matrix}; 1\right]\end{aligned} \tag{2.2}$$

where l is a nonnegative integer such that $l \geqq \max(k, n)$, we obtain

$$\begin{aligned}&\int_0^1 dy\, y^{\alpha}(1-y)^{\beta} F(x, y)\, {}_2F_1(-k, k+\alpha+\beta+1; \alpha+1; y)\\ &\quad = \frac{B(\alpha+1, \beta+1)(\beta+1)_k}{(\alpha+1)_k(\alpha+\beta+2)_k} \sum_{l=k}^{\infty} \frac{(-l)_k a_l}{(\alpha+\beta+k+2)_l} \frac{(\rho x)^l}{l!}\\ &\qquad \cdot \sum_{j=0}^{k} \frac{(-k)_j(k+\alpha+\beta+1)_j}{(-l)_j(\beta+1)_j j!} \sum_{n=j}^{l} \binom{l}{n} (-n)_j \left(\frac{1-x}{x}\right)^n.\end{aligned} \tag{2.3}$$

We should point out that l in (2.2) actually equals $m+n$ in (2.1) and hence $l \geqq n$, since $m \geqq 0$. Also, since $y^m(1-y)^n$ is a polynomial of degree $m+n=l$ in y the integral in (2.1) must vanish if $l < k$, by orthogonality. Noting that the last sum on

the r.h.s. (right-hand side) of (2.3) is simply a binomial series which sums to $l!x^{-l}(x-1)^j/(l-j)!$ and that

$$
\begin{aligned}
&{}_2F_1(-k, k+\alpha+\beta+1; \beta+1; 1-x) \\
&\quad = (-1)^k \frac{(\alpha+1)_k}{(\beta+1)_k} {}_2F_1(-k, k+\alpha+\beta+1; \alpha+1; x),
\end{aligned} \tag{2.4}
$$

we obtain the connection relation

$$
\begin{aligned}
&\int_0^1 K(x, y)\, {}_2F_1(-k, k+\alpha+\beta+1; \alpha+1; y)\, dy \\
&\quad = \lambda_k\, {}_2F_1(-k, k+\alpha+\beta+1; \alpha+1; x)
\end{aligned} \tag{2.5}
$$

where

$$K(x, y) = y^\alpha (1-y)^\beta F(x, y), \tag{2.6}$$

and

$$\lambda_k = B(\alpha+1, \beta+1) \sum_{l=0}^{\infty} \frac{a_{k+l}\rho^{k+l}}{(\alpha+\beta+2)_{2k+l} l!}. \tag{2.7}$$

In terms of the symmetric kernel

$$G(x, y) = K(x, y)\{x^\alpha y^{-\alpha}(1-x)^\beta (1-y)^{-\beta}\}^{1/2} \tag{2.8}$$

and the normalization constant

$$N_k^{(\alpha,\beta)} = \frac{(2k+\alpha+\beta+1)\Gamma(k+\alpha+1)\Gamma(k+\alpha+\beta+1)}{k!\Gamma^2(\alpha+1)\Gamma(k+\beta+1)} \tag{2.9}$$

equation (2.5) reads

$$\int_0^1 G(x, y) f_k(y)\, dy = \lambda_k f_k(x), \tag{2.10}$$

where

$$f_k(x) = \{N_k^{(\alpha,\beta)} x^\alpha (1-x)^\beta\}^{1/2}\, {}_2F_1(-k, k+\alpha+\beta+1; \alpha+1; x), \tag{2.11}$$

$k = 0, 1, \cdots$, are the complete system of orthonormal eigenfunctions of the symmetric kernel $G(x, y)$ over $L_2(0, 1)$.

3. The bilinear formulas. Let us assume that the sequence $\{a_k\}$ is such that $\sum \lambda_k^2 < \infty$, $G(x, y)$ is square-integrable, positive and even continuous, at least in any interval $\varepsilon_1 \leqq x, y \leqq 1-\varepsilon_2$, $\varepsilon_1, \varepsilon_2 > 0$. In the case of hypergeometric coefficients the conditions for $\sum \lambda_k^2 < \infty$ can be easily established. The λ_k takes the form

$$\lambda_k = B(\alpha+1, \beta+1) \sum_{l=0}^{\infty} \frac{(\alpha_1)_{k+l} \cdots (\alpha_p)_{k+l}}{(\beta_1)_{k+l} \cdots (\beta_q)_{k+l}} \cdot \frac{\rho^{k+l}}{(\alpha+\beta+2)_{2k+l} l!}. \tag{3.1}$$

Obviously λ_k is an entire function of ρ if $p < q+2$. If $p = q+2$ the series on the right of (3.1) converges for $|\rho| < 1$; if, further, $\gamma \equiv \operatorname{Re}(\alpha+\beta+2+\sum_{i=1}^{q} \beta_i - \sum_{j=1}^{q+2} \alpha_j) > 0$ the series converges even when $|\rho| = 1$. By

making an asymptotic analysis of λ_k for large k it can be shown that $\sum|\lambda_k|$ converges for $p \leqq q+2$ and $|\rho|<1$. Further, if $\gamma>\frac{1}{2}$ then $\sum|\lambda_k|$ converges for $p=q+2$, $|\rho|=1$, while if $\gamma>\frac{1}{4}\sum|\lambda_k|$ may not converge, but $\sum\lambda_k^2$ does.

Assuming, then, that the conditions for Mercer's theorem are all satisfied, the connection relation (2.10) immediately leads to the bilinear formula

$$F(x,y)=\sum_{k=0}^{\infty}\frac{2k+\alpha+\beta+1}{k+\alpha+\beta+1}\cdot\frac{(\alpha+1)_k(\alpha+\beta+2)_k}{(\beta+1)_k(\alpha+\beta+2)_{2k}k!}\sum_{l=0}^{\infty}\frac{a_{k+l}\rho^{k+l}}{(\alpha+\beta+2+2k)_l l!}$$
$$\cdot {}_2F_1(-k,k+\alpha+\beta+1;\alpha+1;x)\,{}_2F_1(-k,k+\alpha+\beta+1;\alpha+1;y). \tag{3.2}$$

For the hypergeometric coefficients (1.6) we get the sum for the F-function defined in (1.4):

$$F\left[\begin{matrix}\alpha_1,\cdots,\alpha_p;\ —;\ —\\ \beta_1,\cdots,\beta_q;\alpha+1;\beta+1\end{matrix};\rho xy,\rho(1-x)(1-y)\right]$$
$$=\sum_{k=0}^{\infty}\frac{2k+\alpha+\beta+1}{k+\alpha+\beta+1}\cdot\frac{(\alpha+1)_k(\alpha+\beta+2)_k\rho^k}{(\beta+1)_k(\alpha+\beta+2)_{2k}k!} \tag{3.3}$$
$$\cdot {}_pF_{q+1}\left(\begin{matrix}\alpha_1+k,\cdots,\alpha_p+k\\ \beta_1+k,\cdots,\beta_q+k,\alpha+\beta+2+2k\end{matrix};\rho\right)$$
$$\cdot {}_2F_1(-k,k+\alpha+\beta+1;\alpha+1;x)\,{}_2F_1(-k,k+\alpha+\beta+1;\alpha+1;y).$$

In terms of the standard Jacobi polynomials

$$P_k^{(\alpha,\beta)}(x)=\frac{(\alpha+1)_k}{k!}\,{}_2F_1\left(-k,k+\alpha+\beta+1;\alpha+1;\frac{1-x}{2}\right) \tag{3.4}$$

equation (3.3) reads

$$F\left[\begin{matrix}\alpha_1,\cdots,\alpha_p;\ —;\ —\\ \beta_1,\cdots,\beta_q;\alpha+1;\beta+1\end{matrix};\frac{\rho(1-x)(1-y)}{4},\frac{\rho(1+x)(1+y)}{4}\right]$$
$$=\sum_{k=0}^{\infty}\frac{2k+\alpha+\beta+1}{k+\alpha+\beta+1}\cdot\frac{(\alpha+\beta+2)_k k!}{(\alpha+1)_k(\beta+1)_k(\alpha+\beta+2)_{2k}}\cdot\frac{(\alpha_1)_k\cdots(\alpha_p)_k}{(\beta_1)_k\cdots(\beta_q)_k}\rho_k \tag{3.5}$$
$${}_pF_{q+1}\left(\begin{matrix}\alpha_1+k,\cdots,\alpha_p+k\\ \beta_1+k,\cdots,\beta_q+k,\alpha+\beta+2+2k\end{matrix};\rho\right)\cdot P_n^{(\alpha,\beta)}(x)P_n^{(\alpha,\beta)}(y),$$
$$-1<x<1,\quad -1<y<1.$$

In many cases of interest this formula will be valid even when $|x|$ or $|y|$ equals or exceeds 1.

Special cases. (i) $p=2$, $q=0$. In this case the F-series becomes the Appell function F_4 defined in (1.3) and we get Feldheim's result [10].

$$F_4\left(\alpha_1,\alpha_2;\alpha+1,\beta+1;\frac{\rho(1-x)(1-y)}{4},\frac{\rho(1+x)(1+y)}{4}\right)$$
$$=\sum_{k=0}^{\infty}\frac{k!(\alpha+\beta+2)_k}{(\alpha+1)_k(\beta+1)_k(\alpha+\beta+2)_{2k}}\cdot\frac{2k+\alpha+\beta+1}{k+\alpha+\beta+1}(\alpha_1)_k(\alpha_2)_k\rho^k \tag{3.6}$$
$$\cdot {}_2F_1(\alpha_1+k,\alpha_2+k;\alpha+\beta+2+2k;\rho)P_k^{(\alpha,\beta)}(x)P_k^{(\alpha,\beta)}(y).$$

This may be considered as an immediate generalization of Bailey's formula (1.2). If $\operatorname{Re}(\alpha+\beta+2-\alpha_1-\alpha_2)>0$ then this is valid also for $\rho=1$

$$
\begin{aligned}
&F_4\left(\alpha_1, \alpha_2; \alpha+1, \beta+1; \frac{(1-x)(1-y)}{4}, \frac{(1+x)(1+y)}{4}\right) \\
&\quad=\frac{\Gamma(\alpha+\beta+2)\Gamma(\alpha+\beta+2-\alpha_1-\alpha_2)}{\Gamma(\alpha+\beta+2-\alpha_1)\Gamma(\alpha+\beta+2-\alpha_2)} \\
&\quad\cdot \sum_{k=0}^{\infty} \frac{k!(\alpha+\beta+2)_k(\alpha_1)_k(\alpha_2)_k}{(\alpha+1)_k(\beta+1)_k(\alpha+\beta+2-\alpha_1)_k(\alpha+\beta+2-\alpha_2)_k} \\
&\qquad\cdot \frac{2k+\alpha+\beta+1}{k+\alpha+\beta+1}\cdot P_k^{(\alpha,\beta)}(x)P_k^{(\alpha,\beta)}(y).
\end{aligned} \tag{3.7}
$$

However, when $\alpha_1=(\alpha+\beta+2)/2$, $\alpha_2=(\alpha+\beta+3)/2$, we have $\alpha+\beta-\alpha_1-\alpha_2+2=-\frac{1}{2}$ so that (3.6) will hold only for $|\rho|<1$.

Using the quadratic transformation ([8, p. 101])

$$
\begin{aligned}
&{}_2F_1\left(\frac{\alpha+\beta+2+2k}{2}, \frac{\alpha+\beta+3+2k}{2}; \alpha+\beta+2+2k; \rho\right) \\
&\quad=(1-\rho)^{-1/2}\{\tfrac{1}{2}+\tfrac{1}{2}(1-\rho)^{1/2}\}^{-(\alpha+\beta+2k+1)}
\end{aligned} \tag{3.8}
$$

and setting

$$
\rho^{-1/2}=\tfrac{1}{2}(t^{1/2}+t^{-1/2}), \tag{3.9}
$$

we obtain, after some simplifications, Bailey's formula (1.2).

The case $\alpha_1=1$, $\alpha_2=\beta+1$ is perhaps the most interesting since it leads to the integrated version of a well-known addition theorem referred to in the Introduction. Since this case calls for some detailed investigation of the properties of $F_4(1, \beta+1; \alpha+1, \beta+1; u, v)$ as well as the Jacobi function of the second kind we shall treat it separately in the last section.

(ii) $p=1$, $q=0$. There are no convergence difficulties in this case and we have the bilinear formula

$$
\begin{aligned}
&\sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \frac{(\alpha_1)_{m+n}\rho^{m+n}}{(\alpha+1)_m(\beta+1)_n m!n!}\left\{\frac{(1-x)(1-y)}{4}\right\}^m\left\{\frac{(1+x)(1+y)}{4}\right\}^n \\
&\quad=\sum_{k=0}^{\infty}\frac{k!(\alpha_1)_k(\alpha+\beta+2)_k}{(\alpha+1)_k(\beta+1)_k(\alpha+\beta+2)_{2k}}\cdot\frac{2k+\alpha+\beta+1}{k+\alpha+\beta+1}\rho^k \\
&\qquad\cdot {}_1F_1(\alpha_1+k; \alpha+\beta+2+2k; \rho)P_k^{(\alpha,\beta)}(x)P_k^{(\alpha,\beta)}(y).
\end{aligned} \tag{3.10}
$$

This includes as special cases some well-known results with Bessel functions. Let $\alpha_1\to\infty$, $\rho\to 0^+$ in such a way that $\rho\alpha_1=t$, a finite positive number. Then

$$
\lim_{\substack{\alpha_1\to\infty \\ \rho\to 0^+}} (\alpha_1)_{m+n}\rho^{m+n}=t^{m+n}
$$

and

$$\lim_{\substack{\alpha_1\to\infty \\ \rho\to 0^+}} {}_1F_1(\alpha_1+k;\alpha+\beta+2+2k;\rho)=$$

$$\Gamma(\alpha+\beta+2+2k)t^{-(1/2)(\alpha+\beta+2k+1)}\cdot I_{\alpha+\beta+2k+1}(2\sqrt{t})$$

(see, for example, [8, p. 266]), where $I_\nu(z)$ is the modified Bessel function of order ν. Using the relation

$$I_\nu(z)=\frac{(z/2)^\nu}{\Gamma(\nu+1)}\,{}_0F_1\left(-;\nu+1;\frac{z^2}{4}\right)=\frac{(z/2)^\nu e^{-z}}{\Gamma(\nu+1)}\,{}_1F_1\left(\nu+\frac{1}{2};2\nu+1;2z\right) \tag{3.11}$$

([9, p. 5]), and simplifying, we obtain

$$\begin{aligned}&\sqrt{t}I_\alpha(\sqrt{(1-x)(1-y)}t)I_\beta(\sqrt{(1+x)(1+y)}t)\\ &=\left[\frac{(1-x)(1-y)}{4}\right]^{\alpha/2}\left[\frac{(1+x)(1+y)}{4}\right]^{\beta/2}\sum_{k=0}^{\infty}\frac{k!\Gamma(\alpha+\beta+2+k)}{\Gamma(\alpha+1+k)\Gamma(\beta+1+k)}\\ &\quad\cdot\frac{2k+\alpha+\beta+1}{k+\alpha+\beta+1}I_{\alpha+\beta+2k+1}(2\sqrt{t})P_k^{(\alpha,\beta)}(x)P_k^{(\alpha,\beta)}(y).\end{aligned} \tag{3.12}$$

We note that this result could have been obtained directly from (3.2) by setting $a_k=(t/\rho)^k$. The limiting procedures employed here would then be unnecessary. For $t>0$ and $|x|\leqq 1$, $|y|\leqq 1$, the kernel on the left is positive. Writing $x=\cos 2\varphi$, $y=\cos 2\psi$, $2\sqrt{t}=z$, we get

$$\begin{aligned}&\frac{1}{2}zI_\alpha(z\sin\varphi\sin\psi)I_\beta(z\cos\varphi\cos\psi)\\ &=\sin^\alpha\varphi\sin^\alpha\psi\cos^\beta\varphi\cos^\beta\psi\\ &\quad\cdot\sum_{k=0}^{\infty}\frac{k!\Gamma(\alpha+\beta+2+k)}{\Gamma(\alpha+1+k)\Gamma(\beta+1+k)}\cdot\frac{2k+\alpha+\beta+1}{k+\alpha+\beta+1}\\ &\qquad\cdot I_{\alpha+\beta+2k+1}(z)P_k^{(\alpha,\beta)}(\cos 2\varphi)P_k^{(\alpha,\beta)}(\cos 2\psi).\end{aligned} \tag{3.13}$$

For $t<0$ the kernel on the left of (3.12) no longer remains positive, but if we write $2\sqrt{-t}=z$ then the modified Bessel functions I_ν reduce to ordinary Bessel functions J_ν, and (3.13) becomes

$$\begin{aligned}&\frac{1}{2}zJ_\alpha(z\sin\varphi\sin\psi)J_\beta(z\cos\varphi\cos\psi)\\ &=\sin^\alpha\varphi\sin^\alpha\psi\cos^\beta\varphi\cos^\beta\psi\\ &\quad\cdot\sum_{k=0}^{\infty}\frac{k!\Gamma(\alpha+\beta+2+k)}{\Gamma(\alpha+1+k)\Gamma(\beta+1+k)}\cdot\frac{2k+\alpha+\beta+1}{k+\alpha+\beta+1}(-1)^k\\ &\qquad\cdot J_{\alpha+\beta+2k+1}(z)P_k^{(\alpha,\beta)}(\cos 2\varphi)P_k^{(\alpha,\beta)}(\cos 2\psi).\end{aligned} \tag{3.14}$$

This formula was discovered by Bateman in 1904 (Watson [24, p. 370]).

In the rest of this subsection we shall discuss a few more special cases of (3.10).

Formula (3.10) *for* $\alpha_1 = \frac{1}{2}(\alpha+\beta+2)$. An interesting formula is obtained from (3.10) if we set

$$\alpha_1 = \frac{\alpha+\beta+2}{2}.$$

By virtue of (3.11) we have

$$\begin{aligned} &{}_1F_1\left(\frac{\alpha+\beta+2+2k}{2};\ \alpha+\beta+2+2k;\ \rho\right) \\ &\quad = \Gamma\left(\frac{\alpha+\beta+3}{2}+k\right)\left(\frac{\rho}{4}\right)^{-((\alpha+\beta+1)/2+k)} e^{\rho/2}\cdot I_{(\alpha+\beta+1)/2+k}\left(\frac{\rho}{2}\right). \end{aligned}$$

Hence (3.10) becomes

$$\begin{aligned} &e^{-z}(z/2)^{(\alpha+\beta+1)/2}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{((\alpha+\beta+2)/2)_{m+n}}{(\alpha+1)_m(\beta+1)_n m!\,n!} \\ &\qquad\cdot\left\{\frac{z(1-x)(1-y)}{2}\right\}^m\left\{\frac{z(1+x)(1+y)}{2}\right\}^n \\ &\qquad = \Gamma\left(\frac{\alpha+\beta+1}{2}\right)\sum_{k=0}^{\infty}\frac{k!(\alpha+\beta+1)_k(k+(\alpha+\beta+1)/2)}{(\alpha+1)_k(\beta+1)_k} \\ &\qquad\quad\cdot I_{(\alpha+\beta+1)/2+k}(z)P_k^{(\alpha,\beta)}(x)P_k^{(\alpha,\beta)}(y). \end{aligned} \tag{3.15}$$

Setting $y=-1$, multiplying by $(1-x)^\alpha(1+x)^\beta P_k^{(\alpha,\beta)}(x)$ and integrating over x from -1 to 1 we get

$$\begin{aligned} &\Gamma\left(\frac{\alpha+\beta+1}{2}\right)\frac{(-1)^k(\alpha+\beta+1)_k(k+(\alpha+\beta+1)/2)}{(\alpha+1)_k}h_k^{(\alpha,\beta)}I_{(\alpha+\beta+1)/2+k}(z) \\ &\quad = (z/2)^{(\alpha+\beta+1)/2}\int_{-1}^{1}dx\,e^{-zx}(1-x)^\alpha(1+x)^\beta P_k^{(\alpha,\beta)}(x) \\ &\quad\quad\cdot {}_1F_1\left(\frac{\alpha-\beta}{2};\ \alpha+1;\ z(x-1)\right) \end{aligned}$$

where

$$h_k^{(\alpha,\beta)} = \frac{2^{\alpha+\beta+1}\Gamma(k+\alpha+1)\Gamma(k+\beta+1)}{k!\,\Gamma(k+\alpha+\beta+1)(2k+\alpha+\beta+1)}.$$

Simplifying and then changing z to $-iz$ we obtain the expression for an ordinary Bessel function in terms of an integral over a Jacobi polynomial:

$$\pi^{1/2}\frac{\Gamma(\alpha+1)}{\Gamma((\alpha+\beta+2)/2)}\frac{\Gamma(k+\beta+1)}{k!}(2/z)^{(\alpha+\beta+1)/2}i^k J_{(\alpha+\beta+1)/2+k}(z)$$

$$= \int_0^\pi d\theta\, e^{iz\cos\theta}\sin\theta(1-\cos\theta)^\alpha(1+\cos\theta)^\beta P_k^{(\alpha,\beta)}(\cos\theta) \tag{3.16}$$

$$\cdot\, {}_1F_1\left(\frac{\alpha-\beta}{2};\alpha+1;2iz\sin^2\theta/2\right).$$

Case of Gegenbauer polynomials. Equation (3.16) reduces to the well-known Gegenbauer generalization of Poisson's integral when $\alpha=\beta=\nu-\frac{1}{2}$ (see, for example, [9, p. 57]). In this special case the left-hand side of (3.15) reduces to

$$e^{zxy}\,{}_0F_1\left(-;\nu+\frac{1}{2};\frac{z^2}{4}(1-x^2)(1-y^2)\right).$$

Using the Gegenbauer polynomials $C_k^\nu(x)$ defined by

$$C_k^\nu(x)=\frac{(2\nu)_k}{(\nu+\frac{1}{2})_k}P_k^{(\nu-1/2,\,\nu-1/2)}(x) \tag{3.17}$$

and the relation (3.11), we obtain the bilinear sum

$$\frac{\Gamma(\nu+\frac{1}{2})}{\Gamma(\nu)}\left(\frac{z}{2}\right)^{1/2}(\sin\theta\sin\varphi)^{(1/2)(1-2\nu)}I_{\nu-1/2}(z\sin\theta\sin\varphi)\,e^{z\cos\theta\cos\varphi}$$

$$= \sum_{k=0}^{\infty}\frac{(k+\nu)k!}{(2\nu)_k}I_{\nu+k}(z)C_k^\nu(\cos\theta)C_k^\nu(\cos\varphi), \tag{3.18}$$

where we have written $\cos\theta$ and $\cos\varphi$ for x and y respectively. This result can be extended to imaginary z-values to obtain the corresponding sum for the ordinary Bessel functions:

$$\frac{\Gamma(\nu+\frac{1}{2})}{\Gamma(\nu)}\left(\frac{z}{2}\right)^{1/2}(\sin\theta\sin\varphi)^{(1/2)(1-2\nu)}J_{\nu-1/2}(z\sin\theta\sin\varphi)\,e^{iz\cos\theta\cos\varphi}$$

$$= \sum_{k=0}^{\infty}\frac{(k+\nu)k!}{(2\nu)_k}(-1)^{k/2}J_{\nu+k}(z)C_k^\nu(\cos\theta)C_k^\nu(\cos\varphi). \tag{3.19}$$

This is, of course, a well-known formula, due to Gegenbauer ([24, p. 370]).

Equation (3.15) *for arbitrary* α, β. For arbitrary values of α, β with $\operatorname{Re}\alpha$, $\beta>-1$ it seems the best one can do is to express the double series on the left of (3.15) as an integral over a product of Bessel functions.

First note that

$$\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{(\alpha_1)_{m+n}}{(\alpha+1)_m(\beta+1)_n m!n!}u^m v^n$$

$$=\sum_{m=0}^{\infty}\frac{(\alpha_1)_m}{(\beta+1)_m m!}(v-u)^m\,{}_2F_1\left(-m,\,m+\alpha+\beta+1;\,\alpha+1;\,\frac{u}{u-v}\right)$$

$$=\sum_{m=0}^{\infty}\frac{(\alpha_1)_m}{(\alpha+1)_m(\beta+1)_m}(v-u)^m P_m^{(\alpha,\beta)}\left(\frac{v+u}{v-u}\right).$$

Then, using Bateman's generating function ([20, p. 256])

$$\sum_{n=0}^{\infty}\frac{P_n^{(\alpha,\beta)}(x)t^n}{(\alpha+1)_n(\beta+1)_n}={}_0F_1\left(-;\,\alpha+1;\,\frac{t(x-1)}{2}\right){}_0F_1\left(-;\,\beta+1;\,\frac{t(x+1)}{2}\right)\tag{3.20}$$

and the integral representation

$$\Gamma(\alpha_1+m)=2\int_0^{\infty}e^{-\xi^2}\xi^{2(\alpha_1+m)-1}\,d\xi,\qquad \alpha_1+m>0,$$

we get

$$\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{\left(\frac{\alpha+\beta+2}{2}\right)_{m+n}}{(\alpha+1)_m(\beta+1)_n m!n!}\left\{\frac{z(1-x)(1-y)}{2}\right\}^m\left\{\frac{z(1+x)(1+y)}{2}\right\}^n$$

$$=\frac{2}{\Gamma((\alpha+\beta+2)/2)}\int_0^{\infty}e^{-\xi^2}\xi^{\alpha+\beta+1}\,{}_0F_1\left(-;\,\alpha+1;\,\frac{z(1-x)(1-y)}{2}\xi^2\right)\tag{3.21}$$

$$\cdot\,{}_0F_1\left(-;\,\beta+1;\,\frac{z(1+x)(1+y)}{2}\xi^2\right)d\xi.$$

Incidentally, (3.20) also follows from (3.5) if we set $p=0$, $q=0$, $\rho=2t/y$ and let $y\to\infty$. In the process we use the limit $\lim_{y\to\infty}y^{-k}P_k^{(\alpha,\beta)}(y)=2^{-k}(\alpha+\beta+1)_{2k}/(\alpha+\beta+1)_k k!$ [22, p. 63, (4.21.6)].

Assuming that $\operatorname{Re} z<0$, then replacing z by $-z$ and expressing the ${}_0F_1$ functions in terms of Bessel functions, and simplifying, we finally obtain

$$\sum_{k=0}^{\infty}\frac{k!(\alpha+\beta+1)_k(k+(\alpha+\beta+1)/2)}{(\alpha+1)_k(\beta+1)_k}$$

$$\cdot(-1)^k I_{(\alpha+\beta+1)/2+k}(z)P_k^{(\alpha,\beta)}(\cos\theta)P_k^{(\alpha,\beta)}(\cos\varphi)$$

$$=\frac{\Gamma(\alpha+1)\Gamma(\beta+1)}{\Gamma((\alpha+\beta+1)/2)\Gamma((\alpha+\beta+2)/2)}\tag{3.22}$$

$$\cdot e^z(8z)^{-1/2}(2\sin(\theta/2)\sin(\varphi/2))^{-\alpha}(2\cos(\theta/2)\cos(\varphi/2))^{-\beta}$$

$$\cdot\int_0^{\infty}e^{-\xi^2/(8z)}\xi J_\alpha(\xi\sin(\theta/2)\sin(\varphi/2))J_\beta(\xi\cos(\theta/2)\cos(\varphi/2))\,d\xi.$$

(iii) $p=3$, $q=1$. Let n be a fixed nonnegative integer. Then set $\alpha_1=-n$, $\alpha_2=n+a+b+1$, $\alpha_3=\alpha+1$, $\beta_1=a+1$. Equation (3.5) reduces to

$$F\left[\begin{matrix}-n, n+a+b+1, \alpha+1; —; — \\ a+1; \alpha+1; \beta+1\end{matrix}; \frac{\rho(1-x)(1-y)}{4}, \frac{\rho(1+x)(1+y)}{4}\right]$$
$$= \sum_{k=0}^{\infty} \frac{2k+\alpha+\beta+1}{k+\alpha+\beta+1} \cdot \frac{(\alpha+\beta+2)_k k!}{(\beta+1)_k(\alpha+\beta+2)_{2k}} \cdot \frac{(-n)_k(n+a+b+1)_k}{(a+1)_k} \rho^k$$
$$\cdot {}_3F_2\left(\begin{matrix}-n+k, n+a+b+k+1, \alpha+k+1 \\ a+k+1, \alpha+\beta+2+2k\end{matrix}; \rho\right) P_k^{(\alpha,\beta)}(x) P_k^{(\alpha,\beta)}(y).$$

The left-hand side can be shown to reduce to

$$\sum_{l=0}^{n} \frac{(-n)_l(n+a+b+1)_l}{(a+1)_l l!} \left\{\frac{\rho(x+y)}{2}\right\}^l P_l^{(\alpha,\beta)}\left(\frac{1+xy}{x+y}\right) \Big/ P_l^{(\beta,\alpha)}(1).$$

Thus we obtain the formula

$$\frac{(a+1)_n}{n!} \sum_{l=0}^{n} \frac{(-n)_l(n+a+b+1)_l}{(a+1)_l l!} \left\{\frac{\rho(x+y)}{2}\right\}^l P_l^{(\alpha,\beta)}\left(\frac{1+xy}{x+y}\right) \Big/ P_l^{(\beta,\alpha)}(1)$$
$$= \sum_{k=0}^{n} g(n,k;\rho) \rho^k P_n^{(\alpha,\beta)}(x) P_k^{(\alpha,\beta)}(y) / P_k^{(\alpha,\beta)}(-1) \tag{3.23}$$

where

$$g(n,k;\rho) = \frac{(a+1)_n(n+a+b+1)_k(\alpha+\beta+1)_k}{(a+1)_k(\alpha+\beta+1)_{2k}(n-k)!}$$
$$\cdot {}_3F_2\left[\begin{matrix}-n+k, n+k+a+b+1, \alpha+k+1 \\ a+k+1, \alpha+\beta+2+2k\end{matrix}; \rho\right]. \tag{3.24}$$

If we set $\rho=1$, $y=-1$ in (3.23) we obtain Feldheim's projection relation

$$P_n^{(\alpha,\beta)}(x) = \sum_{k=0}^{n} g(n,k;1) P_k^{(\alpha,\beta)}(x). \tag{3.25}$$

4. An application. Let us set $\alpha_1=1$, $\alpha_2=\beta+1$ and $\rho=2/(1+z)$ in (3.6). Then since

$${}_2F_1\left(n+1, n+\beta+1; 2n+\alpha+\beta+2; \frac{2}{1+z}\right)$$
$$= 2^{-n-\alpha-\beta} \frac{\Gamma(2n+\alpha+\beta+1)}{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)} (z-1)^{\alpha}(z+1)^{n+\beta+1} Q_n^{(\alpha,\beta)}(z), \tag{4.1}$$

where $Q_n^{(\alpha,\beta)}(z)$ is the Jacobi function of the second kind (see, for example, Szegö

[22]) we get, on simplifying,

$$F_4\left(1, \beta+1; \alpha+1, \beta+1; \frac{(1-x)(1-y)}{2(1+z)}, \frac{(1+x)(1+y)}{2(1+z)}\right)$$

$$= 2(z-1)^{\alpha}(z+1)^{\beta+1} \sum_{k=0}^{\infty} N_k^{(\alpha,\beta)} R_k^{(\alpha,\beta)}(x) R_k^{(\alpha,\beta)}(y) Q_k^{(\alpha,\beta)}(z)/P_k^{(\alpha,\beta)}(1) \tag{4.2}$$

with

$$R_k^{(\alpha,\beta)}(x) = P_k^{(\alpha,\beta)}(x)/P_k^{(\alpha,\beta)}(1), \tag{4.3}$$

and

$$N_k^{(\alpha,\beta)} = \left[\int_{-1}^{1} (1-x)^{\alpha}(1+x)^{\beta} R_k^{(\alpha,\beta)}(x) R_k^{(\alpha,\beta)}(x)\, dx\right]^{-1}$$

$$= \frac{2^{-\alpha-\beta-1}\Gamma(\alpha+k+1)\Gamma(\alpha+\beta+k+1)}{\Gamma^2(\alpha+1)\Gamma(\beta+k+1)k!}(2k+\alpha+\beta+1). \tag{4.4}$$

Note the slight difference with $N_k^{(\alpha,\beta)}$ defined in (2.9).

Let us keep x, y in (4.2) fixed with $|x|<1$, $|y|<1$ and consider both sides as functions of z. In the z-plane cut from $-\infty$ to 1 the Jacobi functions of the first and second kind are connected by the following relation:

$$P_n^{(\alpha,\beta)}(z) = \frac{i}{\pi} \lim_{\varepsilon \to 0+} [e^{i\pi\alpha} Q_n^{(\alpha,\beta)}(z+i\varepsilon) - e^{-i\pi\alpha} Q_n^{(\alpha,\beta)}(z-i\varepsilon)], \qquad -1<z<1. \tag{4.5}$$

This property was used by Durand [6], [7] in deriving Nicholson-type integrals for Gegenbauer and Jacobi functions. Using (4.5) in (4.2) we obtain

$$K(x, y, z) \equiv \sum_{k=0}^{\infty} N_k^{(\alpha,\beta)} R_k^{(\alpha,\beta)}(x) R_k^{(\alpha,\beta)}(y) R_k^{(\alpha,\beta)}(z)$$

$$= (1-z)^{-\alpha}(1+z)^{-\beta-1} \cdot \frac{i}{2\pi} \lim_{\varepsilon \to 0+} [A(x, y, z+i\varepsilon) - A(x, y, z-i\varepsilon)], \tag{4.6}$$

where we are denoting $A(x, y, z)$ for the F_4 function on the left of (4.2), for abbreviation.

The Appell function in (4.2) is absolutely convergent if

$$\left|\frac{(1-x)(1-y)}{2(1+z)}\right|^{1/2} + \left|\frac{(1+x)(1+y)}{2(1+z)}\right|^{1/2} < 1. \tag{4.7}$$

With

$$x = \cos 2\varphi, \quad y = \cos 2\psi, \quad z = \cos 2\theta, \quad 0<\varphi, \psi, \theta < \pi/2,$$
$$a = \sin\varphi \sin\psi, \quad b = \cos\varphi \cos\psi, \quad c = \cos\theta, \quad 0<a, b, c<1, \tag{4.8}$$

this condition means

$$a+b<c. \tag{4.9}$$

But if this inequality is satisfied then $A(x, y, z)$ is single-valued on the real axis and so the r.h.s. of (4.6) vanishes.

The inequality (4.9) may be violated in either of the two ways:

$$\text{(4.10)}\qquad \begin{aligned} &\text{(i)}\quad 0<c<|b-a|,\\ &\text{(ii)}\quad |a-b|<c<a+b. \end{aligned}$$

Using a well-known transformation of an F_4 function into an F_1 function (see Bailey [4, p. 102, exercise 20(ii)]), we get

$$\text{(4.11)}\qquad \begin{aligned} &F_4(1,\beta+1;\alpha+1,\beta+1;a^2/c^2,b^2/c^2)\\ &\quad=(1-s)(1-t)F_1(1;\alpha-\beta,1-\alpha;\alpha+1;s,st) \end{aligned}$$

where

$$\text{(4.12)}\qquad -\frac{s}{(1-s)(1-t)}=a^2/c^2,\qquad -\frac{t}{(1-t)(1-s)}=b^2/c^2.$$

Solving for s and t we find that they can be real or complex but that $s/t=a^2/b^2$ must be real and positive. Let us first consider the case when they are both real. Equations (4.12) then imply

$$\text{(4.13)}\qquad \begin{aligned} &\text{either}\quad \text{(I)}\quad 0<t<1 \quad\text{and}\quad s>1,\\ &\text{or}\quad \text{(II)}\quad 0<s<1 \quad\text{and}\quad t>1,\\ &\text{or}\quad \text{(III)}\quad s<0,\qquad t<0. \end{aligned}$$

It can be easily shown that s, t are real provided

$$\text{(4.14)}\qquad \Delta\equiv(z-a_+)(z-a_-)\geqq 0$$

where $a_\pm=\cos 2(\varphi\pm\psi)$. In terms of a, b, c this becomes

$$\text{(4.15)}\qquad \Delta\equiv 4(c+a+b)(c+a-b)(c-a+b)(c-a-b).$$

Hence Δ is positive if $c>a+b$ or $c<|b-a|<a+b$. Since the case $c>a+b$ has already been dealt with before we need only consider the case $c<|b-a|$. Solving (4.12) we get

$$\text{(4.16)}\qquad s=\frac{a^2+b^2-c^2\pm\sqrt{\Delta}/2}{2b^2},\qquad t=\frac{a^2+b^2-c^2\pm\sqrt{\Delta}/2}{2a^2}.$$

As a function of z, $\sqrt{\Delta}$ has a branch point at a_- and another at a_+. In the z-plane cut along the real axis from min $(a_\pm)$ to max $(a_\pm)$, $\sqrt{\Delta}$ is single-valued and so for Re $z>a_\pm$ we may choose either of the signs in (4.16). As it turns out the choice of the $-$ sign is somewhat more convenient because it produces only one singularity in the integral representation of $F_1(1;\alpha-\beta,1-\alpha;\alpha+1;s,st)$ instead of two if we had chosen the $+$ sign. So we take

$$\text{(4.17)}\qquad s=\frac{a^2+b^2-c^2-\sqrt{\Delta}/2}{2b^2},\qquad t=\frac{a^2+b^2-c^2-\sqrt{\Delta}/2}{2a^2}.$$

If $c<|b-a|$, it can be easily seen that $s>0$, $t>0$ and

$$1-st=\frac{\sqrt{\Delta}}{2b^2}s>0,$$

but $s<1$, $t>1$ if $c<b-a$ and $s>1$, $t<1$ if $c<a-b$. Since for $s<1$, $st<1$ the Appell function on the r.h.s. of (4.11) is analytic we get

$$K(x, y, z)=0 \quad \text{if } 0<c<b-a. \tag{4.18}$$

For real $\sqrt{\Delta}$ we therefore need to consider only the case $c<a-b$. Let us consider the integral representation of F_1 [4]:

$$F_1(1;\alpha-\beta, 1-\alpha;\alpha+1; s, st)=\alpha\int_0^1 du(1-u)^{\alpha-1}(1-us)^{\beta-\alpha}(1-stu)^{\alpha-1}. \tag{4.19}$$

Note that the convergence of this integral requires $\operatorname{Re}\alpha>0$, but the factor α in front implies that the r.h.s. of (4.19) exists with the milder requirement $\operatorname{Re}\alpha>-1$.

We now split up the integral in (4.19) into two parts I_1 and I_2, say, where

$$\begin{aligned} I_1&=\alpha\int_0^{1/s} du(1-u)^{\alpha-1}(1-us)^{\beta-\alpha}(1-ust)^{\alpha-1},\\ I_2&=\alpha\int_{1/s}^1 du(1-u)^{\alpha-1}(1-us)^{\beta-\alpha}(1-ust)^{\alpha-1}. \end{aligned} \tag{4.20}$$

Obviously the I_1 part is single-valued and hence contributes nothing to $K(x, y, z)$. For the I_2 part we take $\arg(1-s)=-\pi$ in the upper half-plane and $+\pi$ in the lower half-plane. Simplifying, we obtain,

$$\begin{aligned} &\lim_{\varepsilon\to 0+}[A(x, y, z+i\varepsilon)-A(x, y, z-i\varepsilon)]\\ &=(I_1^+-I_2^-)(1-s)(1-t)\\ &=2i\alpha\sin\pi(\beta-\alpha)\\ &\quad\cdot(s-1)^{\beta+1}(1-t)^{\alpha}s^{-\alpha}\int_0^1 du\,u^{\beta-\alpha}(1-u)^{\alpha-1}\left\{1-\frac{t(s-1)}{1-t}u\right\}^{\alpha-1}. \end{aligned} \tag{4.21}$$

Convergence of the integral on the right requires $\operatorname{Re}(\beta-\alpha+1)>0$. It is also obvious that $\beta-\alpha$ cannot be zero or an integer for a nonzero contribution to $K(x, y, z)$. We thus get

$$\begin{aligned} K(x; y, z)&=2^{-\alpha-\beta-1}\frac{\sin\pi(\alpha-\beta)}{\pi}\cdot\frac{\Gamma(\beta-\alpha+1)\Gamma(\alpha+1)}{\Gamma(\beta+1)}\\ &\quad\cdot c^{-2\beta-2}(1-c^2)^{-\alpha}(s-1)^{\beta+1}\left(\frac{1-t}{s}\right)^{\alpha}\\ &\quad\cdot{}_2F_1\left(\beta-\alpha+1, 1-\alpha;\beta+1;\frac{t(s-1)}{1-t}\right). \end{aligned} \tag{4.22}$$

Using the formula for the quadratic tranformation of the ${}_2F_1$ function:

$$
\begin{aligned}
&{}_2F_1(\alpha_1, \alpha_2; \alpha_1-\alpha_2+1; v) \\
&\quad = (1+v)^{-\alpha_1}\, {}_2F_1\left(\frac{\alpha_1}{2}, \frac{\alpha_1+1}{2}; \alpha_1-\alpha_2+1; 4v(1+v)^{-2}\right), \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad |v|<1
\end{aligned} \tag{4.23}
$$

(see, for example, [8, p. 113, (34)]), we obtain

$$
\begin{aligned}
K(x, y, z) &= \frac{\sin \pi(\alpha-\beta)}{\pi} \frac{\Gamma(\beta-\alpha+1)\Gamma(\alpha+1)}{2^{\alpha+\beta+1}\Gamma(\beta+1)} \\
&\quad \cdot (1-c^2)^{-\alpha} c^{-2\beta-2} (s-1)^{\beta+1} \left(\frac{1-t}{s}\right)^{\alpha} (1+v)^{\alpha-\beta-1} \\
&\quad \cdot {}_2F_1\left(\frac{\beta-\alpha+1}{2}; \frac{\beta-\alpha+2}{2}; \beta+1; 4v(1+v)^{-2}\right)
\end{aligned} \tag{4.24}
$$

where $v = t(s-1)/(1-t)$.

This can be simplified by setting

$$
B = \frac{b^2+c^2-a^2}{2bc} < -1 \quad \text{for } c < a-b. \tag{4.25}
$$

Hence

$$
\begin{aligned}
&s = 1-\frac{c}{b}(B+\sqrt{B^2-1}), \qquad t = b^2/a^2 - \frac{bc}{a^2}(B+\sqrt{B^2-1}), \\
&v = \frac{t(s-1)}{1-t} = \frac{b^2}{c^2}(s-1)^2 = (B+\sqrt{B^2-1})^2, \\
&\frac{4v}{(1+v)^2} = 1-\left(\frac{1-v}{1+v}\right)^2 = 1-\frac{B^2-1}{B^2} = B^{-2}.
\end{aligned}
$$

We thus get the integrated addition formula

$$
\begin{aligned}
K(x, y, z) &= \frac{\sin \pi(\alpha-\beta)}{\pi} \Gamma(\beta-\alpha+1) \cdot \frac{\Gamma(\alpha+1)}{2^{\alpha+\beta+1}\Gamma(\beta+1)} (1-c^2)^{-\alpha} \\
&\quad \cdot a^{-2\alpha} (a^2-b^2-c^2)^{\alpha-\beta-1}\, {}_2F_1\left(\frac{\beta-\alpha+1}{2}, \frac{\beta-\alpha+2}{2}; \beta+1; B^{-2}\right),
\end{aligned} \tag{4.26}
$$

for $c < a-b$ and $\operatorname{Re}(\beta-\alpha+1) > 0$.

However, this can be analytically continued to the region $\operatorname{Re}(\alpha-\beta) > 0$ by using the relation $\Gamma(z)\Gamma(1-z) = \pi/\sin \pi z$. Thus

$$
\begin{aligned}
K(x, y, z) &= \frac{\Gamma(\alpha+1)(1-c^2)^{-\alpha} a^{-2\alpha} (a^2-b^2-c^2)^{\alpha-\beta-1}}{\Gamma(\alpha-\beta)\Gamma(\beta+1) 2^{\alpha+\beta+1}} \\
&\quad \cdot {}_2F_1\left(\frac{\beta-\alpha+1}{2}, \frac{\beta-\alpha+2}{2}; \beta+1; B^{-2}\right)
\end{aligned} \tag{4.27}
$$

for $c < a-b$ and $\operatorname{Re}(\alpha-\beta) > 0$, (see Gasper [12]).

Finally we shall consider the case (iii) of (4.10), i.e., $|a-b|<c<a+b$. Since the sum of any two of the numbers a, b, c is greater than the third they may be interpreted as forming the sides of a triangle. In this case Δ is negative and $\min(a\pm)<z<\max(a\pm)$. The function $\sqrt{\Delta}$ is double-valued across the cut and we take the branch $i\sqrt{-\Delta}$ in the upper half-plane and $-i\sqrt{-\Delta}$ in the lower half-plane. Thus, in the upper half-plane

$$F_1(1;\alpha-\beta,1-\alpha;\alpha+1;s,st)=\alpha\int_0^1 du\,(1-u)^{\alpha-1}(1-us)^{\beta-\alpha}(1-stu)^{\alpha-1}$$

where, now

$$s=\frac{a^2+b^2-c^2+i\sqrt{-\Delta}/2}{2b^2},\qquad t=\frac{a^2+b^2-c^2+i\sqrt{-\Delta}/2}{2a^2}. \tag{4.28}$$

Introducing the angle variable χ:

$$\tan\chi=\frac{\sqrt{-\Delta}/2}{a^2+b^2-c^2},\quad\text{so that}\quad s=\frac{a}{b}e^{i\chi},\quad t=\frac{b}{a}e^{i\chi}, \tag{4.29}$$

we get, for the upper half-plane

$$\begin{aligned}A(x,y,z+i\varepsilon)&=(1-s)(1-t)F_1(1;\alpha-\beta,1-\alpha;\alpha+1;s,st)\\&=-\frac{c^2}{ab}\alpha\,e^{i\chi}\int_0^1 du\,(1-u)^{\alpha-1}\left(1-\frac{a}{b}e^{i\chi}u\right)^{\beta-\alpha}(1-e^{2i\chi}u)^{\alpha-1}\\&=-\frac{c^2}{ab}\alpha\int_0^{e^{i\chi}} dv\,[(1-e^{i\chi}v)(1-e^{-i\chi}v)]^{\alpha-1}\left(1-\frac{a}{b}v\right)^{\beta-\alpha}.\end{aligned}$$

The contribution from the lower half-plane is just the complex conjugate of this. Hence

$$\begin{aligned}K(x,y,z)&=-\frac{i}{2\pi}(1-z)^{-\alpha}(1+z)^{-\beta-1}\frac{c^2\alpha}{ab}\int_{e^{-i\chi}}^{e^{i\chi}} dv\,[(1-e^{i\chi}v)\\&\qquad\cdot(1-e^{-i\chi}v)]^{\alpha-1}\left(1-\frac{a}{b}v\right)^{\beta-\alpha}\\&=\frac{c^2}{2\pi ab}(1-z)^{-\alpha}(1+z)^{-\beta-1}(2\sin\chi)^{2\alpha-1}\left(1-\frac{a}{b}e^{-i\chi}\right)\frac{\Gamma(\alpha)\Gamma(\alpha+1)}{\Gamma(2\alpha)}\\&\qquad\cdot{}_2F_1\left(\alpha-\beta,\alpha;2\alpha;\frac{2ia\sin\chi}{b-a\,e^{-i\chi}}\right).\end{aligned} \tag{4.30}$$

By setting

$$B=\frac{b^2+c^2-a^2}{2bc}=\cos 2\tau,\qquad 0<B<1, \tag{4.31}$$

so that $s = 1-(c/b)\,e^{-2i\tau} = (a/b)\,e^{ix}$ and using the transformation formulas [8],

$$
\begin{aligned}
{}_2F_1(\alpha_1, \alpha_2; 2\alpha_2; z) &= [\tfrac{1}{2}+\tfrac{1}{2}(1-z)^{1/2}]^{-2\alpha_1} \\
&\quad \cdot {}_2F_1\left(\alpha_1, \alpha_1-\alpha_2+\tfrac{1}{2}; \alpha_2+\tfrac{1}{2}; \left[\frac{1-(1-z)^{1/2}}{1+(1-z)^{1/2}}\right]^2\right), \qquad (4.32)\\
{}_2F_1(\alpha_1, \alpha_2; \alpha_3; z) &= (1-z)^{-\alpha_1}\, {}_2F_1\left(\alpha_1, \alpha_3-\alpha_2; \alpha_3; \frac{z}{z-1}\right),
\end{aligned}
$$

we obtain, after some simplification, the final result

$$
\begin{aligned}
K(x, y, z) &= \frac{\Gamma(\alpha+1)a^{-2\alpha}(1-c^2)^{-\alpha}(bc)^{\alpha-\beta-1}}{\Gamma(\alpha+\frac{1}{2})\Gamma(\frac{1}{2})2^{\alpha+\beta+2}}(1-B^2)^{\alpha-1/2} \\
&\quad \cdot {}_2F_1(\alpha-\beta, \alpha+\beta; \alpha+\tfrac{1}{2}; \tfrac{1}{2}(1-B)), \qquad (4.33)\\
&\quad |a-b|<c<a+b \quad \text{and} \quad \operatorname{Re}\alpha > -\tfrac{1}{2}.
\end{aligned}
$$

To summarize the results of this section we have

$$
K(x, y, z) = \begin{cases} 0, & \text{if } a<|b-c|, \quad ((4.9) \text{ and } (4.18)) \\ (4.26) \text{ or } (4.27), & \text{if } c<a-b, \\ (4.33), & \text{if } |a-b|<c<a+b. \end{cases}
$$

These results are, of course, precisely the same as those of Gasper [11], [12].

REFERENCES

[1] W. Al-Salam and A. Verma, *Some orthogonality preserving operators*, Proc. Amer. Math. Soc., 23 (1969), pp. 136–139.

[2] R. Askey, *Summability of Jacobi series*, Trans. Amer. Math. Soc., 179 (1973), pp. 71–84.

[3] ———, *Orthogonal polynomials and special functions*, vol. 21, SIAM Series of Regional Conference Lectures, Society for Industrial and Applied Mathematics, Philadelphia, 1975.

[4] W. N. Bailey, *Generalized Hypergeometric Series*, Stechert-Hafner Service Agency, New York and London, 1964.

[5] R. D. Cooper, M. R. Hoare and H. Rahman, Stochastic processes and special functions: on the probabilistic origin of some positive kernels associated with classical orthogonal polynomials, J. Math. Anal. Appl., to appear.

[6] L. Durand, *Nicholson-type integrals for products of Gegenbauer functions and related topics*, Theory and Application of Special Functions, R. Askey, ed., Academic Press, New York, 1975, pp. 353–374.

[7] ———, *Nicholson-type integrals for products of Jacobi functions I. Summary of results*, this Journal, to appear.

[8] A. Erdélyi, W. Magnus, F. Oberhettinger and F. G. Tricomi, eds., *Higher Transcendental Functions*, vol. 1, McGraw-Hill, New York, 1953.

[9] ———, *Higher Transcendental Functions*, vol. II, McGraw-Hill, New York, 1953.

[10] E. Feldheim, *Contributions à la theorie des polynomes de Jacobi*, Mat. Fiz. Lapok, 48 (1941), pp. 453–504. (In Hungarian, French summary.)

[11] G. Gasper, *Positivity and the convolution structure for Jacobi series*, Ann. of Math., 93 (1971), pp. 112–118.

[12] ———, *Banach algebras for Jacobi series and positivity of a kernel*, Ibid., 95 (1972), pp. 261–280.

[13] ———, *Nonnegativity of a discrete Poisson kernel for the Hahn polynomials*, J. Math. Anal. Appl., 42 (1973), pp. 438–451.
[14] M. E. H. ISMAIL, *Connection relations and bilinear formulas for the classical orthogonal polynomials*, Ibid., 57 (1977), pp. 487–496.
[15] ———, private communication.
[16] P. M. MORSE AND H. FESHBACH, *Method of Theoretical Physics*, vol. I, McGraw-Hill, New York, 1953.
[17] M. RAHMAN, *Construction of a family of positive kernels from Jacobi polynomials*, this Journal, 7 (1976), pp. 92–116.
[18] ———, *A five-parameter family of positive kernels from Jacobi polynomials*, this Journal, 7 (1976), pp. 386–413.
[19] ———, *Some positive kernels and bilinear sums for Hahn polynomials*, this Journal, 7 (1976), pp. 414–435.
[20] R. D. RAINVILLE, *Special Functions*, Macmillan, New York, 1960.
[21] L. J. SLATER, *Generalized Hypergeometric Functions*, Cambridge University Press, Cambridge, England, 1966.
[22] G. SZEGÖ, *Orthogonal polynomials*, Colloquium Publications, vol. 13, American Mathematical Society, Providence, RI, 1959.
[23] F. G. TRICOMI, *Integral equations*, Interscience, New York, 1957.
[24] G. N. WATSON, *A treatise on the theory of Bessel functions*, 2nd ed., Cambridge University Press, Cambridge, England, 1962.

SIAM J. MATH. ANAL.
Vol. 8, No. 6, November 1977

COMPARISON THEOREMS FOR SECOND ORDER RICCATI EQUATIONS WITH APPLICATIONS*

L. ERBE†

Abstract. The purpose of this paper is to obtain comparison theorems for the second order Riccati equation $r''+3rr'+r^3+p(t)=0$, where p is nonnegative and continuous on $[a, b)$, $0<a<b\leqq+\infty$. This yields several new comparison theorems and disconjugacy criteria for the third order linear equation $y'''+p(t)y=0$.

1. Introduction.

Consider the pair of third order linear equations

$$y'''+q_1(t)y=0, \tag{1}$$

$$z'''+q_2(t)y=0, \tag{2}$$

and the associated second order Riccati equations

$$r''+3rr'+r^3+q_1(t)=0, \qquad r=\frac{y'}{y}, \tag{3}$$

$$w''+3ww'+w^3+q_2(t)=0, \qquad w=\frac{z'}{z}, \tag{4}$$

where q_1, q_2 are continuous on an interval I. It is known [2], [4] that (1) is disconjugate on the interval I (i.e., no nontrivial solution of (1) has more than two zeros on I, counted with multiplicity) iff there exist functions $\alpha, \beta \in C^2(I)$ with $\alpha(t)<\beta(t)$ on I and

$$\alpha''+f_1(t, \alpha, \alpha')\geqq 0\geqq\beta''+f_1(t, \beta, \beta') \quad \text{on } I, \tag{5}$$

where $f_1(t, r, r')=3rr'+r^3+q_1(t)$. This is, in turn, equivalent to the existence of solutions y_1, y_2 of (1) with $y_1>0$, $y_2>0$ and $W(y_1, y_2)=y_1y_2'-y_1'y_2\neq 0$ on I [2], [4]. In the case that q_1 does not change sign on I, then disconjugacy of (1) is equivalent to disconjugacy of the adjoint equation [3]

$$y'''-q_1(t)y=0. \tag{1*}$$

Moreover, in this case (1) (and (1*)) are disconjugate if there exists $\beta\in C^2(I)$, $\beta(t)>0$ on I, and

$$\beta''+f_1(t, \beta, \beta')\leqq 0, \qquad t\in I, \tag{6}$$

since one can choose $\alpha\equiv 0$ in (5).

In [3] it is shown that if (1) is disconjugate on I and if $0\leqq q_2(t)\leqq q_1(t)$ on I, then it follows that (2) is disconjugate on I. By comparison with the third order

* Received by the editors December 12, 1975, and in revised form June 6, 1976.

† Department of Mathematics, University of Alberta, Edmonton, Alberta, Canada T6G 2G1.
This research was supported by the National Research Council of Canada, Grant NRC-A-7673.

Euler equation in which $q_1(t) = kt^{-3}$, $k > 0$, we find that

$$\limsup_{t\to\infty} t^3 q_2(t) > \frac{2}{3\sqrt{3}} \Rightarrow (2) \text{ is oscillatory} \tag{7}$$

and

$$\limsup_{t\to\infty} t^3 q_2(t) < \frac{2}{3\sqrt{3}} \Rightarrow (2) \text{ is disconjugate} \tag{8}$$

on some interval $[t_0, +\infty)$.

We prove in § 2 comparison theorems which relate the existence of a positive solution of (3) with the existence of a positive solution of (2). This yields immediately new comparison theorems for equation (2) by appropriate choices for $q_1(t)$. In particular, we are able to conclude that (2) is disconjugate on $[a, +\infty)$, $a > 0$ provided $q_2(t) = k(1+\phi(t))t^{-3}$ where $0 < k \leqq 2/(3\sqrt{3})$ and $\phi(t)$ is a continuous ω-periodic function with $\int_a^{\omega+a} \phi(t)\,dt = 0$, $\int_a^t \phi(t)\,dt \leqq 0$, $t > a$, and $\phi(t) \geqq -1$. In § 3 below we give additional integral tests for disconjugacy to which no known tests apply. Similar results were recently obtained by Stafford and Heidel [6] for the second order linear equation via a first order Riccati equation and their methods have been extended to certain first order matrix Riccati equations in [5].

2. Comparison theorems.

THEOREM 2.1. *Let $q_1, q_2 \in C[a, b)$ (or $[a, b]$), $0 < a < b \leqq +\infty$ with $q_1(t) \geqq 0$, $q_2(t) \geqq 0$ and $\not\equiv 0$ on $[a, b)$. Assume (3) has a positive solution r on $[a, b)$ satisfying*

$$1 \leqq tr(t), \qquad t \geqq a, \tag{9}$$

and

$$ar(a) < 2 < 3ar(a) - \tfrac{3}{2}a^2(r(a))^2 - a^2 r'(a). \tag{10}$$

Assume also that

$$\int_a^t s^3 q_2(s)\,ds \leqq \int_a^t s^3 q_1(s)\,ds, \qquad t \geqq a. \tag{11}$$

Then (4) has a positive solution w on $[a, b)$ and (2) is disconjugate on $[a, b)$.

Proof. In (3) and (4) we make the substitutions

$$u(t) = tr(t), \tag{12}$$

$$v(t) = tw(t) \tag{13}$$

to obtain

$$t^2u'' - 2tu' + 3tuu' = 3u^2 - u^3 - 2u - t^3 q_1(t), \tag{14}$$

and

$$t^2v'' - 2v' + 3tvv' = 3v^2 - v^3 - 2v - t^3 q_2(t), \tag{15}$$

respectively. Note that $u(t) \geqq 1$ on $[a, b)$ by (9) and $1 \leqq u(a) < 2$ by (10). We show first that $1 \leqq u(t) < 2$ for all $t \geqq a$. If not, there exists a first point $t_1 > a$ such that $u(t_1) = 2$ and $u'(t_1) \geqq 0$. Integrating (14) by parts on $[a, t_1]$ yields, after some

rearranging,

$$H(t_1)=\int_a^{t_1} g(u(s))\,ds-\int_a^{t_1} s^3q_1(s)\,ds+H(a) \tag{16}$$

where

$$H(t)=t^2u'(t)+tu(t)(\tfrac{3}{2}u(t)-4) \tag{17}$$

and

$$g(u)=\tfrac{9}{2}u^2-u^3-6u. \tag{18}$$

Note first that $g(u)$ is strictly increasing in u for $1<u<2$ and hence for $a\leqq t<t_1$ we have $g(u(t))<g(u(t_1))=g(2)=-2$ so that from (16) we get

$$t_1^2u'(t_1)-2t_1<-2(t_1-a)-\int_a^{t_1} s^3q_1(s)\,ds+H(a) \tag{19}$$

which implies

$$0\leqq t_1^2u'(t_1)<2a+H(a)-\int_a^{t_1} s^3q_1(s)\,ds. \tag{20}$$

But $2a+H(a)=a(2+a^2r'(a)-3ar(a)+\frac{3}{2}a^2(r(a))^2)<0$ by (10) which contradicts (20) since $q_1\geqq 0$. Therefore, we conclude that $1\leqq u(t)<2$ on $[a, b)$.

Now let $v(t)$ be the solution to (15) with $v(a)=u(a)$ and $v'(a)>u'(a)$ and such that

$$2<v(a)(4-\tfrac{3}{2}v(a))-av'(a). \tag{21}$$

Since (21) holds (i.e., is equivalent to the right hand part of (10)) with u replacing v, it is clear that (21) may be satisfied for $v'(a)$ sufficiently close to $u'(a)$. Therefore, by the first part of the argument above, it follows that $v(t)<2$ as long as $v(t)\geqq 1$. We now claim that $v(t)>u(t)$ for $t>a$. It will then follow that $v(t)$ exists on all of $[a, b)$ and satisfies $u(t)<v(t)<2$ on (a, b). (Note that $v(t)$ is extendable as long as it is bounded; cf. e.g. [1, Cor. 1.4.1]). Suppose then there exists a first point $t_2>a$ such that $v(t_2)=u(t_2)$ and $v'(t_2)\leqq u'(t_2)$. Then integrating (14) and (15) by parts from a to t_2 and subtracting gives

$$\begin{aligned} t_2^2(v'(t_2)-u'(t_2)) &= \int_a^{t_2}[g(v(s))-g(u(s))]\,ds-\int_a^{t_2} s^3(q_2(s)-q_1(s))\,ds \\ &\quad +a^2(v'(a)-u'(a)). \end{aligned} \tag{22}$$

But the right hand side of (22) is positive by (11) and the fact that $v'(a)>u'(a)$ and $g(v(t))-g(u(t))>0$ on (a, t_2). This contradicts $v'(t_2)-u'(t_2)\leqq 0$ and so we conclude $u(t)<v(t)<2$ on (a, b). Therefore, $w(t)=v(t)/t$ is a positive solution of equation (4) existing on $[a, b)$ and this implies that equation (2) is disconjugate on $[a, b)$. This complétes the proof of the theorem.

More general changes of variable yield additional comparison results. The following result includes the previous theorem as a special case. For completeness, we include a brief proof, noting the differences.

THEOREM 2.2. *Let q_1, q_2 be as in Theorem* 2.1 *and let $\rho \in C^2[a, b)$ with $\rho(t)>0$, $\rho'(t)>0$, and $\rho''(t)\leqq 0$ on $[a, b)$. Assume* (3) *has a solution r on $[a, b)$ satisfying*

$$\rho'(t)\leqq \rho(t)r(t), \qquad t\geqq a, \tag{23}$$

$$\rho'(a)\leqq \rho(a)r(a)<2\rho'(a), \tag{24}$$

and

$$G(a)+3\rho(a)(\rho'(a))^2<\int_a^t q_1(s)(\rho(s))^3\,ds-\int_a^t(\rho'(s))^3\,ds+M(t), \qquad t\geqq a, \tag{25}$$

where

$$G(a)=(\rho(a))^3\left(\tfrac{3}{2}(r(a))^2+r'(a)\right)-3(\rho(a))^2\rho'(a)r(a)$$

and

$$M(t)=2\rho''(t)(\rho(t))^2+\rho(t)(\rho'(t))^2.$$

Assume further that

$$\int_a^t(\rho(s))^3q_2(s)\,ds\leqq\int_a^t(\rho(s))^3q_1(s)\,ds, \qquad t\geqq a. \tag{26}$$

Then (4) *has a solution on $[a, b)$ and equation* (2) *is disconjugate on $[a, b)$.*

Proof. Let $u(t)=\rho(t)r(t)$ and $v(t)=\rho(t)w(t)$ in (3) and (4) and we get

$$\rho^2u''-2\rho\rho'u'+3uu'\rho=-u^3+3u^2\rho'-2u\rho'^2+u\rho\rho''-q_1\rho^3, \tag{27}$$

and

$$\rho^2v''-2\rho\rho'v'+3vv'\rho=-v^3+3v^2\rho'-2v\rho'^2+v\rho\rho''-q_2\rho^3, \tag{28}$$

respectively. The proof is similar to the proof of Theorem 2.1: Integrating (27) by parts from a to $t>a$ and rearranging gives

$$G(t)=\int_a^t g_1(u(s))\,ds-\int_a^t 3u(s)\rho(s)\rho''(s)\,ds-\int_a^t q_1(s)(\rho(s))^3+G(a), \tag{29}$$

where

$$g_1(u)=\tfrac{9}{2}u^2\rho'-6u\rho'^2-u^3,$$

$$G(t)=\tfrac{3}{2}(u(t))^2\rho(t)-4\rho(t)\rho'(t)u(t)+(\rho(t))^2u'(t).$$

Note that for each fixed t, $g_1(u)$ is strictly increasing in u for $\rho'(t)<u<2\rho'(t)$. We may show that (23), (24), and (25) imply

$$\rho'(t)\leqq u(t)<2\rho'(t) \quad \text{on } [a, b). \tag{30}$$

For if not, then (30) holds on a subinterval $[a, t_1)$ for some $t_1>a$ and

$u(t_1)=2\rho'(t_1)$, $u'(t_1)\geqq 2\rho''(t_1)$. Then on $[a,t_1)$ we have $g_1(u(t))<g_1(2\rho'(t))=-2(\rho(t))^3$ and

$$-\int_a^{t_1} 3u\rho\rho''\,ds<-\int_a^{t_1} 6\rho\rho'\rho''\,ds=-3\rho(t_1)(\rho'(t_1))^2+3\int_a^{t_1}(\rho'(s))^3\,ds \tag{31}$$
$$+3\rho(a)(\rho'(a))^2.$$

Since $\int_a^{t_1} g_1(u(s))\,ds<-2\int_a^{t_1}(\rho'(s))^3\,ds$ we conclude from (29) that

$$G(t_1)<\int_a^{t_1}(\rho'(s))^3\,ds-\int_a^{t_1}(\rho(s))^3q_1(s)\,ds-3\rho(t_1)(\rho'(t_1))^2 \tag{32}$$
$$+3\rho(a)(\rho'(a))^2+G(a).$$

But $G(t_1)\geqq 2(\rho(t_1))^2\rho''(t_1)-2\rho(t_1)(\rho'(t_1))^2=M(t_1)-3\rho(t_1)(\rho'(t_1))^2$ so that (32) implies

$$M(t_1)<\int_a^{t_1}(\rho'(s))^3\,ds-\int_a^{t_1}q_1(s)(\rho(s))^3\,ds+G(a)+3\rho(a)(\rho'(a))^2, \tag{33}$$

contradicting (25). Therefore, we conclude that (30) holds on $[a,b)$. Thus, if $v(t)$ is the solution of (28) satisfying $v(a)=u(a)$, $v'(a)>u'(a)$ and such that (25) holds now with v replacing u, then we conclude that $v(t)<2\rho'(t)$ as long as $v(t)>\rho'(t)$.

In fact, it follows that $v(t)>u(t)$ on $[a,b)$ for if there exists $t_2>a$ with $v(t_2)=u(t_2)$, $v'(t_2)\leqq u'(t_2)$, then we get, as in Theorem 2.1,

$$\begin{aligned}(\rho(t_2))^2(v'(t_2)-u'(t_2))&=\int_a^{t_2}(g_1(v(s))-g_1(u(s))\,ds)\\&\quad-\int_a^{t_2}3\rho(s)\rho''(s)(v(s)-u(s))\,ds\\&\quad-\int_a^{t_2}(\rho(s))^3(q_2(s)-q_1(s))\,ds\\&\quad+(\rho(a))^2(v'(a)-u'(a))\end{aligned} \tag{34}$$

and again the left hand side of (34) is $\leqq 0$ and the right hand side is positive. This completes the proof.

3. Examples and applications.

Example 3.1. Let $q_1(t)=(2/(3\sqrt{3}))t^{-3}$. Then (3) has the solution $r(t)=\lambda/t$, where $\lambda=1+\sqrt{3}/3<2$. Hence, $u(t)=tr(t)=\lambda$ and condition (10) becomes $2<4\lambda-\frac{3}{2}\lambda^2$ which is clearly satisfied for $1<\lambda<2$. Hence, if $q_2(t)\geqq 0$ and

$$\frac{1}{t-a}\int_a^t s^3q_2(s)\,ds\leqq\frac{2}{3\sqrt{3}},\qquad t>a, \tag{35}$$

then (4) has a solution $w(t)$ with $\lambda<tw(t)<2$ on $(a,+\infty)$ and (2) is disconjugate on $[a,+\infty)$. As an example, $q_2(t)=kt^{-3}(1+\phi(t))$ satisfies (35) if $k\leqq 2/(3\sqrt{3})$, $\phi(t)$ is

ω-periodic, continuous, $\phi(t) \geqq -1$, and $\int_a^t \phi(s)\,ds \leqq 0$, $t > a$. Thus if $\phi(t) = -\sin(t-a)$, then (2) is disconjugate if $0 < k \leqq 2/(3\sqrt{3})$.

The above example can be generalized by an application of Theorem 2.2.

Example 3.2. Let $\rho(t) = t^\delta$, $0 < \delta < 1$, and $q_1(t) = (2/(3\sqrt{3}))t^3$ so that $r(t) = \lambda/t$, $\lambda = 1 + \frac{1}{3}\sqrt{3}$, as in Example 3.1. Condition (23) obviously holds and (24) will hold if $1 + \frac{1}{3}\sqrt{3} < 2\delta < 2$. Condition (25) becomes, after some calculation,

$$h(\delta) \equiv 8\delta^3 - 15\delta^2 + 12\delta - 4 - \frac{\sqrt{3}}{3}(3\delta - 2)^2 < 0, \qquad \left(\delta > \frac{2}{3}\right), \tag{36}$$

and

$$\psi(\delta) \equiv 4\delta^3 - 6\delta^2 + 2\delta + \frac{1}{3\sqrt{3}} > 0. \tag{37}$$

It is not difficult to verify that both (36) and (37) hold for $1 + \frac{1}{3}\sqrt{3} < 2\delta < 2$, and hence, for δ in this range, it follows that (2) will be disconjugate on $[a, +\infty)$ provided

$$\int_a^t s^{3\delta} q_2(s)\,ds \leqq \frac{2}{3\sqrt{3}(3\delta - 2)}\left(t^{3\delta-2} - a^{3\delta-2}\right). \tag{38}$$

For $\delta = 1$, this is Example 3.1 above.

Remark. We wish to conclude by noting that the above theorems are sharp since they include, as a special case, the Euler equation. The examples above may not be handled by any comparison or disconjugacy theorems for third order linear equations known to the author.

REFERENCES

[1] S. BERNFELD AND V. LAKSHMIKANTHAM, *An Introduction to Nonlinear Boundary Value Problems*, Academic Press, New York, 1974.

[2] L. ERBE, *Disconjugacy conditions for the third order linear differential equation*, Canad. Math. Bull., 12 (1969), pp. 603–613.

[3] M. HANAN, *Oscillation criteria for third order linear differential equations*, Pacific J. Math, 11 (1961), pp. 919–944.

[4] P. HARTMAN, *Principal solutions of disconjugate n'th order linear differential equations*, Amer. J. Math., 91 (1969), pp. 306–362.

[5] R. A. JONES, *Comparison theorems for matrix Riccati equations*, SIAM J. Appl. Math., 29 (1975), pp. 77–90.

[6] R. A. STAFFORD AND J. W. HEIDEL, *A new comparison theorem for scalar Riccati equations*, Bull. Amer. Math. Soc., 80 (1974), pp. 754–757.

SIAM J. MATH. ANAL.
Vol. 8, No. 6, November 1977

BOUNDS ON THE NUMBER OF POLES OF THE SOLUTION OF A GENERAL RICCATI EQUATION*

A. RONVEAUX,† M-C. DUMONT-LEPAGE† AND J. HABAY‡

Abstract. An elementary technique is used to give upper and lower bounds to the number of poles of the solutions of a general Riccati equation and asymptotic properties are also derived. Different applications of these bounds in the theory of second order linear differential equations give new results on the number of zeros or extrema of the solutions, on the characterization of eigenvalues and on the stability intervals in the periodic cases.

Introduction. The relation between second order linear differential equations and the Riccati equation has been of course well known since Bernoulli [1]. However, for a Riccati equation with variable coefficients the exact location of the poles of the solutions is still an unsolved problem. Yet this problem is an important one because, as we will see in the second part of this work, rather different applications are shown to be equivalent to the problem of finding the poles of an appropriate Riccati equation.

This work consists of counting the number of poles, inside a finite interval, of a general Riccati equation written as a "phase equation" [2], [3] and of applying these results to 5 different situations.

In part I we give general upper and lower bounds for the number of poles inside a finite interval. Aside from explicit and various bounds, the most interesting theoretical result of this part states that using relatively weak hypotheses, the number of poles inside a given interval increases asymptotically as the square root of a fundamental parameter. In Part II we apply our bounds to the following situations:

§ 5: Number of extrema of $u(x)$: $u''(x)+Q(x)u(x)=0$.

§ 6: Number of eigenvalues smaller than a given value.

§ 7: Number of zeros of some special functions.

§ 8: Stability interval for the periodic equation $u''(x)+\lambda Q(x)u(x)=0$.

§ 9: Stability interval for the periodic equation $u''(x)+(\lambda+Q(x))u(x)=0$.

The method of majoration is quite elementary and generalizes a technique used by Calogero [4] in an interesting problem of quantum mechanics: counting the number of bound states of a given potential for the Schrödinger equation.

The so-called "Riccati phase equation" [15] used throughout this work is more appropriate than the usual Riccati equation satisfied by the logarithmic derivative for the two following reasons.

1. The "phase," which is an appropriate homographical transformation of the logarithmic derivative, is systematically used to relate the location of the poles of the solution to the particular quantity of interest (eigenvalue, limit of the stability zone, etc.)

* Received by the editors May 30, 1975, and in revised form February 26, 1976.

† Department de Physique, Facultés Universitaires Notre-Dame de la Paix, B-5000 Namur, Belgium.

‡ Department de Mathématiques, Facultés Universitaires Notre-Dame de la Paix, B-5000 Namur, Belgium.

2. The Riccati equation obtained is in the "factorized" form of a perfect square which is essential for the technical developments.

The conditions obtained in the present work for the existence of n poles generalize in some way conditions given before for the existence of one pole [3], [6].

We also want to point out that the "phases" used in this work do not have the same meaning as in the work of Borüvka [7], where the phases are defined by a Prüfer substitution.

PART I. UPPER AND LOWER BOUNDS FOR THE NUMBER OF POLES

1. General content. Let us consider the second order differential equation

$$(p(x)u'(x))'+r(x)u(x)=V(x)u(x) \tag{1}$$

with the following assumptions:

$$\begin{aligned} &p(x)>0 \quad \text{on }]a,b[, \\ &V(x)\leqq 0 \quad \text{on }]a,b[, \qquad 0\leqq a<b, \\ &V(x), r(x), p(x) \text{ and } p'(x) \text{ are piecewise} \\ &\qquad\qquad \text{continuous in }]a,b[. \end{aligned} \tag{2}$$

Let $u_i(x)$, $i=1,2$, be two linearly independent solutions of (1) when $V(x)\equiv 0$, such that the Wronskian of $u_1(x)$ and $u_2(x)$ is negative:

$$W(x)=u_1(x)u_2'(x)-u_2(x)u_1'(x)=\frac{K}{p(x)}<0. \tag{3}$$

The solution of the second order differential equation (1) can be written as follows, using the Lagrange method of variation of constants:

$$u(x)=C_1(x)u_1(x)+C_2(x)u_2(x). \tag{4}$$

The so-called "phase" $S(x)$ defined as

$$S(x)\equiv\frac{C_2(x)}{C_1(x)} \tag{5}$$

satisfies the Riccati equation [3]

$$S'(x)=\frac{dS(x)}{dx}=\frac{V(x)}{p(x)W(x)}(u_1(x)+S(x)u_2(x))^2 \tag{6}$$

Let us also assume that the initial or boundary condition on $u(x)$, connected to the phase by the relation

$$\frac{u(x)}{u'(x)}=\frac{u_1(x)+S(x)u_2(x)}{u_1'(x)+S(x)u_2'(x)}, \tag{7}$$

is such that $S(a)=0$.

We may now give upper and lower bounds on the number of poles of the phase $S(x)$.

2. Upper bounds on the number of poles of the phase.

2.1. First type of upper bounds. Following Calogero [4], let us now introduce the new function $g(x)$ through

$$S(x)=\rho(\tan g(x)-\alpha) \tag{8}$$

where ρ and α are positive parameters which we shall specify later. Because $S(a)=0$, $g(a)$ equals arctan α, and let us assume

$$0<g(a)<\pi/2 \tag{9}$$

in order to define the function $g(x)$ uniquely. Equations (6) and (8) imply that

$$g'(x)=\frac{V(x)u_2^2(x)}{\rho p(x)W(x)}[f(x)\cos g(x)+\rho \sin g(x)-\rho\alpha \cos g(x)]^2, \tag{10}$$

with

$$f(x)\equiv\frac{u_1(x)}{u_2(x)}. \tag{11}$$

Thus, $g(x)$ is an increasing positive function. Finally, $g(x)$ satisfies the following inequality:

$$g'(x)\leqq\frac{V(x)u_2^2(x)}{\rho p(x)W(x)}[(f(x)-\rho\alpha)^2+\rho^2]. \tag{12}$$

This new function $g(x)$ allows us to characterize the number of poles of the phase function $S(x)$. Indeed, by (8), a pole of the phase $S(x)$ is given by any solution of

$$g(x)=\pi/2+k\pi. \tag{13}$$

Therefore, the first pole of $S(x)$ occurs at the point x_1 defined as follows:

$$g(x_1)=\pi/2, \qquad x_1>a, \tag{14}$$

and there are at most

$$[g(b)-\pi/2]/\pi \tag{15}$$

poles between x_1 and b. (The number of poles in $[x_1, b]$ is exactly equal to the integral part of the quantity (15)). Thus, it is clear that the total number of poles belonging to the interval $[a, b]$ satisfies the following inequality:

$$n\leqq\frac{1}{2}+\frac{g(b)}{\pi}. \tag{16}$$

With relation (12) we may write:

$$n\leqq\frac{1}{2}+\frac{1}{\pi}g(a)+\frac{1}{\pi\rho}(I_2-2\rho\alpha I_1+\rho^2(1+\alpha^2)I_0) \tag{17}$$

where

$$I_i = \int_a^b \frac{V(x)}{p(x)W(x)} u_2^2(x) f^i(x)\, dx, \qquad i = 0, 1, 2. \tag{18}$$

(We shall verify the existence of these integrals in every particular case).

At this point the choice of ρ and α is still arbitrary. The best choice is the one that minimizes the right-hand side of (17). We shall proceed in two stages. We first choose the ρ value which gives a null derivative of the right-hand side of (17). Then we have the first general upper bound on n:

$$U: \qquad n \leqq \frac{1}{2} + \frac{1}{\pi} \arctan \alpha + \frac{2}{\pi}(\sqrt{I_0 I_2 (1+\alpha^2)} - \alpha I_1). \tag{19}$$

The most efficient choice for α would be the value which minimizes the right-hand side of (19). But that would involve solving a fourth degree equation in α, and we prefer to choose first

$$\alpha = 0. \tag{20}$$

This gives a particular upper bound U_1 on n:

$$U_1: \qquad n \leqq \frac{1}{2} + \frac{2}{\pi}\sqrt{I_0 I_2}. \tag{21}$$

Another choice of α is the one which minimizes the expression $\sqrt{I_0 I_2 (1+\alpha^2)} - \alpha I_1$ (neglecting $(1/\pi)$ arctan α), i.e.:

$$\alpha = \frac{I_1}{(I_0 I_2 - I_1^2)^{1/2}}. \tag{22}$$

This choice is possible if and only if I_1 and $I_0 I_2 - I_1^2$ are positive functions.

The condition $I_1 > 0$ is true if

$$f(x) \geqq 0, \tag{23}$$

and from the Schwarz inequality, we deduce that $I_0 I_2$ is never smaller than I_1^2.

Thus from (9), we may write:

$$U_2: \qquad n \leqq 1 + \frac{2}{\pi}(I_0 I_2 - I_1^2)^{1/2}, \qquad I_1 > 0, \tag{24}$$

which is a second form of the required upper bounds on n.

The bounds U_1 and U_2 are easy to compute and in practice the smallest one should be used.

2.2. Second type of upper bounds [4]. When (1) verifies additional assumptions, we may find another upper bound on the number of poles of $S(x)$ which is usually better than the bound U.

Define

$$h(x) \equiv \frac{V(x)}{p(x)W(x)} u_2^2(x) \tag{25}$$

and let us introduce the following additional conditions on $[a, b]$:

(26)
(i) $f(x)$ finite;
(ii) $f'(x) = -W/u_2^2$ bounded;
(iii) $h(x)$ piecewise differentiable and $h'(x) \leqq 0$;
(iv) $[h(x)]^{1/2}$ integrable.

These conditions can easily be transferred to conditions on $u_1(x)$, $u_2(x)$ and $V(x)$.

Let us suppose now that the phase function $S(x)$ has n poles in $[a, b]$ denoted by $p_i (i = 1, \cdots, n)$.

We know that $S(x)$ is an increasing function satisfying a Riccati equation and such that $S(a) = 0$. Thus, $S(x)$ has at least n zeros in $[a, b]$ at the points,

$$x = z_i, \qquad i = 1, \cdots, n, \tag{27}$$

with

$$z_1 = a, \qquad z_i < p_i < z_{i+1}.$$

In order to give a limit on the integer n, let us introduce a new function $y(x)$ defined as follows:

$$y(x) = f(x) + S(x). \tag{28}$$

By the properties of $f(x)$ and $S(x)$ we see that $y(x)$ verifies a Riccati equation which has n poles on $[a, b]$ at the points

$$x = p_i, \qquad i = 1, \cdots, n, \tag{29}$$

and has at least $n-1$ zeros and at most $n+1$ zeros (because it is an increasing function).

The exact number of zeros depends on the values $y(a)$ and $y(b)$. Let us call s_i, $i = 1, \cdots, n_z$, the zeros of $y(x)$ belonging to $[a, p_n]$. Therefore we may write

(30)
$$s_i < p_i < s_{i+1} \quad \text{if } f(a) \leqq 0 \quad (n_z = n),$$
$$s_i < p_{i+1} < s_{i+1} \quad \text{if } f(a) > 0 \quad (n_z = n-1).$$

If we define the following integrals

$$J_i = \int_{s_i}^{p_j} (h(x))^{1/2}\, dx \tag{31}$$

where $j = i$ if $f(a) \leqq 0$ and $j = i+1$ if $f(a) > 0$,

$$J = \int_a^b (h(x))^{1/2}\, dx \tag{32}$$

it is obvious that

$$J \geqq \sum_{i=1}^{n_z} J_i. \tag{33}$$

Let us again make a change of dependent variable $y(x)$ by:

$$t(x) = h(x)^{1/2} y(x). \tag{34}$$

We immediately find the Riccati equation

$$\frac{t'(x)}{\sqrt{h(x)}} = \frac{h'(x)}{2h(x)\sqrt{h(x)}} t(x) + f'(x) + t^2(x). \tag{35}$$

But we know that $h'(x)/[2h(x)\sqrt{h(x)}]$ is negative and that $y(x)$ takes on positive values for any x in $[s_i, p_j]$ because $y(x)$ increases. This property being true for $t(x)$ we may write, using conditions (26)(ii),

$$\frac{t'(x)}{\sqrt{h(x)}} \leqq F + t^2(x) \quad \forall x : s_i \leqq x < p_j \tag{36}$$

where

$$0 \leqq f'(x) \leqq F. \tag{37}$$

Integrating (36) between s_i and p_j, we find

$$\left[\frac{1}{\sqrt{F}} \arctan g \frac{t(x)}{\sqrt{F}}\right]_{t(s_i)}^{t(p_j)} \leqq J_i, \tag{38}$$

so that

$$J_i \geqq \frac{1}{\sqrt{F}} \frac{\pi}{2}. \tag{39}$$

Here we must distinguish two cases.

1. Let us first suppose that $f(a) \leqq 0$. Then (33) becomes

$$J \geqq \frac{n}{\sqrt{F}} \frac{\pi}{2}. \tag{40}$$

Thus we have the first required upper bound of the second type:

$$U_1': \qquad n \leqq \sqrt{F} \frac{2}{\pi} \int_a^b \left(\frac{V(x)u_2^2(x)}{p(x)W(x)}\right)^{1/2} dx \qquad \text{if } \frac{u_1(a)}{u_2(a)} \leqq 0. \tag{41}$$

2. If $f(a) > 0$, equation (33) becomes

$$J \geqq \frac{n-1}{\sqrt{F}} \frac{\pi}{2} \tag{42}$$

which is equivalent to

$$U_2': \qquad n \leqq \sqrt{F} \frac{2}{\pi} \int_a^b \left(\frac{V(x)u_2^2(x)}{p(x)W(x)}\right)^{1/2} dx + 1 \qquad \text{if } \frac{u_1(a)}{u_2(a)} > 0. \tag{43}$$

2.3. Conclusion. The bounds U_1' and U_2' lead us to obtain an important result on the growth of n. In order to see this, let us give another form to $V(x)$, and let us write

$$V(x) = \nu v(x) \tag{44}$$

where $v(x)$ is a reference negative function giving the shape of $V(x)$, and ν is a positive constant representing the *strength* of $V(x)$.

Then the upper bound U_1', for instance, becomes

$$n \leqq \sqrt{\nu}\sqrt{F}\frac{2}{\pi}\int_a^b \left(\frac{v(x)u_2^2(x)}{p(x)W(x)}\right)^{1/2} dx. \tag{45}$$

We can interpret this as follows:

(46) *For a perturbing function $V(x)=\nu n(x)$ whose shape $v(x)$ is fixed, the number of poles of the phase function increases at most as the square root of the strength parameter ν.*

3. Lower bounds on the number of poles of $S(x)$. In this section we suppose that there is only one hypothesis in addition to those mentioned above, in (2), (3) and (7), that is,

$$f(x)<\infty, \qquad a \leqq x \leqq b. \tag{47}$$

Later we shall see that we can dispense with this assumption, but we keep it here because it simplifies the general discussion. On the other hand, in order to make this discussion more general and to infer easily some important conclusions, let us work with the Riccati equation

$$s'(x,\mu)=\frac{V(x)}{p(x)W(x)}\frac{u_2^2(x)}{\phi(\mu)}(\phi(\mu)f(x)+s(x,\mu))^2 \tag{48}$$

deduced from (6) by the change of notations:

$$s(x,\mu)=S(x)\phi(\mu) \tag{49}$$

where μ is a positive parameter and $\phi(\mu)$ is finite.

In this section as in the preceding one, we transform $s(x,\mu)$ into a trigonometric function whose number of poles is easier to compute [4]:

$$\beta \tan g(x)=\phi(\mu)f(x)+s(x,\mu) \tag{50}$$

where β is a new positive parameter. The increasing function $g(x)$ is such that $g(a)$ never equals the values of $(\pi/2+k\pi; k\in Z)$. Therefore, in order to define $g(x)$ without ambiguity let us take

$$-\pi/2<g(a)<\pi/2. \tag{51}$$

Our hypotheses allow us to conclude that the total number of poles n of the phase function $S(x)$ on the whole interval $[a, b]$ is exactly equal to the total number of poles of the function $\tan(g(x))$ on the same interval. Thus we may write as before that

$$n=\left\{\left\{\frac{g(b)}{\pi}+\frac{1}{2}\right\}\right\}, \tag{52}$$

where the symbol $\{\{a\}\}$ means the positive integral part of a, i.e., $\{\{a\}\}=[a]$ if $a>0$, $\{\{a\}\}=0$ if $a\leqq 0$. Furthermore we notice that $g(x)$ is solution of the differential equation

$$g'(x)=\frac{\phi(\mu)f'(x)\cos^2 g(x)}{\beta}+\frac{V(x)u_2^2(x)\beta}{p(x)W(x)\phi(\mu)}\sin^2 g(x) \tag{53}$$

with the boundary condition

$$g(a) = \arctan \frac{\phi(\mu) f(a)}{\beta}. \tag{54}$$

We may now write the obvious inequality:

$$g'(x) \geqq \min_x \left(\frac{\phi(\mu)}{\beta} f'(x), \frac{V(x)}{p(x)W(x)} \frac{u_2^2(x)\beta}{\phi(\mu)} \right) \tag{55}$$

The relations (52), (54) and (55) finally give the expression of the general lower bound on n:

$$L; \quad n \geqq \left\{ \left\{ \frac{1}{2} + \frac{1}{\pi} \arctan \left(\frac{\phi(\mu) f(a)}{\beta} \right) + \frac{1}{\pi} \int_a^b \min_x \left(\frac{\phi(\mu)}{\beta} f'(x), \frac{V(x)\beta}{p(x)W(x)} \frac{u_2^2(x)}{\phi(\mu)} \right) dx \right\} \right\}. \tag{56}$$

We see that the lower bound L depends on the values of the parameters β, μ and on the form of the function $\phi(\mu)$. If for example we put

$$\phi(\mu) = \mu^{1/2} \quad \text{and} \quad \mu = \nu \tag{57}$$

where ν is defined as in (44), the relation (56) becomes:

$$L_1: \quad n \geqq \left\{ \left\{ \frac{1}{2} + \frac{1}{\pi} \arctan \left(\frac{\nu^{1/2} f(a)}{\beta} \right) + \frac{1}{\pi} \sqrt{\nu} \int_a^b \min_x \left(\frac{1}{\beta} f'(x), \frac{v(x) u_2^2(x) \beta}{p(x) W(x)} \right) dx \right\} \right\}, \tag{58}$$

and we obtain the very important conclusion:

(59) *The number of poles of the phase function increases at least as the square root of the strength of the perturbation.*

4. General conclusion. The main conclusion of this part is contained in (46) and (59), namely:

(60) *For a perturbing function $V(x) = \nu v(x)$ whose shape $v(x)$ is fixed, the number of poles of the phase function increases asymptotically as the square root of the strength parameter ν.*

Remark. From this conclusion, it is clear that the upper bound of type U cannot be very good for large ν because I_1 and $(I_0 I_2)^{1/2}$ behaves like ν and not $\nu^{1/2}$ as the bounds of type U'.

PART II. APPLICATIONS

5. Number of extrema of $u(x)$: $u''(x) + Q(x)u(x) = 0$. Let us consider the second order differential equation

$$u''(x) + Q(x)u(x) = 0 \tag{61}$$

with

$$Q(x) = -V(x) \geqq 0, \qquad a = 0, \qquad u(0) = 0, \qquad u'(0) \neq 0, \tag{62}$$

and let us try to give bounds on the number of extrema of the solution $u(x)$ by using the theoretical discussion of Part I [8].

It is easy to see that the assumptions (2) are true in this case. The two linearly independent solutions u_1, u_2 of negative Wronskian are:

$$u_1 = x, \qquad u_2 = 1. \tag{63}$$

The Riccati phase equation (6) becomes

$$S'(x) = Q(x)(x + S(x))^2. \tag{64}$$

Let us set

$$\gamma(x) = \frac{u(x)}{u'(x)}; \tag{65}$$

then

$$\gamma(x) = x + S(x) \tag{66}$$

and the poles of $S(x)$ are poles of $\gamma(x)$ so that the number of extrema of the solution $u(x)$ equals the number of poles of $S(x)$.

The integrals I_i defined in (17) become now:

$$I_i = \int_0^b Q(x) x^i \, dx, \qquad i = 0, 1, 2. \tag{67}$$

Thus we first obtain the upper bound

$$U_1: \qquad n \leqq \frac{1}{2} + \frac{2}{\pi}\left[\int_0^b Q(x)\, dx \int_0^b Q(x) x^2 \, dx\right]^{1/2}. \tag{68}$$

The integral I_1 being positive, we can apply the second upper bound U_2, which becomes

$$U_2: \qquad n \leqq 1 + \frac{2}{\pi}\left\{\int_0^b Q(x)\, dx \int_0^b Q(x) x^2 \, dx - \left[\int_0^b Q(x) x \, dx\right]^2\right\}^{1/2}. \tag{69}$$

It is now obvious that the assumptions (26) hold with $f(0) = 0$ and $F = 1$. We therefore obtain the second upper bound (U') as:

$$n \leqq \frac{2}{\pi} \int_0^b \sqrt{Q(x)}\, dx \tag{70}$$

or

$$n \leqq \frac{2\sqrt{\nu}}{\pi} \int_0^b \sqrt{q(x)}\, dx. \tag{71}$$

Let us at once set $Q(x) = \nu q(x)$, $\phi(\mu) = \sqrt{\mu}$, and $\mu = \nu$, so that we may obtain the lower bound (L)

$$n \geqq \left\{ \left\{ \frac{1}{2} + \frac{\sqrt{\nu}}{\pi} \int_0^b \min_x \left(\frac{1}{\beta}, q(x)\beta \right) dx \right\} \right\}, \tag{72}$$

where β is still arbitrary.

From Elbert's [9] and Makai's [10] works on the number of zeros of solutions of (61) and from the obvious relation between the number of zeros (n_Z) and the number of extrema (n_E) of the oscillatory solution of (61)

$$n_E \geqq n_Z - 1 \tag{73}$$

it is easy to verify that in some cases ($Q(x)$ small) our upper bounds on n_E give upper bounds on n_Z better than Elbert's one.

6. Number of eigenvalues smaller than a fixed number Λ. Let us try to characterize the spectrum of the operator

$$L = -\frac{d^2}{dx^2} + Q(x). \tag{74}$$

We therefore study the equation

$$u''(x) + (\lambda - Q(x))u(x) = 0 \tag{75}$$

with the boundary conditions

$$\begin{aligned} &u(0) = 0, \\ &u(X) \neq 0 \quad \text{(where } X \text{ is an arbitrary positive given value)} \end{aligned} \tag{76}$$

and with the following assumptions on $Q(x)$:

$$\begin{aligned} &Q(x) \in C^0[0, \infty], \\ &Q(x) \text{ is a steadily increasing function to infinity,} \\ &Q(0) = 0. \end{aligned} \tag{77}$$

Let us compute the number $N(\Lambda)$ of eigenvalues (λ_i) of the operator L, whose values are less than a known number Λ. Let us put

$$\lambda_n \leqq \Lambda < \lambda_{n+1}, \tag{78}$$

and X, and X_n such that

$$\begin{aligned} &Q(X) = \Lambda, \\ &u(X) \neq 0 \quad \text{(hypothesis (76))}, \\ &Q(X_n) = \lambda_n. \end{aligned} \tag{79}$$

We know from Titchmarsh [11] that X_n is less than X and that the number of zeros and of extrema of the eigenfunction $\psi_n(x)$ are the same on $[0, X]$ as on $[0, X_n]$. Because the number of zeros different from $x = 0$ of the eigenfunction $\psi_n(x)$ and its number of extrema differ from one unit, we can say that $N(\Lambda)$ equals the

number of extrema N_E of $\psi_n(x)$. We thus can use the results of § 5 to compute N_E.

We also have:

$$\begin{aligned} &V(x)=Q(x)-\Lambda \quad \text{(negative on } [0, X]), \\ &u_1(x)=x, \\ &u_2(x)=1 \quad (W=-1). \end{aligned} \tag{80}$$

We obtain therefore

$$U_1: \qquad N(\Lambda)\leqq\frac{1}{2}+\frac{2}{\pi}\left(\int_0^X [\Lambda-Q(x)]\,dx \int_0^X [\Lambda-Q(x)]x^2\,dx\right)^{1/2} \tag{81}$$

$$\begin{aligned} U_2: \qquad N(\Lambda)\leqq 1+\frac{2}{\pi}\Bigg(&\int_0^X [\Lambda-Q(x)]\,dx \int_0^X [\Lambda-Q(x)]x^2\,dx \\ &-\left(\int_0^X [\Lambda-Q(x)]x\,dx\right)^2\Bigg)^{1/2}. \end{aligned} \tag{82}$$

Assumptions (26) also hold with $f(0)=0$ and $F=1$ so that we have

$$U_1': \qquad N(\Lambda)\leqq\frac{2}{\pi}\int_0^X \sqrt{\Lambda-Q(x)}\,dx \equiv 2I. \tag{83}$$

But Titchmarsh gives a lower bound using the same integral as (83), which is

$$N(\Lambda)\geqq\frac{1}{\pi}\int_0^X \sqrt{\Lambda-Q(x)}\,dx-1; \tag{84}$$

we thus have the important relation

$$I-1\leqq N(\Lambda)\leqq 2I. \tag{85}$$

We also can find another lower bound than Titchmarsh's one, and it is

$$L: \qquad N(\Lambda)\geqq\left\{\left\{\frac{1}{2}+\frac{1}{\pi}\int_0^X \min_x\left(\frac{\phi(\mu)}{\beta}, \frac{(\Lambda-Q(x))\beta}{\phi(\mu)}\right)dx\right\}\right\}. \tag{86}$$

Since $[1-q(x)]$ decreases from 1 to 0 between $x=0$ and $x=X$, formula (86) becomes

$$L: \qquad N(\Lambda)\geqq\left\{\left\{\frac{1}{2}+\frac{\sqrt{\Lambda}}{\pi}\int_0^X (1-q(x))\,dx\right\}\right\} \tag{87}$$

with

$$\begin{aligned} &\phi(\mu)=\mu^{1/2} \quad \text{and} \quad \mu=\Lambda, \\ &Q(x)=\Lambda q(x), \qquad \beta=1. \end{aligned} \tag{88}$$

Numerical example. Let us set

$$Q(x)=x^2, \tag{89}$$

which evidently satisfies the conditions (77). We know that in this case the

spectrum of the operator is defined as follows [11]:

$$\lambda_n = 2n+1.$$

Then we have $X=\sqrt{\Lambda}$ and thus

$$U_1': \qquad N(\Lambda)\leqq\frac{\Lambda}{2},$$

$$L: \qquad N(\Lambda)\geqq\left\{\left\{\frac{1}{2}+\frac{2\Lambda}{3\pi}\right\}\right\},$$

and Titchmarsh's lower bound gives

$$N(\Lambda)\geqq\Lambda/4-1.$$

If we set, for example, $\Lambda=10$, we know that $N(\Lambda)$ exactly equals 5 while the best upper bound (U_1') gives

$$N(\Lambda)\leqq\Lambda/2=5$$

and the lower bound and the Titchmarsch's one give

$$N(\Lambda)\geqq 2.$$

7. Number of zeros of some special functions. Our purpose in this section is to give bounds on the number of zeros of some special functions which are solutions of linear second order differential equations [12].

a) Let us first study the Bessel cylindrical differential equation ($\rho>0$)

$$u''(x)+\frac{1}{x}u'(x)-\frac{\rho^2}{x^2}u(x)=-\lambda^2 u(x), \qquad 0\leqq x\leqq 1, \tag{90}$$

whose regular solution is the Bessel function of the first kind $J_\rho(\lambda x)$. We know that:

$$\begin{aligned}&\text{the point } x=1 \text{ is a zero of } J_\rho(\lambda x)\\ &\text{iff } \lambda \text{ is a zero of } J_\rho(x).\end{aligned} \tag{91}$$

Thus it will be easy to compute the number of zeros of $J_\rho(x)$ whose values are less than λ.

Let the two linearly independent solutions of the Euler equation, u_1 and u_2 be:

$$u_1(x)=x^\rho, \qquad u_2(x)=x^{-\rho}-x^\rho, \qquad W=-2\rho/x. \tag{92}$$

It is then easy to see that $u(0)$ finite and $u(1)=0$ give

$$S(0)=0 \quad \text{and} \quad S(1)=\infty. \tag{93}$$

Thus, the number of zeros less than λ of $J_\rho(x)$ exactly equals the number of poles of $S(x)$ in [0, 1]. The phase equation (6) is now:

$$S'(x)=\frac{\lambda^2 x}{2\rho}[x^\rho+S(x)(x^{-\rho}-x^\rho)]^2 \tag{94}$$

and we can try to deduce the bounds U, U' and L.

Let us remark that the existence of I_0 is ensured only if $\rho < 1$. Under this condition we may write

$$U_1: \qquad n \leqq \frac{1}{2} + \frac{1}{\pi} \frac{\lambda^2}{\sqrt{2}(1+\rho)\sqrt{1-\rho}}; \tag{95}$$

and since I_1 is positive,

$$U_2: \qquad n \leqq 1 + \frac{\lambda^2}{2\pi\sqrt{1-\rho^2}}. \tag{96}$$

$f(x)$ and $f'(x)$ do not satisfy the conditions (26) so that we can apply neither (41) nor (43).

For the computation of a lower bound, the assumption (47) can be relaxed if we remark that (50) allows us to say that:

$$0 \leqq n_e - n \leqq 1 \tag{97}$$

where n_e is the computed number of poles of $\tan g(x)$ and n is the number of poles of $S(x)$.

Now L becomes

$$n \geqq \left\{ \left\{ \frac{1}{2} + \frac{\lambda}{\pi} \int_0^1 \min_x \left(\frac{x^{2\rho-1}}{(1-x^{2\rho})^2}, \frac{(1-x^{2\rho})^2}{x^{2\rho-1}} \right) dx \right\} \right\} - 1 \tag{98}$$

if we set

$$\phi(\mu) = \mu \quad \text{with } \mu = \lambda, \qquad \beta = 2\rho,$$

and there is no restriction on the values of ρ.

Now we can relax the restriction $0 < \rho < 1$ by replacing ρ by $-\rho$, and we can show by the same argument that in that case the upper and lower bounds are still valid inside the larger interval

$$0 < |\rho| < 1. \tag{99}$$

b) Let us now make same study on the regular Coulomb wave function $F_L(\eta, r)$. For this reason, we study the following Coulomb differential equation [13]:

$$u''(r) - \frac{L(L+1)}{r^2} u(r) = -\left(k^2 - \frac{2\eta k}{r} \right) u(r), \qquad 0 \leqq x \leqq 1, \tag{100}$$

with

$$r > 0, \quad -\infty < \eta < +\infty, \quad L \geqq 0, \tag{101}$$

whose regular solution is $F_L(\eta, kr)$. As for the Bessel function, we can say that $r = 1$ is a zero of $F_L(\eta, kr)$ iff k is a zero of $F_L(\eta, r)$.

In order to have $V(x) \leqq 0$ on $[0, 1]$, we restrict the analysis to $\eta < 0$. The linearly independent solutions $u(r)$ are:

$$u_1 = r^{L+1}, \qquad u_2 = r^{-L} - r^{L+1}, \qquad W = -(2L+1). \tag{102}$$

We have $S(0)=0$ and on the other hand, the condition that $F_L(\eta, kr)$ vanishes at $x=1$ implies that $S(1)=\infty$. Thus the number of zeros less than k of the Coulomb wave function $F_L(\eta, r)$ exactly equals the number of poles of $S(x)$ located on $[0, 1]$. We can thus give upper and lower bounds on this number n.

The integrals I_i defined in (17) are:

$$
\begin{aligned}
I_0 &= \frac{k^2(2L+1)}{(1-2L)(2L+3)} + \frac{k\eta(2L+1)}{L(L+1)} \qquad (L<\tfrac{1}{2}),\\
I_1 &= \frac{k^2}{2(2L+3)} - \frac{k\eta}{L+1} \qquad (L\geqq 0),
\end{aligned} \tag{103}
$$

and

$$
I_2 = \frac{k^2}{(2L+1)(2L+3)} - \frac{k\eta}{(2L+1)(L+1)} \qquad (L\geqq 0).
$$

We can then give the first upper bound on n:

$$
U_1: \qquad n \leqq \frac{1}{2} + \frac{2}{\pi}\sqrt{I_0 I_2}. \tag{104}
$$

Since I_1 is positive we also have

$$
U_2: \qquad n \leqq 1 + \frac{2}{\pi}(I_0 I_2 - I_1^2)^{1/2}. \tag{105}
$$

The lower bound is

$$
\begin{aligned}
L: \qquad n \geqq \Bigg\{ \Bigg\{ \frac{1}{2} + \frac{1}{\pi}\int_0^1 \min_r \Bigg(& \frac{1}{(r^{-L}-r^{L+1})^2},\\
& \left(k^2 - \frac{2k\eta}{r}\right)(r^{-L}-r^{L+1})^2 \Bigg)\, dr \Bigg\} \Bigg\} - 1
\end{aligned} \tag{106}
$$

with

$$
\beta = 2L+1, \qquad \phi(\mu) = 1.
$$

For $\eta = 0$, the equation (101) reduces to the Riccati spherical Bessel equation [13] whose solution has the same zeros as the spherical Bessel function, which have themselves the same zeros as the cylindrical Bessel function of order $\rho = 2L+1$. So, if we set

$$
\eta = 0, \qquad L = \rho - \tfrac{1}{2}, \tag{107}
$$

these bounds reduce to the previous one on the zeros of the Bessel function, as we can check immediately.

8. Characterization of stability intervals for $y''(x)+\lambda Q(x)y(x)=0$. This section is concerned with the stability zones for the linear periodic equation [14]

$$
y''(x) = -\lambda Q(x) y(x) \tag{108}
$$

$$Q(x) = Q(x+L), \qquad \lambda Q(x) \geqq 0. \tag{109}$$

The characteristic numbers λ_i' and $\lambda_j (1 \leqq i;\ 0 \leqq j)$ are defined respectively by the following boundary value problem [15]:

$$\begin{aligned} &y''(x) + \lambda' Q(x) y(x) = 0 \\ &y(0) = -y(L), \qquad y'(0) = -y'(L) \end{aligned} \tag{110}$$

(solutions $2L$ periodic) or

$$\begin{aligned} &y''(x) + \lambda Q(x) y(x) = 0 \\ &y(0) = y(L), \qquad y'(0) = y'(L) \end{aligned} \tag{111}$$

(solutions L periodic).

We will characterize the so-called stability intervals $]\lambda_n', \lambda_{n-1}[$ or $]\lambda_n, \lambda_{n+1}'[$ defined by:

$$\begin{aligned} &y(x, \lambda) \text{ is bounded for all } x \\ &\quad \text{iff there exists an } n \text{ such that } \lambda \text{ belongs to} \\ &\quad]\lambda_n', \lambda_{n-1}[\quad \text{or} \quad]\lambda_n, \lambda_{n+1}'[. \end{aligned} \tag{112}$$

Let us choose

$$y(0) = 0, \qquad y'(0) \neq 0. \tag{113}$$

We want to set up a Riccati equation giving bounds on λ_n and λ_n'; so let us define

$$\Gamma(x) = \frac{y(x)}{y'(x)} \quad \text{or} \quad \Gamma(x) = x + S(x). \tag{114}$$

The boundary conditions of problems (110) and (111), with the condition (113) impose:

$$S(0) = 0, \qquad S(L) = -L. \tag{115}$$

In order to introduce poles at $x = L$, let us make the homographic transformation

$$A(x) = \frac{S(x)}{S(x) + L}. \tag{116}$$

The function $A(x)$ satisfies the Riccati equation [16]

$$A'(x) = \frac{\lambda Q(x)}{L} (x + A(x)(L - x))^2 \tag{117}$$

with now

$$A(0) = 0, \qquad A(L) = \infty. \tag{118}$$

We know that λ belongs to the $(n+1)$st stability zone, with $n+1$ being even iff λ belongs to $]\lambda_{n+1}', \lambda_n[$ where λ_{n+1}' and λ_n respectively are the $(2n)$th and $(2n+1)$st values allowing $A(x)$ to satisfy the conditions (118); while if $n+1$ is odd the $(n+1)$st stability zone is $]\lambda_n, \lambda_{n+1}'[$ and the $(2n)$th and $(2n+1)$st values giving conditions (118) are λ_n and λ_{n+1}'. So, in order to characterize the $(n+1)$th stability

interval we will compute the number of poles of the solution $A(x)$ in the interval $[0, L]$.

Consequently (112) can be reduced to:

$$\begin{aligned} &\lambda \text{ belongs to the } (n+1)\text{st } (n=0,\cdots) \text{ stability zone} \\ &\text{iff } A(x,\lambda) \text{ has at least } 2n \text{ poles} \\ &\text{and at most } 2n+1 \text{ poles.} \end{aligned} \tag{119}$$

Relation between n and λ can therefore be obtained using Part I in the form:

$$2n \geqq L(\lambda), \qquad 2n+1 \leqq U(\lambda). \tag{120}$$

Let us now give $U(\lambda)$ and $L(\lambda)$ explicitly and let us try to write them in the forms (21), (24), (41) or (43) and (56). The corresponding equation (6) here is the Riccati equation (117) and the linearly independent solutions u_i are

$$u_1(x) = x, \qquad u_2(x) = L - x. \tag{121}$$

Thus, $W(x) = -L$ is negative on $[0, L]$ and the function $V(x)/(p(x))$ corresponding to $-\lambda Q(x)$ is also negative.

As in the preceding example, we cannot apply the bound U' because $u_2(x)$ vanishes at $x = L$. Moreover $f(x)$ has a singularity at the point $x = L$ and we can derive the lower bound with the same argument as before.

Consequently, the conditions (119) become (a):

$$\begin{aligned} &\lambda \geqq \lambda'_{n+1} \quad (\text{if } (n+1)\text{is even) or} \\ &\lambda \geqq \lambda_n \quad (\text{if } (n+1) \text{ is odd}) \end{aligned}$$

implies

$$\left\{\left\{\frac{1}{2} + \frac{\sqrt{\lambda}}{\pi}\int_0^L \min_x \left(\frac{1}{(L-x)^2}\right), Q(x)(L-x)^2\,dx\right\}\right\} \leqq 2n+1 \tag{122}$$

if we set

$$\begin{aligned} &\mu = \lambda \quad \text{and} \quad \phi(\mu) = \sqrt{\mu}, \\ &\beta = L, \end{aligned}$$

and (b):

$$\begin{aligned} &\lambda \leqq \lambda_n \quad (\text{if } (n+1) \text{ is even) or} \\ &\lambda \leqq \lambda'_{n+1} \quad (\text{if } (n+1) \text{ is odd}) \end{aligned}$$

implies

$$2n+1 \leqq \frac{1}{2} + \frac{2}{\pi}\frac{\lambda}{L}\left(\int_0^L Q(x)(L-x)^2\,dx \int_0^L q(x)x^2\,dx\right)^{1/2} \tag{123 (i)}$$

or

$$\text{(ii)} \qquad 2n+1 \leqq 1+\frac{2}{\pi}\frac{\lambda}{L}\left(\int_0^L Q(x)(L-x)^2\,dx \int_0^L Q(x)x^2\,dx - \left(\int_0^L Q(x)x(L-x)\,dx\right)^2\right)^{1/2}. \tag{124}$$

9. Characterization of stability intervals for $y''(x)+(\lambda+Q(x))y(x)=0$. The same study as before can be made for the linear periodic equation

$$\begin{aligned} &y''(x)+\lambda y(x) = -Q(x)y(x) \\ &Q(x+L)=Q(x), \qquad Q(x)\geqq 0. \end{aligned} \tag{125}$$

For a fixed value of λ, the solution $y(x,\lambda)$ is bounded iff there exists an integer n such that λ belongs to one of the two intervals:

$$]\lambda_{2n}, \lambda'_{2n+1}[, \qquad]\lambda'_{2n}, \lambda_{2n-1}[,$$

where the characteristic numbers λ_j, λ'_i $(1 \leqq i;\ 0 \leqq j)$ are defined as in the preceding application [17]. Let us choose

$$\begin{aligned} &\lambda = -l^2, \qquad l>0 \\ &y(0)=0; \qquad y'(0)\neq 0. \end{aligned} \tag{126}$$

Let us again set $\Gamma(x)=y(x)/(y'(x))$, so that we have the following conditions on $S(x)$:

$$S(0)=-1, \qquad S(L)=-e^{2lL}. \tag{127}$$

By introduction of the homographical transformation

$$A(x)=\frac{S(x)+1}{S(x)+e^{2lL}}, \tag{128}$$

we have the Riccati equation

$$A'(x)=\frac{Q(x)}{2l(e^{2lL}-1)}\{2\,Shlx+A(x)\,e^{lx}(e^{-2lx}e^{2lL}-1)\}^2, \qquad A(0)=0, \qquad A(L)=\infty. \tag{129}$$

Equivalently to (119), we have:

$$\begin{aligned} &\lambda \text{ belongs to the } n\text{th stability zone} \\ &\quad \text{iff } A(x) \text{ has at least } 2n-1 \text{ poles} \\ &\quad\quad \text{and at most } 2n \text{ poles} \end{aligned} \tag{130}$$

and, as before, the relations between n and λ are:

$$(2n-1)\geqq L(\lambda) \quad \text{and} \quad 2n \leqq U(\lambda). \tag{131}$$

The corresponding equation (6) now is (129) and the functions $u_i(x)$ are

$$\begin{aligned} &u_1(x)=2\,Sh\,lx, \qquad W(u_1,u_2)=2l(1-e^{2lL})<0, \\ &u_2(x)=e^{lx}(e^{-2lx}e^{2lL}-1). \end{aligned} \tag{132}$$

Thus we have

$$I_0 = \int_0^L \frac{Q(x)}{2l(e^{2lL}-1)} e^{2lx}(e^{2l(L-x)}-1)^2\,dx,$$

$$I_1 = \int_0^L \frac{Q(x)}{l(e^{2lL}-1)} Sh\,(lx)\,e^{lx}(e^{2l(L-x)}-1)\,dx, \tag{133}$$

$$I_2 = \int_0^L \frac{2Q(x)}{l(e^{2lL}-1)} Sh^2\,(lx)\,dx.$$

Consequently, by the same argument as before, we immediately may write (a):

$$\lambda = -l^2 > \lambda_n' \quad \text{(if } n \text{ is even) or}$$

$$\lambda > \lambda_{n-1} \quad \text{(if } n \text{ is odd)}$$

implies

$$\left\{\left\{\frac{1}{2}+\frac{1}{\pi}\int_0^L \min_x \left(\frac{2l(e^{2lL}-1)}{e^{2lx}(e^{-2lx}e^{2lL}-1)^2},\right.\right.\right. \tag{134}$$
$$\left.\left.\left.\frac{Q(x)\,e^{2lx}(e^{-2lx}e^{2lL}-1)^2}{2l(e^{2lL}-1)}\right)dx\right\}\right\} \leqq 2n$$

and (b):

$$\lambda < \lambda_{n-1} \quad \text{(if } n \text{ is even) or}$$

$$\lambda < \lambda_n' \quad \text{(if } n \text{ is odd)}$$

implies

$$\text{(i)} \qquad 2n < \frac{1}{2}+\frac{2}{\pi}\sqrt{I_0 I_1} \tag{135}$$

or

$$\text{(ii)} \qquad 2n < 1+\frac{2}{\pi}\sqrt{I_0 I_2 - I_1^2}. \tag{136}$$

(Let us remark that we have choosen $\beta/(\phi(\mu)) = 1$.)

It is interesting to note that if we put

$$\lambda + Q(x) = \Lambda q(x) \tag{137}$$

we find the results of the preceding section if we compute the limit as $\lambda \to 0$.

10. Conclusions. The domain of applicability of the methods we used is quite large and of course not limited to the given mathematical examples we work with. As indicated before these techniques were introduced by Calogero [4] to count the number of bound states of a central potential.

The generalization we obtain here allows us to extend the Calogero's work to more general Schrödinger equations. We already obtain results for the first bound state in the spheroidal case [18] and for the first band in the periodic case [19].

Conditions on the nth bound state in the spheroidal case will be published elsewhere.

Furthermore, scalar phase equations can be written for a 2×2 first order linear system of differential equations [5], [8], and matrix phase equations can also be constructed for linear second order matrix differential systems of arbitrary dimension [2], [20]. It is probably true that the methods we used in this work are partly applicable to these more general phase equations.

REFERENCES

[1] G. N. WATSON. *A Treatise on the Theory of Bessel Functions*, Cambridge at the University Press, second ed., Cambridge, England, 1966.

[2] F. CALOGERO, *Variable Phase Approach to Potential Scattering*, Academic Press, New York, 1967.

[3] A. RONVEAUX, *Approximations de l'équation de phase et applications à la théorie quantique des collisions*, C.R. Acad. Sci. Paris Ser. A, 266 (1968), pp. 306–308.

[4] F. CALOGERO, *Upper and lower limits for the number of bound states in a given central potential*, Comm. Math. Phys., 1 (1965), pp. 80–88.

[5] A. RONVEAUX, *Phase equation in quantum mechanics*, Amer. J. Phys., 37 (1969), pp. 135–141.

[6] ———, *Even and odd spectra of one-dimensional symmetrical potential. I: Schrödinger theory*, Ibid., 40 (1972), pp. 888–892.

[7] O. BORŮVKA, *Linear Differential Transformations of the Second order*, The English Universities Press Ltd., London, 1971.

[8] A. RONVEAUX, *Equations différentielles du second ordre: Distances entre zéro et extremum des solutions*, Ann. Soc. Sci. Bruxelles, 84 (1970), pp. 5–20.

[9] A. ELBERT, *On the solutions of the differential equation $y''+q(x)y=0$, where $[q(x)]^r$ is concave. I*, Acta Math. Acad. Sci. Hungar., 20 (1969), pp. 1–11.

[10] E. MAKAI, *Über die Nullstellen von Funktionen, die Lösungen Sturm-Liouville'scher Differentialgleichungen sind*, Comm. Math. Helv., 16 (1943–44), pp. 153–199.

[11] F. C. TITCHMARSH, *Eigenfunctions Expansions*, Oxford at the Clarendon Press, second ed., Oxford, England, 1962.

[12] A. RONVEAUX AND A. MOUSSIAUX, *A priori bound on the first zero of some special functions*, Ann. Soc. Sci. Bruxelles, 88 (1974), pp. 169–175.

[13] M. ABRAMOWITZ AND I. A. STEGUN, *Handbook of Mathematical Functions*, Dover, New York, 1965.

[14] M. G. KREIN, *On certain problems on the maximum and the minimum of characteristic values and on the Lyapunov zones of stability*, Amer. Math. Soc. Transl., 2 (1955), no. 1, pp. 163–187.

[15] A. LIAPOUNOFF, *Sur une équation différentielle linéaire du second ordre*, C.R. Acad. Sci. Paris, 128 (1899), pp. 910–913.

[16] A. RONVEAUX, *Comparative properties of the single, double and periodic potentials*, Lett. Nuovo Cimento, 10 (1974), pp. 523–528.

[17] W. MAGNUS AND S. WINKLER, *Hill's Equation*, Interscience, New York, 1966.

[18] A. RONVEAUX, M-C. DUMONT-LEPAGE AND R. GERARD, *Potentiels sphéroïdaux pour l'equation de Schroedinger I, Caractérisation d'un état lié*, Ann. Inst. H. Poincaré Sér. A, 22 (1975), pp. 291–304.

[19] A. RONVEAUX, *Etats liés à energie nulle en théorie de Schroedinger*, Internat. Congress of Mathematicians, Vancouver, 1974.

[20] ———, *Résolution des équations de Riccati rencontrées en théorie des collisions*, C.R. Acad. Sci. Paris, 270 (1970), pp. 77–79.

SIAM J. MATH. ANAL.
Vol. 8, No. 6, November 1977

LINEAR SYSTEMS OF DIFFERENTIAL EQUATIONS WITH SINGULAR COEFFICIENTS*

STEPHEN L. CAMPBELL†

Abstract. Differential equations of the form $A\dot{x}+Bx=f$ are studied where A, B are $m\times n$ matrices. Explicit solutions are derived for several cases of interest. One such case is when there exists a scalar λ such that $\lambda A+B$ is of full rank. Another includes the case when A, B are normal matrices and one is positive semidefinite. The application of these results to linear autonomous control processes is discussed.

1. Introduction. In [2], closed forms for all solutions of the differential equation

$$A\dot{x}+Bx=f \tag{1}$$

were given for the case when A, B were $n\times n$ matrices and (1) had unique solutions for consistent initial conditions. The results of [2] were applied to a variety of optimal control problems in [3]. As observed in [3], a wider class of control problems could be handled by these methods if (1) could be solved when A, B were $m\times n$ instead of $n\times n$.

Ways of solving (1) for general A, B exist. See, for example, Gantmacher [4]. However, we seek explicit closed form solutions of (1) and not just a method of solution as done in [4].

As in [2], [3], we shall make assumptions on A and B in (1). We then seek to determine: for which f is (1) consistent, for which initial conditions x_0 is (1) consistent, and what is the expression for the general solution? As to be expected the approach of [2] does not seem to extend to cover (1) for all A, B. However, we have been able to explicitly solve (1) for several major cases of interest.

Section 2 will develop some needed results. Section 3 will discuss the case when A, B are $n\times n$. In § 4, (1) is solved when $\lambda A+B$ is one-to-one. The next section completely solves (1) when $\lambda A+B$ is onto. Applications of §§ 1–5 to control theory are discussed in § 6.

Our notation and terminology is the same as [2]. In particular, A^D denotes the Drazin inverse of a square matrix A, $A^\dagger$ denotes the Moore–Penrose inverse of a rectangular matrix A, and X is a (2)-inverse of A if $XAX=X$.

An $n\times n$ matrix A is called EP if its range is perpendicular to its null space, or equivalently, $AA^\dagger=A^\dagger A$ [1]. If A is EP and has rank r, it is sometimes called EP_r. Normal, and hence Hermitian, matrices are always EP.

We shall use $\oplus$ to denote an orthogonal sum. Suppose $\mathbb{C}^n=M_1\oplus M_2$ where M_1, M_2 are subspaces. Then $T=T_1\oplus T_2$, T an $n\times n$ matrix, means that if U is a unitary matrix whose first r columns are an orthogonal basis for M_1 and whose next $n-r$ columns are an orthogonal basis for M_2, then

$$U^*TU=\begin{bmatrix}T_1 & 0\\ 0 & T_2\end{bmatrix}.$$

* Received by the editors December 10, 1975, and in revised form July 26, 1976.

† Department of Mathematics, North Carolina State University at Raleigh, Raleigh, North Carolina 27607. This research was supported in part by a grant from the North Carolina Engineering Foundation.

If $f(s)$ is a vector valued function, then $f = f_1 \oplus f_2$ means that $f_1(s) \in M_1, f_2(s) \in M_2$ for all s.

Generalized inverses are also used to study differential equations in [7]. Our results and approach are substantially different from theirs.

In this paper we have omitted many of the more obvious corollaries to our results. Their inclusion would have made this paper unreasonably long. For example, it is a simple matter for the reader to take Theorems 1, 4, 5, 6, 7, 9, 10 and derive the appropriate statements about consistent initial conditions.

2. Reduction to the commuting case. A key step in [2] was to replace A, B, f in (1) by $\hat{A}, \hat{B}, \hat{f}$. This did not change the solutions, but extensive use was made of the fact that $\hat{A}\hat{B} = \hat{B}\hat{A}$. This section will develop the necessary technical lemmas. Their application to differential equations will begin in § 3.

PROPOSITION 1. *Suppose that A, B are $m \times n$ matrices. Let $(\cdot)^0$ denote a (2)-inverse. Then the following are equivalent for $\lambda \neq 0$.*

(i) $(\lambda A + B)^0 A$, $(\lambda A + B)^0 B$, *commute.*

(ii) $(\lambda A + B)(\lambda A + B)^0 A[I - (\lambda A + B)^0(\lambda A + B)] = 0$.

(iii) $(\lambda A + B)(\lambda A + B)^0 B[I - (\lambda A + B)^0(\lambda A + B)] = 0$.

Proof.

$$\begin{aligned}
&\lambda(\lambda A + B)^0 A(\lambda A + B)^0 B - \lambda(\lambda A + B)^0 B(\lambda A + B)^0 A \\
&\quad = (\lambda A + B)^0\{(\lambda A + B)(\lambda A + B)^0 B - B(\lambda A + B)^0(\lambda A + B)\} \\
&\quad = (\lambda A + B)^0\{B - B(\lambda A + B)^0(\lambda A + B)\} \\
&\quad = (\lambda A + B)^0 B\{I - (\lambda A + B)^0(\lambda A + B)\}.
\end{aligned}$$

Thus (i) and (iii) are equivalent. That (ii) and (iii) are equivalent follows from the fact that

$$(\lambda A + B)^0(\lambda A + B)[I - (\lambda A + B)^0(\lambda A + B)] = 0.$$

Hence

$$(\lambda A + B)^0 B[I - (\lambda A + B)^0(\lambda A + b)] = 0$$

if and only if $(\lambda A + B)^0 A[I - (\lambda A + B)^0(\lambda A + B)] = 0$. □

Note that if the Moore–Penrose inverse is used, then (ii) of Proposition 1 says that

$$P_{\mathcal{R}(\lambda A + B)} A P_{\mathcal{N}(\lambda A + B)} = 0 \tag{2}$$

where P_M is the orthogonal projection onto the subspace M.

The next two results will be useful in what follows. Note that if $\mathcal{N}(\lambda A + B) = \mathcal{N}(A) \cap \mathcal{N}(B)$ then (ii) holds in Proposition 1.

PROPOSITION 2. *Suppose that A, B are Hermitian. Then $(\lambda A + B)^\dagger A$, $(\lambda A + B)^\dagger B$ commute for $\lambda \neq 0$ if and only if there exists a $\tilde{\lambda}$ such that $\mathcal{N}(\tilde{\lambda} A + B) = \mathcal{N}(A) \cap \mathcal{N}(B)$. Furthermore, if $\tilde{\lambda}$ exists, then $(\tilde{\lambda} A + B)^\dagger A$, $(\tilde{\lambda} A + B)^\dagger B$ commute.*

Proof. Suppose that A, B are Hermitian and $\lambda \neq 0$. Then $(\lambda A+B)^{\dagger}A$, $(\lambda A+B)^{\dagger}B$ commute if and only if

$$P_{\mathcal{R}(\lambda A+B)}AP_{\mathcal{N}(\lambda A+B)}=0, \qquad P_{\mathcal{R}(\lambda A+B)}BP_{\mathcal{N}(\lambda A+B)}=0 \tag{3}$$

by Proposition 1. We may assume λ is real. Thus (3) is equivalent to

$$AP_{\mathcal{N}(\lambda A+B)}=P_{\mathcal{N}(\lambda A+B)}AP_{\mathcal{N}(\lambda A+B)}, \tag{4}$$

and

$$BP_{\mathcal{N}(\lambda A+B)}=P_{\mathcal{N}(\lambda A+B)}BP_{\mathcal{N}(\lambda A+B)}. \tag{5}$$

If $\mathcal{N}(\lambda A+B)=\mathcal{N}(A)\cap\mathcal{N}(B)$, then (3) obviously follows. Suppose then that (3) holds for λ_1. If $\mathcal{N}(\lambda_1 A+B)=\mathcal{N}(A)\cap\mathcal{N}(B)$ we are done. If not, observe that (4), (5) imply that $\mathcal{N}(\lambda_1 A+B)$ is an invariant, and hence reducing, subspace for both A and B. Then relative to

$$\mathbb{C}^n=\mathcal{R}(\lambda_1 A+B)\oplus\{\mathcal{N}(\lambda_1 A+B)\cap[\mathcal{N}(A)\cap\mathcal{N}(B)]^{\perp}\}\ \oplus\mathcal{N}(A)\cap\mathcal{N}(B)$$

we have

$$A=A_1\oplus A_2\oplus 0, \qquad B=B_1\oplus -\lambda_1 A_2\oplus 0$$

where $\lambda_1 A_1+B_1$ and A_2 are invertible. Since $\lambda_1 A_1+B_1$ is invertible, it is invertible for all by a finite number of λ's. Let $\tilde{\lambda}$ be one of these other λ's. □

PROPOSITION 3. *If* $A, B\in\mathbb{C}^{n\times n}$ *are such that one is* EP *and the other is positive semidefinite, then there exists* λ *such that* $\lambda A+B$ *is invertible if and only if* $\mathcal{N}(A)\cap\mathcal{N}(B)=\{0\}$.

Proof. For convenience assume that A is EP and B is positive semidefinite. The proof when B is EP and A positive semidefinite will be similar. Let $B^{1/2}$ denote the unique positive square root of B. If there exists λ such that $(\lambda A+B)$ is invertible, then clearly $\mathcal{N}(A)\cap\mathcal{N}(B)=\{0\}$. Suppose then that $\mathcal{N}(A)\cap\mathcal{N}(B)=\{0\}$. If $A=0$ or $\mathcal{N}(A)=\{0\}$ we are done. Suppose not. Since A is EP we have that there exists a unitary matrix U such that

$$\tilde{A}=UAU^*=\begin{bmatrix}A_1 & 0\\ 0 & 0\end{bmatrix}, \qquad A_1 \text{ invertible.}$$

Define B_1, B_2, B_3 by

$$\tilde{B}=UBU^*=\begin{bmatrix}B_1 & B_2\\ B_2^* & B_3\end{bmatrix}.$$

We claim that B_3 is invertible. Suppose not. If $B_3\phi=0$, then let $\tilde{\phi}=\begin{bmatrix}0\\ \phi\end{bmatrix}$. We then have $0=\tilde{\phi}^*\tilde{B}\tilde{\phi}=\tilde{\phi}^*\tilde{B}^{1/2}\tilde{B}^{1/2}\tilde{\phi}=\|\tilde{B}^{1/2}\tilde{\phi}\|^2$. Hence $\tilde{B}\tilde{\phi}=0$. But $\tilde{A}\tilde{\phi}=0$ which contradicts $\mathcal{N}(A)\cap\mathcal{N}(B)=\{0\}$. Observe that for large $|\lambda|$, (λA_1+B_1) is invertible

since A_1 is. Now

$$\begin{bmatrix} B_3 & -B_3B_2B_3^{-1} \\ -(\lambda A_1+B_1)B_2^*(\lambda A_1+B_1)^{-1} & (\lambda A_1+B_1) \end{bmatrix}\begin{bmatrix} (\lambda A_1+B_1) & B_2 \\ B_2^* & B_3 \end{bmatrix}$$
$$= \begin{bmatrix} B_3(\lambda A_1+B_1)-B_3B_2B_3^{-1}B_2^* & 0 \\ 0 & -(\lambda A_1+B_1)B_2^*(\lambda A_1+B_1)^{-1}B_2+(\lambda A_1+B_1)B_3 \end{bmatrix}$$
$$= \begin{bmatrix} Q_1(\lambda) & 0 \\ 0 & Q_2(\lambda) \end{bmatrix}.$$

That $\lambda\tilde{A}+\tilde{B}$ (and hence $\lambda A+B$) is invertible for large λ will follow if $Q_1(\lambda)$, $Q_2(\lambda)$ are invertible for large λ. But

$$Q_1(\lambda)=\{B_3-B_3B_2B_3^{-1}B_2^*(\lambda A_1+B_1)^{-1}\}(\lambda A_1+B_1)$$

and $B_3B_2B_3^{-1}B_2^*(\lambda A_1+B_1)^{-1}\to 0$ and $\lambda\to\infty$. Since B_3 is invertible and (λA_1+B_1) is invertible for large λ, we get $Q_1(\lambda)$ is invertible for large λ. Similarly

$$Q_2(\lambda)=(\lambda A_1+B_1)\{B_3-B_2^*(\lambda A_1+B_1)^{-1}B_2\}$$

is invertible for large λ. □

PROPOSITION 4. *If $A, B\in\mathbb{C}^{n\times n}$ are such that one is* EP *and one is positive semidefinite, then there exists λ such that $\mathcal{N}(\lambda A+B)=\mathcal{N}(A)\cap\mathcal{N}(B)$.*

Proof. Since A, B are both EP we have that $\mathcal{N}(A)$ reduces A, $\mathcal{N}(B)$ reduces B. Hence $\mathcal{N}(A)\cap\mathcal{N}(B)$ reduces both A and B. Relative to the decomposition $\mathbb{C}^n=[\mathcal{N}(A)\cap\mathcal{N}(B)]^{\perp}\oplus[\mathcal{N}(A)\cap\mathcal{N}(B)]$ we get that A, B are unitarily equivalent to matrices of the form

$$\begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \begin{bmatrix} B_1 & 0 \\ 0 & 0 \end{bmatrix}$$

respectively where A_1, B_1 are such that one is EP and one is positive semidefinite. Furthermore, $\mathcal{N}(A_1)\cap\mathcal{N}(B_1)=\{0\}$. Proposition 4 now follows from Proposition 3. □

It is not possible to weaken the assumption that one of A, B is positive semi-definite in Propositions 3 and 4 even if the other matrix is required to be Hermitian.

Example 1. Let

$$A=\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad B=\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Then $A=A^*$, $B=B^*$ and $\mathcal{N}(A)\cap\mathcal{N}(B)=\{0\}$. But the vector $[+1, -1, -\lambda]^*\in\mathcal{N}(\lambda A+B)$ for all λ.

Throughout this paper we shall frequently use the following proposition without referring to it explicitly.

PROPOSITION 5. *Suppose that A, B are $m\times n$ matrices. Then one of the following must hold*:

(i) $\mathcal{N}(A)\cap\mathcal{N}(B)=\mathcal{N}(\lambda A+B)$ *for all but a finite number of* λ,

(ii) $\mathcal{N}(A)\cap\mathcal{N}(B)\subsetneqq\mathcal{N}(\lambda A+B)$ *for all* λ.

Proof. Note that $\mathcal{N}(A)\cap\mathcal{N}(B)\subseteq\mathcal{N}(\lambda A+B)$ for all λ. If (ii) holds, we are done. Suppose then that there exists a λ_0 such that $\mathcal{N}(A)\cap\mathcal{N}(B)=\mathcal{N}(\lambda_0 A+B)$. Let $r=\dim\mathcal{N}(A)\cap\mathcal{N}(B)$. Since $\operatorname{Rank}(\lambda_0 A+B)=n-r$, there exists a $(n-r)\times(n-r)$ submatrix with nonzero determinant. Thus the corresponding $(n-r)\times(n-r)$ submatrix of $\lambda A+B$ has nonzero determinant for all but a finite number of λ. Thus $\operatorname{Rank}(\lambda A+B)\geqq n-r$, or $\dim\mathcal{N}(\lambda A+B)\leqq r$, for all but a finite number of λ. Thus (i) follows. □

3. The $n\times n$ case.

THEOREM 1. *Suppose that $A, B\in\mathbb{C}^{n\times n}$ are such that $\mathcal{N}(A)\cap\mathcal{N}(B)$ reduces both A and B. Suppose also that there exists a λ such that $\mathcal{N}(\lambda A+B)=\mathcal{N}(A)\cap\mathcal{N}(B)$. Then $A\dot{x}+Bx=f$, f n-times continuously differentiable, is consistent if and only if $f(t)\in\mathcal{R}(\lambda A+B)$ for all t, that is, $(\lambda A+B)(\lambda A+B)^{\dagger}f=f$. If it is consistent, then all solutions are of the form*

$$
\begin{aligned}
x=\ & \hat{A}^{D}e^{-\hat{A}^{D}\hat{B}t}\int_0^t e^{\hat{A}^{D}\hat{B}s}\hat{f}(s)\,ds \\
& +[(\lambda A+B)^{D}(\lambda A+B)-\hat{A}\hat{A}^{D}]\sum_{m=0}^{k-1}(-1)^m[\hat{A}\hat{B}^{D}]^m\hat{B}^{D}\hat{f}^{(m)} \\
& +e^{-\hat{A}^{D}\hat{B}t}\hat{A}^{D}\hat{A}q+[I-(\lambda A+B)^{D}(\lambda A+B)]g
\end{aligned}
\tag{6}
$$

where $\hat{A}=(\lambda A+B)^{D}A$, $\hat{B}=(\lambda A+B)^{D}B$, $\hat{f}=(\lambda A+B)^{D}f$, q is an arbitrary vector, g an arbitrary vector valued function, and $k=\operatorname{Index}(\hat{A})$.

Proof. Since $\mathcal{N}(A)\cap\mathcal{N}(B)$ reduces both A and B, we have $A=A_1\oplus 0$, $B=B_1\oplus 0$ relative to $\mathbb{C}^n=[\mathcal{N}(A)\cap\mathcal{N}(B)]^{\perp}\oplus[\mathcal{N}(A)\cap\mathcal{N}(B)]$. But $(\lambda A+B)=(\lambda A_1+B_1)\oplus 0$, so that if $\mathcal{N}(\lambda A+B)=\mathcal{N}(A)\cap\mathcal{N}(B)$, we must have that λA_1+B_1 is invertible on $[\mathcal{N}(A)\cap\mathcal{N}(B)]^{\perp}$. Formula (6) now follows from [2, Thm. 7] and [2] guarantees a solution if $(\lambda A+B)(\lambda A+B)^{\dagger}f=f$. On the other hand, if $f=A\dot{x}+Bx$ for some x, then for all t, $f(t)\in\mathcal{R}(A)+\mathcal{R}(B)\subseteq(\mathcal{N}(A)\cap\mathcal{N}(B))^{\perp}=\mathcal{R}((\lambda A+B)(\lambda A+B)^{\dagger})$. Hence $(\lambda A+B)(\lambda A+B)^{\dagger}f=f$. □

Note that the first two terms of (6) are a particular solution of (1), and the last two are the general solution of the associated homogeneous equation. Theorem 1 is a generalization of [2, Thm. 7] since it includes it as a special case.

Since $\lambda A+B$ in Theorem 1 turns out to be EP, one has $(\lambda A+B)^{D}=(\lambda A+B)^{\#}=(\lambda A+B)^{\dagger}$ where $\#$ denotes the group inverse.

From Theorem 1 and Proposition 4 we have the following result.

THEOREM 2. *If A, B are* EP *and one is positive semidefinite, then there exists a λ such that $\mathcal{N}(\lambda A+B)=\mathcal{N}(A)\cap\mathcal{N}(B)$. Thus all solutions of* (1) *are in the form of* (6).

Theorem 2 is proved in [9] for the special case when A, B are both semidefinite. The expression in [9] for the solutions of (1) is less specific than (6).

Since the Drazin inverse is well-behaved with respect to similarity, and matrices of index 1 are similar to EP matrices, one might hypothesize that the assumptions of Theorem 2 can be weakened to assuming that A, B have index one. The next example shows that this is not the case. The difficulty is caused by the fact that multiplying (1) by a singular matrix will often produce a system not equivalent to the original one.

Example 2. Let

$$A=\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix},\qquad B=\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix},\qquad f=0.$$

Then A, B have index one, and $\mathcal{N}(A)\cap\mathcal{N}(B)=\mathcal{N}(\lambda A+B)$ for all λ. Note that (1) with this A, B is just $x_1+x_2=0$. We may take $\lambda=0$. Now $B^D=B$, $B^\dagger=\frac{1}{2}\begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}$. Multiplication of (1) by B^D gives $2B\dot{x}+Bx=0$ or $2\dot{x}_1+2\dot{x}_2+x_1+x_2=0$ which is not equivalent to the original system. Multiplication of (1) by $B^\dagger$ gives $\frac{1}{2}A\dot{x}+\frac{1}{2}Ax=0$ or $\dot{x}_1+\dot{x}_2+x_1+x_2=0$ which also is not equivalent to the original system.

In looking at systems with nonsquare coefficients, we shall be especially interested in two cases: when $(\lambda A+B)$ is one-to-one for some λ, and when $(\lambda A+B)$ is onto for some λ. The first says that solutions of (1) are unique when they exist; the second says that solutions always exist for at least one initial condition.

4. The case when $(\lambda A+B)$ is one-to-one. In this section we shall consider $A\dot{x}+Bx=f$ for the case when $(\lambda A+B)$ is one-to-one. This is equivalent to assuming that solutions, when they exist, are uniquely determined by f and the initial conditions.

THEOREM 3. *Suppose $(\lambda A+B)$ is one-to-one. Then all solutions of $A\dot{x}+Bx=0$ are of the form*

$$x=e^{-\hat{A}^D\hat{B}t}q\quad \textit{where } q\in\mathcal{R}(\hat{A}^D\hat{A})$$

and

$$(7)\qquad [I-(\lambda A+B)(\lambda A+B)^\dagger]A\hat{A}^D\{\hat{A}^{Dm}\hat{B}^m\}q=0\quad \textit{for } m=0,1,\cdots,n.$$

Here $\hat{A}=(\lambda A+B)^\dagger A$, $\hat{B}=(\lambda A+B)^\dagger B$.

Proof. If x is a solution of $A\dot{x}+Bx=0$, then x is a solution of $\hat{A}\dot{x}+\hat{B}x=0$. But $\hat{A}\hat{B}=\hat{B}\hat{A}$ and $\lambda\hat{A}+\hat{B}=I$. Hence $x=e^{-\hat{A}^D\hat{B}t}\hat{A}^D\hat{A}q$ [2]. Substituting back in gives $[-A\hat{A}^D\hat{B}\hat{A}^D\hat{A}+B\hat{A}]e^{-A^DBt}q=0$ for all t. Thus $[-A\hat{A}^D\hat{B}+B\hat{A}\hat{A}^D]e^{-\hat{A}^D\hat{B}t}q=0$ for all t, or equivalently, $[B\hat{A}-A\hat{B}]\hat{A}^D[\hat{A}^D\hat{B}]^mq=0$ for $m=0,1,2,\cdots$. But

$$\begin{aligned}A\hat{B}&=A(\lambda A+B)^\dagger B=A(\lambda A+B)^\dagger(\lambda A+B)-A(\lambda A+B)^\dagger\lambda A\\&=A-\lambda A(\lambda A+B)^\dagger A\\&=A-(\lambda A+B)(\lambda A+B)^\dagger A+B(\lambda A+B)^\dagger A\\&=[I-(\lambda A+B)(\lambda A+B)^\dagger]A+B\hat{A}.\quad\square\end{aligned}$$

COROLLARY 1. *If $\lambda A+B$ is one-to-one, and $\mathcal{N}(\bar{\lambda}A^*+B^*)=\mathcal{N}(A^*)\cap\mathcal{N}(B^*)$, then all solutions of $A\dot{x}+Bx=0$ are of the form $x=e^{-\hat{A}^D\hat{B}t}\hat{A}^D\hat{A}q$ where q is an arbitrary vector.*

Proof. $\mathcal{R}(\lambda A+B)^\perp=\mathcal{N}(\bar{\lambda}A^*+B^*)=\mathcal{N}(A^*)\cap\mathcal{N}(B^*)$. But $\mathcal{R}(A)\perp\mathcal{N}(A^*)$ so $\mathcal{R}(A)\subseteq\mathcal{R}(\lambda A+B)^\perp$. Thus (7) holds for all $q\in\mathcal{R}(AA^D)$. $\square$

Example 3. Let $A=\begin{bmatrix}1\\ 0\end{bmatrix}$, $B=\begin{bmatrix}0\\ 1\end{bmatrix}$. Then $(\lambda A+B)$ is one-to-one and

$\mathcal{N}(\lambda A+B)=\mathcal{N}(A)\cap\mathcal{N}(B)=\{0\}$ for all λ. However, $\mathcal{N}(\bar{\lambda}A^*+B^*)\neq \mathcal{N}(A^*)\cap\mathcal{N}(B^*)$ for all λ. $A\dot{x}+Bx=0$ has only $x=0$ as a solution. Multiplying by $(\lambda A+B)^\dagger=(|\lambda|^2+1)^{-1}[\lambda, 1]$ *we get*

$$\dot{\lambda}(|\lambda|^2+1)^{-1}\dot{x}+(|\lambda|^2+1)^{-1}m=0$$

which has the nonzero solutions $x=e^{-\lambda^{-1}t}q$.

THEOREM 4. *Suppose $(\lambda A+B)$ is one-to-one and $A\dot{x}+Bx=f$ is consistent. Then all solutions of $A\dot{x}+Bx=f$ are of the form*

$$x=e^{-\hat{A}^D\hat{B}t}\hat{A}^D\hat{A}q+\hat{A}^D e^{-\hat{A}^D\hat{B}t}\int_0^t e^{\hat{A}^D\hat{B}s}\hat{f}(s)\,ds$$
$$+(I-\hat{A}\hat{A}^D)\sum_{n=0}^{k-1}(-1)^n(\hat{A}\hat{B}^D)^n\hat{B}^D\hat{f}^{(n)} \tag{8}$$

where $\hat{A}=(\lambda A+B)^\dagger A$, $\hat{B}=(\lambda A+B)^\dagger B$, $k=\text{Index}(\hat{A})$, and $\hat{f}=(\lambda A+B)^\dagger f$.

Proof. If x solves $A\dot{x}+Bx=f$, then x solves $\hat{A}\dot{x}+\hat{B}x=\hat{f}$ and $\lambda\hat{A}+\hat{B}=I$. Thus (8) follows from [2, Thm. 7]. □

Theorem 4 is not as completely satisfying as our other results since we have not stated precisely for which f is $A\dot{x}+Bx=f$ consistent when $\lambda A+B$ is one-to-one. While the general problem appears difficult, we do have the following.

THEOREM 5. *Suppose $\lambda A+B$ is one-to-one and $\mathcal{N}(\lambda A^*+B^*)=\mathcal{N}(A^*)\cap\mathcal{N}(B^*)$. Then $A\dot{x}+Bx=f$ is consistent if and only if $(I-(\lambda A+B)(\lambda A+B)^\dagger)f=0$.*

Proof. Suppose $\lambda A+B$ is one-to-one and $\mathcal{N}(\bar{\lambda}A^*+B^*)=\mathcal{N}(A^*)\cap\mathcal{N}(B^*)$. Now $(\lambda A+B)(\lambda A+B)^\dagger$ is the identity on $\mathcal{R}(\lambda A+B)=\mathcal{N}(\bar{\lambda}A^*+B^*)^\perp=[\mathcal{N}(A^*)\cap\mathcal{N}(B^*)]^\perp\supseteq\mathcal{R}(A)\cup\mathcal{R}(B)$. Thus $(\lambda A+B)(\lambda A+B)^\dagger A=A$ and $(\lambda A+B)(\lambda A+B)^\dagger B=B$ Hence for any x, if we set $f=A\dot{x}+Bx$, we get $(\lambda A+B)(\lambda A+B)^\dagger f=f$. On the other hand, if $(\lambda A+B)(\lambda A+B)^\dagger f=f$, then $A\dot{x}+Bx=f$ is equivalent to $\hat{A}\dot{x}+\hat{B}x=\hat{f}$. Since $\hat{A}\dot{x}+\hat{B}x=\hat{f}$ is consistent from [2], so is $A\dot{x}+Bx=f$. □

The special cases when A or B are one-to-one are of some interest. As shall be pointed out in § 6, B being one-to-one is the case of most interest for the applications we have in mind.

THEOREM 6. *Suppose A is one-to-one. Then $A\dot{x}+Bx=f$ is consistent if and only if f is of the form*

$$f=AA^\dagger h\oplus(I-AA^\dagger)Bg \tag{9}$$

where h is an arbitrary function and

$$g=e^{-A^\dagger Bt}q+e^{-A^\dagger Bt}\int_a^t e^{-A^\dagger Bs}A^\dagger h(s)\,ds, \tag{10}$$

q an arbitrary constant. Conversely, if f has the form (9), *then g given in* (10) *is the general solution.*

Proof. Suppose A is one-to-one. Then $A\dot{x}+Bx=f$ is equivalent to the pair of equations:

$$\dot{x}+A^\dagger Bx=A^\dagger f, \tag{11}$$

and

$$(I-AA^{\dagger})Bx=(I-AA^{\dagger})f. \tag{12}$$

Now $AA^{\dagger}f$ can be chosen arbitrarily, say $AA^{\dagger}h$. Then (1) uniquely determines x giving (10). Substituting x into (12) gives $(I-AA^{\dagger})f$. □

A similar result is possible if B is one-to-one.

THEOREM 7. *Suppose B is one-to-one. Then $A\dot{x}+Bx=f$ is consistent if and only if f is of the form*

$$f=BB^{\dagger}h+(I-BB^{\dagger})A\dot{g} \tag{13}$$

where h is arbitrary and

$$\begin{aligned} g=&e^{-(B^{\dagger}A)^{D}t}(B^{\dagger}A)^{D}B^{\dagger}Aq+e^{-(B^{\dagger}A)^{D}t}(B^{\dagger}A)^{D}\int_{0}^{t}e^{(B^{\dagger}A)^{D}s}B^{\dagger}h(s)\,ds\\ &+[I-(B^{\dagger}A)^{D}(B^{\dagger}A)]\sum_{n=0}^{k-1}(-1)^{n}[B^{\dagger}A]^{n}B^{\dagger}h^{(n)}, \end{aligned} \tag{14}$$

$k=\operatorname{Index} B^{\dagger}A$, q arbitrary. Conversely, if f has the form (13), *then g in* (14) *is the general solution.*

Proof. Suppose B is one-to-one Then $A\dot{x}+Bx=f$ is equivalent to

$$B^{\dagger}A\dot{x}+x=B^{\dagger}f, \tag{15}$$

and

$$(I-BB^{\dagger})A\dot{x}=(I-BB^{\dagger})f. \tag{16}$$

Again $BB^{\dagger}f$ is arbitrary. From (15), x is determined uniquely in terms of $B^{\dagger}f$. Then $(I-BB^{\dagger})f$ must follow from (16). □

5. The case when $(\lambda A+B)$ is onto. To assume that $\lambda A+B$ is onto is the same as assuming that (1) is consistent for all sufficiently smooth f. Solutions will not, in general, be uniquely determined by initial conditions. We shall first solve (1) and then summarize our results.

Let A, B be $m\times n$ matrices. Let λ be such that $\lambda A+B$ is onto. Define $P=(\lambda A+B)^{\dagger}(\lambda A+B)$. Then (1) becomes

$$AP\dot{x}+BPx=f-A(I-P)\dot{x}-B(I-P)x.$$

Or, equivalently,

$$A(\lambda A+B)^{\dagger}[(\lambda A+B)x]+B(\lambda A+B)^{\dagger}[(\lambda A+B)x]=f-A(I-P)\dot{x}-B(I-P)x. \tag{17}$$

But $\lambda[A(\lambda A+B)^{\dagger}]+[B(\lambda A+B^{\dagger})]=I$. Thus (17) is, in terms of $(\lambda A+B)x$, a differential equation of the type solved in [2] and hence has a solution for any choice of $(I-P)x$. Note that $[\lambda[A(\lambda A+B)^{\dagger}]+[B(\lambda A+B)^{\dagger}]]^{-1}=I$. Thus [2, Thm. 7] gives us

THEOREM 8. *Suppose that $\lambda A+B$ is onto and f is n-times differentiable. Let $\hat{A}=A(\lambda A+B)^{\dagger}$, $\hat{B}(\lambda A+B)^{\dagger}$. Let $g=f-A[I-(\lambda A+B)^{\dagger}(\lambda A+B)]\dot{h}-B[I-(\lambda A+B)^{\dagger}(\lambda A+B)]h$ where h is an arbitrary $(n+1)$-times differentiable vector*

valued function. Then all solutions of $A\dot{x}+Bx=f$ *are of the form*

$$x=(\lambda A+B)^{\dagger}\Big\{e^{-\hat{A}^{D}\hat{B}t}\hat{A}\hat{A}^{D}q+\hat{A}^{D}e^{-\hat{A}^{D}\hat{B}t}\int_{0}^{t}e^{\hat{A}^{D}\hat{B}s}g(s)\,ds$$
$$+(I-\hat{A}^{D}\hat{A})\sum_{n=0}^{k-1}(-1)^{n}[\hat{A}\hat{B}^{D}]^{n}\hat{B}^{D}g^{(n)}\Big\}$$
$$+[I-(\lambda A+B)^{\dagger}(\lambda A+B)]h,$$

q an arbitrary constant vector, $k=\operatorname{Index}\hat{A}$.

The formulas in Theorem 8 simplify considerably if A or B are onto. For the applications we shall discuss in the last section, the case when B is onto is the more important.

THEOREM 9. *Suppose that B is onto. Then all solutions of* $A\dot{x}+Bx=f$ *are of the form*:

$$x=B^{\dagger}\Big\{e^{-\hat{C}^{D}t}\hat{C}\hat{C}^{D}q+\hat{C}^{D}e^{-\hat{C}^{D}t}\int_{0}^{t}e^{\hat{C}^{D}s}g(s)\,ds+(I-\hat{C}^{D}\hat{C})\sum_{n=0}^{k-1}(-1)^{n}\hat{C}^{n}g^{(n)}\Big\}$$
$$+[I-B^{\dagger}B]h,$$

h an arbitrary function, q an arbitrary vector, $g=f-A[I-B^{\dagger}B]\dot{h}$, $\hat{C}=AB^{\dagger}$, $k=\operatorname{Index}(C)$.

Theorem 9 comes immediately from Theorem 8 by setting $\lambda=0$ and noting that $\hat{B}=I$.

THEOREM 10. *Suppose that A is onto. Then all solutions of* $A\dot{x}+Bx=f$ *are of the form*

$$x=A^{\dagger}\Big\{e^{-BA^{\dagger}t}q+e^{-BA^{\dagger}t}\int_{0}^{t}e^{BA^{\dagger}s}g(s)\,ds\Big\}+[I-A^{\dagger}A]h \tag{18}$$

where h is an arbitrary function and $g=f-B[I-A_{A}^{\dagger}A]\dot{h}$.

Proof. This one is easier to prove directly. Suppose A is onto and rewrite $A\dot{x}+Bx=f$ as

$$(A\dot{x})+BA^{\dagger}(Ax)=f-B[I-A^{\dagger}A]x. \tag{19}$$

Taking $[I-A^{\dagger}A]x$ arbitrarily we can solve (19) uniquely for Ax, $A^{\dagger}Ax=x$, to get (18). □

In Theorems 8, 9 and 10 we have used without stating the basic fact that $[I-C^{\dagger}C]$ is the orthogonal projection onto $\mathcal{N}(C)$.

6. Applications. In [3], the results of [2] were used to solve a linear autonomous control process with quadratic cost funtional. The key step in [3] was solving explicitly a differential equation of the form

$$A\dot{z}+Bz=0 \tag{20}$$

where

$$A=\begin{bmatrix}I & 0\\ 0 & 0\end{bmatrix}\quad\text{and}\quad B=\begin{bmatrix}B_{1} & B_{2}\\ B_{3} & B_{4}\end{bmatrix}.$$

In [3], the matrices A, B turned out to always be $n \times n$. The results of § 3 can be sometimes used to solve (20).

A different problem that leads to the same type of differential equation is:

$$\dot{x} = Cx + Du, \qquad g = Ex + Fu. \tag{21}$$

Here g is a given output and the problem is to determine x and/or u. Rewrite (21) as

$$\mathcal{A}\dot{z} - \mathcal{B}z = f \tag{22}$$

where

$$\mathcal{A} = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} C & D \\ E & F \end{bmatrix}, \quad z = \begin{bmatrix} x \\ u \end{bmatrix}, \quad f = \begin{bmatrix} o \\ g \end{bmatrix}.$$

One frequently does not want to assume in (22) that $\mathcal{A}$, $\mathcal{B}$ are square. However, it may not be difficult to argue from physical grounds that $\lambda\mathcal{A} + \mathcal{B}$ is one-to-one or $\lambda\mathcal{A} + \mathcal{B}$ is onto. That is, one may know either that (22) is always consistent, or solutions are always unique. In this case we may use §§ 4 or 5 and the approach of [3] to solve (22) explicitly. As in [3] this will lead to feedback controls.

Note that the process in (21) can be made affine., $\dot{x} = Cx + Du + b(t)$, without significantly altering (22).

The control problem (21) has been analyzed using generalized inverses in [5] and in more generality in [6]. However, in [5], [6] conditions are placed on $\mathcal{A}$, $\mathcal{B}$ to assure that (21) is consistent for all g. While this simplifies the mathematics, it may or may not be a reasonable assumption to make as regards a particular problem. That a given plant is incapable of producing an arbitrary output is not unrealistic.

Our approach is able to handle some cases that [5], [6] cannot. However, the problem in [6] is such that while the differential system is in the form of (22) and approachable by the methods of this paper and [2], the matrix $\mathcal{A}$ is not of the form $\begin{bmatrix} A_1 & 0 \\ A_2 & 0 \end{bmatrix}$. Hence the results of [8] cannot always be used and the results of [8] were instrumental in simplifying the computations in [3].

REFERENCES

[1] S. L. CAMPBELL AND C. D. MEYER, JR., *EP operators and generalized inverses*, Canad. Math. Bull, 18 (1975), pp. 327–333.

[2] S. L. CAMPBELL, C. D. MEYER, JR. AND N. J. ROSE, *Applications of the Drazin inverse to linear systems of differential equations*, SIAM J. Appl. Math., 31 (1976), pp. 411–425.

[3] S. L. CAMPBELL, *Optimal control of autonomous linear processes with singular matrices in the quadratic cost functional*, SIAM J. Control, 14 (1976), pp. 1092–1106.

[4] F. R. GANTMACHER, *The Theory of Matrices*, vol. II, Chelsea, New York, 1964.

[5] V. LOVASS-NAGY AND D. L. POWERS, *On output feedback control*, Internat. J. Control, 21 (1975), pp. 1025–1028.

[6] ———, *Matrix generalized inverses in the handling of control problems containing input derivatives*, Internat. J. Systems Sci., 6 (1975), pp. 693–696.

[7] ———, *On rectangular systems of differential equations and their application to circuit theory*, J. Franklin Inst., 299 (1975), pp. 399–407.

[8] C. D. MEYER, JR. AND N. J. ROSE, *The index and the Drazin inverse of block triangular matrices*, SIAM J. Appl. Math., to appear.

[9] K. T. WONG, *The eigenvalue problem $\lambda Tx + Sx$*, J. Differential Equations, 16 (1974), pp. 270–280.

SIAM J. MATH. ANAL.
Vol. 8, No. 6, November 1977

ON THE VALIDITY OF THE TWO-TIMING METHOD FOR LARGE TIMES*

JAMES P. KEENER†

Abstract. In this paper we investigate the validity of the two-timing method when used to approximate the solution of an initial value problem tending to a limit cycle. It is shown that the two-timing method leads to an orbitally valid asymptotic approximation of the stable limit cycles, but that in certain situations, the two-timing method predicts an incorrect domain of attraction. As such, for certain initial values, two-timing predicts entirely incorrect long time behavior.

1. Introduction. The attempt to find asymptotic solutions of initial value problems which are valid for "long times" has a long history, beginning with the work of Poincaré, Linstedt and Van der Pol [9]. Currently, there are two methods of finding approximate solutions which are in wide use: The method of averaging, developed by Kryloff and Bogoliubov [1], and the method of multiple scales (also called two-timing), developed by Kevorkian and Cole [4]. Despite the fact that solutions generated by these methods appear to satisfy the equations uniformly for all time, in general it is known that the approximate solution is a valid approximation only for times of order $O(1/\varepsilon)$ as ε, the expansion parameter, approaches zero [8]. Recently, Greenlee and Snow [10] have taken advantage of special properties of damped second order ordinary differential equations to show that the approximate solutions found by two-timing are pointwise valid on the entire half line $t \in (0, \infty)$.

When the two-timing method is applied to problems with limit cycles, such as the Van der Pol oscillator, one finds that the approximate solution approaches a limit cycle as t approaches infinity [4]. However, there is no guarantee that this is indeed the correct long time behavior of the solution since the approximate solution is pointwise valid only for times of order $O(1/\varepsilon)$.

In this paper we investigate the validity of two-timing in the calculation of limit cycles. In particular, we will show that, even though the approximate solution is not pointwise valid for all times, it is orbitally valid for large times, in that the approximate solution approaches a valid approximation of a stable limit cycle of the differential equation. Two-timing, then, provides an indirect existence proof of stable limit cycles. However, we will also show that, for given initial data, the method of two-timing may pick the wrong limit cycle as its asymptotic limit.

We consider the system of equations

$$\frac{dY}{dt} = PY + \varepsilon F(Y, \varepsilon) \tag{1.1}$$

where Y is a k-vector, $k \geqq 2$, P is a $k \times k$ constant real matrix, and $F(Y, \varepsilon)$ a nonlinear vector function. The equation (1.1) is motivated by problems of oscillators in chemical systems [2], [3] in which the parameter ε is some physical parameter which can be externally adjusted. The matrix P is assumed to have a

* Received by the editors December 11, 1975, and in revised form June 11, 1976.

† Department of Mathematics, University of Arizona, Tucson, Arizona 85721. This research was supported in part by the National Science Foundation under Grant MPS 75-07621.

pair of imaginary eigenvalues and $k-2$ distinct eigenvalues with negative real parts. Then, by means of a standard change of variables, we can assume that

$$P=\begin{pmatrix} 0 & 1 & \\ -1 & 0 & 0 \\ & 0 & -\Lambda \end{pmatrix} \tag{1.2}$$

where the matrix Λ is a diagonal matrix whose entries have positive real parts. The nonlinear function $F(Y,\varepsilon)=(f_i(Y,\varepsilon))$ is assumed to be a polynomial function in Y and ε satisfying

$$f_i(0,\varepsilon)=0, \qquad i=1,2,\cdots,k, \tag{1.3a}$$

$$f_i(Y,\varepsilon)=p_i(y_1,y_2,\varepsilon)+q_i(Y,\varepsilon), \qquad i=1,2, \tag{1.3b}$$

where $p_i(y_1,y_2,\varepsilon)$ are polynomials in y_1 and y_2 with at least one of p_1 or p_2 containing terms of the form $y_1^r y_2^s$ with $r+s$ positive and odd, and where

$$q_i(y_1,y_2,0,\cdots,0;\varepsilon)=0, \qquad i=1,2. \tag{1.3c}$$

The remainder of this paper is organized as follows: § 2 is devoted to an examination of the existence of periodic solutions of (1.1) and in § 3, the linearized stability of these periodic solutions is calculated using Floquet exponents. In § 4 we discuss the two-timing method and its validity in describing limit cycle behavior for initial value problems. Finally in § 5, we show a situation in which the two-timing method predicts an entirely incorrect asymptotic limit cycle behavior.

In the development that follows, it is convenient to refer to iteration schemes rather than perturbation series. This should cause no problem for the practitioner who prefers to think of perturbation series since, in fact, there is no difference in the validity of the two approaches, as will be shown in Lemma 1.4.

DEFINITION 1.1. Suppose the operators $K: B_1 \to B_2$ and $N(u,\varepsilon): B_1 \times \mathbb{R} \to B_2$ are linear invertible and nonlinear, continuous in ε, respectively, on the Banach spaces B_1, B_2. Then

$$\begin{aligned} &U_0 \in B_1, \\ &KU_{n+1}=\varepsilon N(U_n;\varepsilon), \qquad n=0,1,2,\cdots, \end{aligned} \tag{1.4a}$$

defines an iterative scheme for the sequence $\{U_n\}$ of approximate solutions of

$$Ku=\varepsilon N(u;\varepsilon). \tag{1.4b}$$

DEFINITION 1.2. With K and N as in Definition 1.1, suppose that for any sequence $\{u_n\}$ there are operators $N_0, N_1(u_1), N_2(u_1,u_2),\cdots,N_k(u_1,u_2,\cdots,u_k)$ such that

$$\begin{aligned} N(\tilde{U}_k;\varepsilon)=&N_0+\varepsilon N_1(u_1)+\varepsilon^2 N_2(u_1,u_2)+\cdots \\ &+\varepsilon^k N_k(u_1,u_1,\cdots,u_k)+O(\varepsilon^{k+1}) \end{aligned} \tag{1.5}$$

$$\tilde{U}_k=\sum_{j=1}^{k}\varepsilon^j u_j,$$

then

$$
\begin{aligned}
&Ku_{j+1} = N_j(u_1, u_2, \cdots, u_j), \qquad j = 0, 1, 2, \cdots \\
&\tilde{U}_k = \sum_{j=1}^{k} \varepsilon^j u_j
\end{aligned} \tag{1.6}
$$

is said to define a *perturbation series solution* for the equation

$$Ku = \varepsilon N(u; \varepsilon). \tag{1.4b}$$

The convergence of the iteration procedure of Definition 1.1 is easily proven using the contraction mapping principle, and the convergence of the perturbation series is found indirectly by comparison with the iteration scheme.

LEMMA 1.3. *Suppose the operator K and N of Definition* 1.1 *also satisfy*

$$\|K^{-1}\| \leqq k_1 \tag{1.7i}$$

$$
\begin{aligned}
&\|N(u_1; \varepsilon)\| \leqq M_1, \\
&\|N(u_1; \varepsilon) - N(u_2; \varepsilon)\| \leqq M_2 \|u_1 - u_2\| \\
&\qquad\qquad \textit{whenever } \|u_i\| \leqq m, \quad |\varepsilon| \leqq 1.
\end{aligned} \tag{1.7ii}
$$

Then for each ε *satisfying*

$$0 \leqq |\varepsilon| \leqq \varepsilon_0, \ \varepsilon_0 = \min\left\{\frac{m}{k_0 M_1}, \frac{1}{k_0 M_2}, 1\right\},$$

the sequence $\{U_n\}$ *generated by the iteration scheme* (1.4) *converges to the unique solution* $u(\varepsilon)$ *of* $Ku = \varepsilon N(u; \varepsilon)$ *with* $\|u\| \leqq m$. *Furthermore, if* $U_0 = 0$, *then*

$$\|U_n - u(\varepsilon)\| \leqq O(\varepsilon^{n+1}).$$

LEMMA 1.4. *If the iteration scheme* (1.4) *has* $U_0 = 0$, *then the perturbation series* $\{\tilde{U}_n\}$ *of* (1.6) *satisfies*

$$\|\tilde{U}_n - U_n\| \leqq O(\varepsilon^{n+1}).$$

If, in addition, Lemma 1.3 *holds, then by the triangle inequality*

$$\|\tilde{U}_n - u(\varepsilon)\| \leqq O(\varepsilon^{n+1}).$$

The proof of this lemma follows by induction from (1.5), (1.7ii) and the fact that

$$
\begin{aligned}
K(\tilde{U}_{k+1} - U_{k+1}) &= \varepsilon(N_0 + \varepsilon N_1(u_1) + \cdots + \varepsilon^k N_k(u_1, u_2, \cdots, u_k) - N(U_k; \varepsilon)) \\
&= \varepsilon(N(\tilde{U}_k; \varepsilon) - N(U_k; \varepsilon) + O(\varepsilon^{k+1})).
\end{aligned}
$$

Actually, the validity of an approximate solution can be determined with no knowledge of how the solution was generated, provided we know that a solution does exist.

LEMMA 1.5. *Suppose the problem*

$$Ku = \varepsilon N(u; \varepsilon) \tag{1.4b}$$

has a solution $u(\varepsilon)$ *for all* ε *satisfying* $0 \leqq |\varepsilon| \leqq \varepsilon_0$, *and that for* $|\varepsilon| \leqq \varepsilon_0$, $v(\varepsilon)$ *satisfies*

$\|v\| \leqq m$ *and*

$$Kv - \varepsilon N(v, \varepsilon) = r(\varepsilon). \tag{1.8}$$

Then

$$\|u(\varepsilon) - v(\varepsilon)\| \leqq \frac{k_0}{1 - |\varepsilon| k_0 m_2} \|r(\varepsilon)\|$$

so that for

$$|\varepsilon| \leqq \min\left\{\varepsilon_0, \frac{1}{2k_0 m_2}\right\},$$

$$\|u(\varepsilon) - v(\varepsilon)\| \leqq \frac{k_0}{2} \|r(\varepsilon)\|. \tag{1.9}$$

The proof of this lemma follows from the fact that

$$u - v = K^{-1} r(\varepsilon) + \varepsilon K^{-1}(N(u, \varepsilon) - N(v, \varepsilon)).$$

Using (1.7) we find that

$$\|u - v\| \leqq k_0 \|r(\varepsilon)\| + |\varepsilon| k_0 m_2 \|u - v\|$$

from which (1.8) follows immediately.

Practically speaking, then, a function which satisfies equation (1.4b) to a certain order in ε also approximates the actual solution to the same order.

In the proofs that follow, it is sufficient to rewrite the problem (1.1) in terms of operators K and N. The properties (1.5) and (1.7ii) of the nonlinear map N will often follow directly from the assumption that $F(Y, \varepsilon)$ in (1.1) is a polynomial in Y and ε. Thus, most of what follows is the attempt to find the appropriate linear, invertible operator K, followed by a reference to one of preceding lemmas.

2. Periodic solutions. We begin by studying the periodic solutions, if any, of (1.1). Because of the assumed structure of P, periodic solutions of period 2π exist for $\varepsilon = 0$ and have arbitrary amplitude. We expect a slightly different period for $\varepsilon \neq 0$, and introduce the change of independent variable $t^* = wt$, $w = w(\varepsilon)$, so that (1.1) becomes

$$LY \equiv \frac{dY}{dt^*} - PY = \varepsilon F(Y, \varepsilon) + (1 - w)\frac{dY}{dt^*} \tag{2.1}$$

$$Y(0) = Y(2\pi), \qquad w(0) = 1.$$

The operator L defined in (2.1) with periodic boundary conditions has a two dimensional null space spanned by the linearly independent vectors

$$\phi_1(t^*) = \begin{pmatrix} \sin t^* \\ \cos t^* \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \phi_2(t^*) = \begin{pmatrix} \cos t^* \\ -\sin t^* \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{2.2}$$

The vectors $\phi_1(t^*)$, $\phi_2(t^*)$ also span the null space of the adjoint of L. Because of the structure of the null space, we can find solutions of (2.1) by applying the results of bifurcation theory at multiple eigenvalues [6]. Specifically, we expect solutions of (2.1) to be of the form

$$\begin{aligned} &Y(t^*, \varepsilon) = a(\varepsilon)\phi_1(t^*) + b(\varepsilon)\phi_2(t^*) + \varepsilon U(t^*, \varepsilon) \\ &\langle \phi_i, U \rangle = 0, \qquad i = 1, 2, \end{aligned} \tag{2.3}$$

where the inner product in (2.3) is defined by

$$\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{2\pi} \int_0^{2\pi} \mathbf{u}(\tau) \cdot \mathbf{v}(\tau)\, d\tau. \tag{2.4}$$

Furthermore, solutions of (2.1) exist if and only if

$$\left\langle \phi_i, F(Y, \varepsilon) + \left(\frac{1-w}{\varepsilon}\right)\frac{dY}{dt^*} \right\rangle = 0, \qquad i = 1, 2. \tag{2.5}$$

One immediate consequence of (2.5) is

LEMMA 2.1. *A necessary condition for the existence of periodic solutions of* (2.1) *is that there exist numbers* α_0, β_0, W_1 *satisfying the nonlinear eigenvalue problem*

$$\begin{aligned} &\langle \phi_1, F(Y_0, 0) \rangle + W_1\beta_0 = 0, \\ &\langle \phi_2, F(Y_0, 0) \rangle - W_1\alpha_0 = 0, \\ &Y_0 = \alpha_0\phi_1(t^*) + \beta_0\phi_2(t^*). \end{aligned} \tag{2.6}$$

The proof of this lemma follows by letting ε approach zero in (2.5) and defining $W_1 = \lim_{\varepsilon \to 0} (w(\varepsilon) - 1)/\varepsilon$. If (2.6) has a nontrivial solution, α_0, β_0, W_1, then these provide a first approximation to the quantities $a(\varepsilon)$, $b(\varepsilon)$, $w(\varepsilon)$. It is then possible to rewrite (2.5) as an equation for perturbation quantities α_1, β_1, W_2 of the form

$$A\begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} = W_2\begin{pmatrix} -\beta_0 \\ \alpha_0 \end{pmatrix} + \begin{pmatrix} R_1 \\ R_2 \end{pmatrix} \tag{2.7a}$$

where $a(\varepsilon) = \alpha_0 + \varepsilon\alpha_1$, $b(\varepsilon) = \beta_0 + \varepsilon\beta_1$, $w(\varepsilon) = 1 + \varepsilon W_1 + \varepsilon^2 W_2$ and

$$\begin{aligned} &A = \begin{pmatrix} \langle \phi_1, F_y(Y_0, 0)\phi_1 \rangle & \langle \phi_1, F_y(Y_0, 0)\phi_2 \rangle + W_1 \\ \langle \phi_2, F_y(Y_0, 0)\phi_1 \rangle - W_1 & \langle \phi_2, F_y(Y_0, 0)\phi_2 \rangle \end{pmatrix} \\ &y_0 = \alpha_0\phi_1(t^*) + \beta_0\phi_2(t^*), \end{aligned} \tag{2.7b}$$

and R_1, R_2 represent nonlinear remainder terms.

The matrix A is always a singular matrix. Since the system (2.1) is an autonomous system, the shift $t^* \to t^* + \delta$ applied to any known solution yields a new solution. Written in the form (2.3) this variable shift gives new values for $a(\varepsilon)$ and $b(\varepsilon)$ with $a^2(\varepsilon) + b^2(\varepsilon)$ unchanged. By allowing δ to change freely we generate a family of solutions of (2.1) and therefore of (2.6) also. Thus, finding the solutions of (2.6) is always equivalent to finding the roots of an equation of the form $G(\rho, W_1) = 0$ where $\rho = \alpha_0^2 + \beta_0^2$. It follows that any vector $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ satisfying $u_1\alpha_0 + u_2\beta_0 = 0$ belongs to the null space of A.

LEMMA 2.2. *Suppose that the matrices A and A^T have one dimensional null spaces spanned by*

$$\Gamma = \begin{pmatrix} \beta_0 \\ -\alpha_0 \end{pmatrix}, \qquad \Gamma^* = \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} \tag{2.8}$$

respectively, and that α_0, β_0, W_1, satisfy (2.6). *A sufficient condition for the existence of periodic solutions of* (2.1) *is that*

$$\beta_0\gamma_1 - \alpha_0\gamma_2 \equiv \Gamma^T\Gamma^* \neq 0. \tag{2.9}$$

To solve (2.7a) with known R_1, R_2, one must be able to choose W_2 so that the right hand side of (2.7a) is orthogonal to Γ^*, that is

$$W_2(\beta_0\gamma_1 - \alpha_0\gamma_2) = R_1\gamma_1 + R_2\gamma_2. \tag{2.10}$$

Clearly this can be accomplished if (2.9) holds. There is a certain arbitrariness in the solution of (2.7) since one can add to $\begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix}$ arbitrary amounts of $\begin{pmatrix} \beta_0 \\ -\alpha_0 \end{pmatrix}$. The solution can be made unique by requiring, for example, $\alpha_1\beta_0 - \alpha_0\beta_1 = 0$. Such a normalization is clearly necessary since the system (2.1) is autonomous, and the general solution of (2.1) is a two parameter surface of solutions generated by a phase shift δ and parameter ε.

Since $L\phi_i(t^*) = 0$, the equation (2.1) can now be viewed as an equation for the unknown function $\varepsilon U(t^*, \varepsilon)$. In fact, with the above details in mind, we see that equations (2.1), (2.3), (2.7) along with the orthogonality condition (2.10) and the normalization $\alpha_1\beta_0 - \beta_1\alpha_0 = 0$ define an equation of the form (1.4b) for the unknowns $\varepsilon U(t^*, \varepsilon)$, $\varepsilon\alpha_1(\varepsilon)$, $\varepsilon\beta_1(\varepsilon)$, $\varepsilon W_2(\varepsilon)$, provided $U(t^*, \varepsilon)$ is in the Banach space of periodic $C^1[0, 2\pi]$ functions. The observation that (1.7) also holds concludes the proof of Lemma 2.2.

When the nonlinear vector function $F(Y, \varepsilon)$ is a polynomial satisfying (1.3), the inner products in (2.6), (2.7) can be simplified in the following manner. According to (2.3), $Y_0 = a\phi_1 + b\phi_2$ when $\varepsilon = 0$, so that

$$\begin{aligned} y_1 &= a \sin t^* + b \cos t^*, \\ y_2 &= a \cos t^* - b \sin t^* \\ y_i &= 0, \; i = 3, 4, \cdots, k. \end{aligned} \tag{2.11a}$$

Then a polynomial term $y_1^p y_2^q$, $p \geqq 0$, $q \geqq 0$ becomes

$$y_1^p y_2^q = \begin{cases} \text{even harmonics} & \text{if } p+q \text{ even} \\ \alpha_{pq}(a^2+b^2)^{\frac{p+q-1}{2}} y_2 + \text{higher harmonics} & \text{if } p+q \text{ odd, } p \text{ even} \\ \alpha_{pq}(a^2+b^2)^{\frac{p+q-1}{2}} y_2 + \text{higher harmonics} & \text{if } p+q \text{ odd, } p \text{ even} \end{cases} \tag{2.11b}$$

where

$$\alpha_{pq} = \begin{cases} \sum_{j=0}^{q/2} \frac{(-1)^j}{2^{2j+p-1}} \binom{q/2}{j} \left\langle \frac{2j+p}{\frac{2j+p-1}{2}} \right\rangle, & p+q \text{ odd}, q \text{ even}, \\ \sum_{j=0}^{p/2} \frac{(-1)^j}{2^{2j+q-1}} \binom{p/2}{j} \left\langle \frac{2j+q}{\frac{2j+q-1}{2}} \right\rangle, & p+q \text{ odd}, p \text{ even}, \\ 0, & p+q \text{ even}. \end{cases} \tag{2.11c}$$

(For a tabulation of α_{pq} see Appendix C.)

Suppose that the polynomials $p_i(y_1, y_2, \varepsilon)$ satisfy

$$p_i(y_1, y_2, 0) = \sum_{p,q} \beta_{pq}^i y_1^p y_2^q. \tag{2.12a}$$

Then one can easily see that

$$f_i(Y_0, 0) = g_{i1} y_1 + g_{i2} y_2 + \text{higher harmonics}, \qquad i = 1, 2 \tag{2.12b}$$

where

$$\begin{aligned} g_{i1} &= \sum_{\substack{p+q \text{ odd} \\ q \text{ even}}} \beta_{pq}^i \alpha_{pq} (a^2+b^2)^{(p+q-1)/2}, \\ g_{i2} &= \sum_{\substack{p+q \text{ odd} \\ p \text{ even}}} \beta_{pq}^i \alpha_{pq} (a^2+b^2)^{(p+q-1)/2}. \end{aligned} \tag{2.12c}$$

Finally we can calculate

$$\begin{aligned} \langle \phi_1, F(Y_0, 0) \rangle &= \tfrac{1}{2}(g_{11}+g_{22})a + \tfrac{1}{2}(g_{21}-g_{12})b, \\ \langle \phi_2, F(Y_0, 0) \rangle &= \tfrac{1}{2}(g_{12}-g_{21})a + \tfrac{1}{2}(g_{11}+g_{22})b \end{aligned} \tag{2.13a}$$

and the system (2.6) reduces to

$$\begin{aligned} \tfrac{1}{2}(g_{11}+g_{22})a + [\tfrac{1}{2}(g_{21}-g_{12}) + W_1]b &= 0, \\ [\tfrac{1}{2}(g_{12}-g_{21}) - W_1]a + \tfrac{1}{2}(g_{11}+g_{22})b &= 0 \end{aligned} \tag{2.13b}$$

where the functions g_{ij} are polynomial functions of $\rho = a^2+b^2$. The determinant of the system (2.13) is

$$\tfrac{1}{4}(g_{11}+g_{22})^2 + (\tfrac{1}{2}(g_{21}-g_{12}) + W_1)^2 = 0. \tag{2.13c}$$

Clearly Lemma 2.1 is satisfied whenever α_0, β_0, W_1 are chosen so that $\alpha_0^2+\beta_0^2 = \rho$ is a root of the polynomial equation

$$\begin{aligned} f(\rho) &\equiv g_{11}(\rho) + g_{22}(\rho) = 0, \\ W_1 &= \tfrac{1}{2} g(\rho) \equiv \tfrac{1}{2}(g_{12}(\rho) - g_{21}(\rho)). \end{aligned} \tag{2.14}$$

To examine more closely the condition (2.9) when $F(Y, \varepsilon)$ is a polynomial vector function, notice that we can reduce the matrix A in (2.7), using (2.13) and

(2.14), to

$$A = \begin{pmatrix} \alpha_0^2 f'(\rho_0) + \alpha_0\beta_0 g'(\rho_0) & \alpha_0\beta_0 f'(\rho_0) + \beta_0^2 g'(\rho_0) \\ \alpha_0\beta_0 f'(\rho_0) - \alpha_0^2 g'(\rho_0) & \beta_0^2 f'(\rho_0) - \alpha_0\beta_0 g'(\rho_0) \end{pmatrix}. \tag{2.15}$$

For $\rho_0 \neq 0$ and $f'^2(\rho_0) + g'^2(\rho_0) \neq 0$, the null space of A is spanned by

$$\Gamma = \begin{pmatrix} \beta_0 \\ -\alpha_0 \end{pmatrix} \tag{2.16a}$$

and the null space of A^T is spanned by

$$\Gamma^* = \begin{pmatrix} g'(\rho_0)\alpha_0 - f'(\rho_0)\beta_0 \\ f'(\rho_0)\alpha_0 + g'(\rho_0)\beta_0 \end{pmatrix}. \tag{2.16b}$$

Then the sufficient condition (2.9) reduces to

$$\rho_0 f'(\rho_0) \neq 0. \tag{2.17}$$

This leads to the following application of Lemma 1.3.

THEOREM 2.3. *Suppose that the polynomial function $F(Y, \varepsilon)$ satisfies* (1.3) *and that the polynomials $f(\rho) \not\equiv 0$, $g(\rho)$ are defined by* (2.11), (2.12) *and* (2.14). *For each nontrivial root $\rho_0 \neq 0$ of $f(\rho) = 0$ satisfying $f'(\rho_0) \neq 0$, there are positive constants m_0, m_1, ε_0 such that for each ε in $0 \leqq |\varepsilon| < \varepsilon_0$, the problem* (1.1) *has nontrivial periodic solutions of the form*

$$\begin{aligned} &Y(t^*, \varepsilon) = a(\varepsilon)\phi_1(t^*) + b(\varepsilon)\phi_2(t^*) + \varepsilon U(t^*, \varepsilon), \qquad \|U\| \leqq m_1, \\ &w(\varepsilon) = 1 + \varepsilon W_1(\varepsilon), \quad |W_1| \leqq m_2, \quad \langle \phi_i, U \rangle = 0, \qquad i = 1, 2 \\ &t^* = w(\varepsilon)t \end{aligned} \tag{2.18}$$

where $a^2(0) + b^2(0) = \rho_0$ and $W_1(0) = \frac{1}{2}g(\rho_0)$.

For future reference, we state the following application of Lemma 1.5.

COROLLARY 2.4. *Suppose the hypotheses of Theorem* 2.3 *hold and suppose the functions $z(t^*, \varepsilon)$, $\theta(\varepsilon)$ are bounded and continuous in ε for $|\varepsilon| \leqq \varepsilon_0$ and satisfy*

$$\begin{aligned} &Lz - \varepsilon F(z, \varepsilon) + (\theta - 1)\frac{dz}{dt^*} = r(t^*, \varepsilon), \\ &z(0) = z(2\pi), \qquad \|r(t^*, \varepsilon)\| \leqq R(\varepsilon) \end{aligned} \tag{2.19}$$

for $|\varepsilon| \leqq \varepsilon_0$, $t^* \in [0, 2\pi]$, where $R(\varepsilon) = o(\varepsilon)$. Then there are positive constants K_1, K_2, ε_1 and a periodic solution $u(t^*, \varepsilon)$, $w(\varepsilon)$ of (1.1) *satisfying*

$$\|z(t^*, \varepsilon) - u(t^*, \varepsilon)\| \leqq K_1 R(\varepsilon), \qquad |\theta(\varepsilon) - w(\varepsilon)| \leqq K_2 R(\varepsilon) \tag{2.20}$$

for all $|\varepsilon| \leqq \varepsilon_1 \leqq \varepsilon_0$.

This simply states that functions $z(t^*, \varepsilon)$, $\theta(\varepsilon)$ which approximately satisfy the equation (1.1) approximate a solution of (1.1) to the same order of approximation. At first glance it may appear that this corollary follows from Lemma 1.5 with no further comment. Such is not the case. In fact, since (1.1) may have more than one solution, and since the representation of each solution is not unique, we must first determine the function $u(t^*, \varepsilon)$ with which we wish to compare $z(t^*, \varepsilon)$. In particular, there is a certain amount of freedom in the choice of α_0, β_0 and in the

normalization of $\alpha_1\beta_0-\beta_1\alpha_0$, and these must be chosen correctly so that the hypotheses (1.8) applies.

For $z(t^*, \varepsilon)$ given, define

$$a(\varepsilon)=\langle\phi_1(t^*), z\rangle, \qquad b(\varepsilon)=\langle\phi_2(t^*), z\rangle.$$

Then, define $u(t^*, \varepsilon)$ to be that periodic solution of (1.1) for which

$$\alpha_0=a(0), \qquad \beta_0=b(0)$$

and for which the normalization

$$\alpha_1\beta_0-\beta_1\alpha_0=(1/\varepsilon)(a(\varepsilon)b(0)-b(\varepsilon)a(0))$$

holds. Then, for the mapping defined by (2.1), (2.3), (2.7), (2.10) and the above normalization, the hypotheses of Lemma 1.5 hold, implying the stated result.

3. Stability and Floquet exponents. An investigation of the linearized stability of periodic solutions of (1.1) reduces to the determination of the Floquet exponents [5]. In the present situation, the Floquet exponents are defined to be those numbers μ_i, $i=1, 2, \cdots, k$, for which the linear problem

$$dU/dt=(P-\mu I)U+\varepsilon F_y(Y, \varepsilon)U \tag{3.1}$$

has periodic solutions, where $Y=Y(t, \varepsilon)$ is a known periodic solution of (1.1). Applying the transformation $t^*=wt$, where $w=w(\varepsilon)$ is the known function for which $Y(t^*, \varepsilon)$ has period 2π, the equation (3.1) becomes

$$\begin{aligned} &LU\equiv\frac{dU}{dt^*}-PU=-\mu U+\varepsilon F_y(Y(t^*, \varepsilon), \varepsilon)U+(1-w)\frac{dU}{dt^*}, \\ &U(0)=U(2\pi). \end{aligned} \tag{3.2}$$

The solution $Y(t, \varepsilon)$ is said to be (linearly) *stable* if the numbers $\mu_i=\mu_i(\varepsilon)$, $i=2, 3, \cdots, k$ have negative real part. Notice that since the system (1.1) is autonomous, the function $U=dy/dt$ satisfies (3.1) for $\mu=0$. Thus, one of the Floquet exponents, μ_1, is zero. The remaining Floquet exponents are given in

THEOREM 3.1. *Let $Y(t^*, \varepsilon)$, $w(\varepsilon)$ be a periodic solution of* (2.1) *found in Theorem* 2.3. *There exists a positive constant $\varepsilon_1\leqq\varepsilon_0$ such that for all ε satisfying $|\varepsilon|\leqq\varepsilon_1$, the Floquet exponents $\mu_i(\varepsilon)$ are given by*

$$\begin{aligned} &\mu_1=0, \\ &\mu_2=\varepsilon\rho_0 f'(\rho_0)+O(\varepsilon^2), \\ &\mu_j=-\lambda_{j-2}+O(\varepsilon), \qquad 2<j\leqq k, \end{aligned} \tag{3.3}$$

where λ_i is the i-th diagonal element of Λ.

COROLLARY 3.2. *A periodic solution $Y(t^*, \varepsilon)$, $w(\varepsilon)$ given by Theorem* 2.3 *is (linearly) stable if and only if $f'(\rho_0)<0$.*

The proof of (3.3) for $j>2$ uses Lemma 1.3 and follows arguments commonly found in branching theory for simple eigenvalues. For $j>2$, define ϕ_j by $\phi_j=(\delta_{ij})$. Then the first approximation to the jth solution U_j, $j>2$ is the constant

$$U_j^0=(\delta_{ij})=\phi_j, \qquad \mu_j^0=-\lambda_{j-2} \tag{3.4}$$

which, since λ_{j-2} is a simple eigenvalue of P, spans the one dimensional null space of the linear operator $L_j = d/dt^* - (P - \mu_j^0 I)$. Expressing $U_j(t^*)$ and μ_j as

$$\begin{aligned} &U_j(t^*) = \phi_j + \varepsilon V_j(t^*, \varepsilon), \qquad \langle \phi_j, V_j \rangle = 0, \\ &\mu_j = -\lambda_{j-2} + \varepsilon \eta_j(\varepsilon) \end{aligned} \tag{3.5a}$$

we can rewrite (3.2) in the form

$$L_j(\varepsilon V_j) = \varepsilon \eta_j U_j + \varepsilon F_y U_j + (1-w)\frac{d(\varepsilon V_j)}{dt^*} \tag{3.5b}$$

where $\varepsilon\eta_j$ is determined by the orthogonality condition

$$\mu_j - \mu_j^0 = \varepsilon \langle \phi_j, F_y(y, \varepsilon) U \rangle + (1-w)\left\langle \phi_j, \frac{dU}{dt^*} \right\rangle. \tag{3.5c}$$

Notice that for any vector function $u = (u_j)$, the inner product with ϕ_j, $j > 2$ reduces to

$$\langle \phi_j, u \rangle = \frac{1}{2\pi} \int_0^{2\pi} u_j(t^*)\, dt^*. \tag{3.5d}$$

The equations (3.5) are now seen to be in the form of (1.4b) for the unknowns $\varepsilon V_j(t^*, \varepsilon)$, $\varepsilon\eta_j(\varepsilon)$ where $V_j(t^*, \varepsilon)$ is required to be periodic in $C^1[0, 2\pi]$. Furthermore, the hypotheses of Lemma 1.3 are satisfied so that the theorem is proven for $j > 2$.

To find the one remaining Floquet exponent $\mu_2(\varepsilon)$, we notice that for $\mu = 0$ the operator $L = d/dt^* - (P - \mu I)$ is exactly the operator encountered in § 2. As such, the process of finding periodic solutions must involve the two linearly independent vectors $\phi_1(t^*)$ and $\phi_2(t^*)$ of (2.2) which span the two dimensional null space of L. We write the solution $U(t^*, \varepsilon)$ as

$$\begin{aligned} &U(t^*, \varepsilon) = c(\varepsilon)\phi_1(t^*) + d(\varepsilon)\phi_2(t^*) + V(t^*, \varepsilon), \\ &\langle \phi_i, V \rangle = 0, \qquad i = 1, 2. \end{aligned} \tag{3.6}$$

Recalling the orthogonality condition necessary to guarantee the invertibility of L, we note that the pair U, μ must satisfy

$$-\mu\langle \phi_i, U \rangle + \varepsilon \langle \phi_i, F_y(Y, \varepsilon) U \rangle + (1-w)\langle \phi_i, dU/dt^* \rangle = 0, \qquad i = 1, 2. \tag{3.7a}$$

With the assumed form of the solution $U(t^*, \varepsilon)$ in (3.6), the orthogonality condition (3.7) reduces to the set of linear equations

$$\begin{aligned} &\begin{pmatrix} \varepsilon\langle \phi_1, F_y(Y, \varepsilon)\phi_1 \rangle - \mu & \varepsilon\langle \phi_1, F_y(Y, \varepsilon)\phi_2 \rangle - (1-w) \\ \varepsilon\langle \phi_2, F_y(Y, \varepsilon)\phi_1 \rangle + (1-w) & \varepsilon\langle \phi_2, F_y(Y, \varepsilon)\phi_2 \rangle - \mu \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} \\ &\qquad = -\varepsilon^2 \begin{pmatrix} \langle \phi_1, F_y(Y, \varepsilon) V \rangle \\ \langle \phi_2, F_y(Y, \varepsilon) V \rangle \end{pmatrix}. \end{aligned} \tag{3.7b}$$

To lowest order in ε, the equation (3.7b) becomes

$$(\varepsilon A - \mu I)\begin{pmatrix} c \\ d \end{pmatrix} = 0 \tag{3.8}$$

so that, to lowest order in ε, $\mu(\varepsilon)$ is an eigenvalue of εA, where the matrix A is given by (2.7b). From (2.15) we know that $\det A = 0$ and $\text{trace}(A) = \rho_0 f'(\rho_0)$, so that the two eigenvalues of εA are

$$\mu_1^0 = 0, \qquad \mu_2^0 = \varepsilon\rho_0 f'(\rho_0).$$

For the eigenvalue μ_2^0, the null space of $\varepsilon A - \mu_2^0$ and its transpose are spanned by

$$\Gamma = \begin{pmatrix} f'(\rho_0)\alpha_0 + g'(\rho_0)\beta_0 \\ f'(\rho_0)\beta_0 - g'(\rho_0)\alpha_0 \end{pmatrix}, \qquad \Gamma^* = \begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix} \tag{3.9}$$

respectively. Thus, we set $\gamma \equiv \begin{pmatrix} c \\ d \end{pmatrix} = \Gamma + \Gamma_1(\varepsilon)$, $\mu_2 = \mu_2^0 + \varepsilon^2\eta_2(\varepsilon)$, where $(\Gamma^*, \Gamma_1) = 0$, and rewrite (3.7b) as

$$(\varepsilon A - \mu_2^0 I)(\varepsilon\Gamma_1) = \varepsilon^2\eta_2(\Gamma + \varepsilon\Gamma_1) - \varepsilon^2 R. \tag{3.10a}$$

Since $(\Gamma^*, \Gamma) = \rho_0 f'(\rho_0) \neq 0$, the orthogonality condition implied by (3.8), (3.9) is satisfied by choosing

$$\eta_2(\Gamma^*, \Gamma) = (\Gamma^*, R). \tag{3.10b}$$

With these details accounted for, equations (3.2), (3.6), (3.10) can be viewed in the framework of Lemma 1.3 as equations for $V(t^*, \varepsilon)$, $\varepsilon\Gamma_1(\varepsilon)$, $\varepsilon^2\eta_2(\varepsilon)$ with $V(t^*, \varepsilon)$ periodic and in $C^1[0, 2\pi]$. Again the estimates (1.7) hold and the proof of the theorem is complete.

4. Two-timing and the initial value problem. In this section we consider the problem (1.1) as an initial value problem subject to the arbitrary, but fixed, initial data

$$Y(0, \varepsilon) = Y_0. \tag{4.1}$$

One method of finding approximate solutions for $\varepsilon \geqq 0$ is the two-timing method which assumes, for purposes of computation, that there are two independent time scales

$$t^* = w(\varepsilon)t, \qquad \tau = \varepsilon t \tag{4.2}$$

the fast and slow times, respectively, which are present in the problem. Using these as independent variables, the differential equation (1.1) is converted into the partial differential equation

$$LY \equiv \frac{\partial Y}{\partial t^*} - PY = \varepsilon F(Y, \varepsilon) + (1 - w)\frac{\partial Y}{\partial t^*} - \varepsilon\frac{\partial Y}{\partial \tau}. \tag{4.3}$$

One then proceeds to find a power series solution of (4.3) which remains bounded for all positive values of t^* and τ.

We want to show that approximate solutions of (1.1), (4.1) generated by the two-timing hypothesis (via iterations or perturbations) have a certain validity for large times in that they are "orbitally valid". In other words, we want to show that after a long time (i.e., $\tau \to \infty$) the approximate solution of (4.3) approaches a limit cycle which is indeed a stable periodic solution of equation (1.1). To do so we must demonstrate that approximate solutions of (4.3) approach approximate solutions

of (2.1) in the sense of (2.19). Thus, we need to show that for approximate solutions of (4.3), derivatives with respect to τ become unimportant as $\tau \to \infty$.

As in § 2, we note that the operator L has a periodic null space spanned by functions whose t^* dependence is $\phi_1(t^*)$ and $\phi_2(t^*)$. The operator L also has null vectors in t^* which are comprised of $e^{-\Lambda t^*}$. Thus, we expect to find solutions of (4.3) of the form

$$Y(t^*, \tau, \varepsilon) = P(t^*, \tau, \varepsilon) + E(t^*, \tau, \varepsilon) \tag{4.4a}$$

where

$$P(t^* + 2\pi, \tau, \varepsilon) = P(t^*, \tau, \varepsilon)$$

and $E(t^*, \tau, \varepsilon)$ is exponentially decaying in t^*. Furthermore, the periodic part of the solution can be written as

$$P(t^*, \tau, \varepsilon) = A(\tau, \varepsilon)\phi_1(t^*) + B(\tau, \varepsilon)\phi_2(t^*) + \varepsilon U(t^*, \tau, \varepsilon) \tag{4.4b}$$

where

$$\langle \phi_i, U \rangle = \lim_{T \to \infty} \frac{1}{T} \int_0^T \phi_i(t^*) \cdot U(t^*, \tau, \varepsilon)\, dt^* = 0.$$

When the function U is periodic, the above inner product is equivalent to the inner product (2.4). With the inner product of (4.4b), a necessary condition for the existence of bounded solutions of (4.3) is that

$$\left\langle \phi_i, \varepsilon F(Y, \varepsilon) + (1 - w)\frac{\partial Y}{\partial t^*} - \varepsilon \frac{\partial Y}{\partial \tau} \right\rangle = 0, \qquad i = 1, 2. \tag{4.5}$$

The equation (4.5) gives two ordinary differential equations in τ which must be satisfied by the solution of (4.3).

The first step of any approximating procedure is to write the solution of $LY = 0$, $Y(0, \varepsilon) = Y_0$ as

$$Y_0(t^*, \tau) = A_0(\tau)\phi_1(t^*) + B_0(\tau)\phi_2(\tau^*) + CE_0(t^*) \tag{4.6a}$$

where

$$E_0(t^*) = \begin{pmatrix} 0 \\ 0 \\ e^{-\Lambda t^*} \end{pmatrix}, \qquad C = (c_j) = (\xi_i \delta_{ij}). \tag{4.6b}$$

The initial data are satisfied by requiring

$$A_0(0) = y_{0_1}, \quad B_0(0) = y_{0_2}, \quad \xi_1 = \xi_2 = 0, \quad \xi_i = y_{0_i} \qquad i > 2. \tag{4.6c}$$

Notice that one can allow the diagonal matrix C to depend on τ, but τ dependence is not necessary to guarantee the existence of bounded approximations. On the other hand, the functions $A_0(\tau)$ and $B_0(\tau)$ are specifically used to eliminate secular terms and hence maintain bounded solutions, and their τ dependence is at the heart of the two-timing method.

The problem of secularity for periodic functions as opposed to exponentially decaying functions is exemplified by comparing the typical secular terms $\varepsilon t \sin t$

and $\varepsilon t e^{-t}$. Both functions can be considered "secular" because of the presence of the term εt, but $\varepsilon t e^{-t}$ is nonetheless bounded for all times, and causes no particular difficulty as one term in some expansion. For example, the regular perturbation expansion of the solution of $y' + (1+\varepsilon)y = 0$, which contains powers of εt, is uniformly bounded for all time, whereas the regular perturbation expansion of $y'' + (1+\varepsilon)y = 0$ is unbounded. The reason, then, that C is allowed to be independent of τ is

LEMMA 4.1. *Suppose* $f(t) \in \mathbb{R}^{k-2}$ *satisfies*

$$\|f(t)\|_\infty \leqq K_0 e^{-\gamma_0 t}, \qquad t \geqq 0, \quad \gamma_0 \geqq 0. \tag{4.7}$$

Then there exist nonnegative constants K_1, γ_1 *such that the solution* $y(t)$ *of*

$$\hat{L}y \equiv \frac{dy}{dt} + \Lambda y = f(t), \qquad y(0) = y_0$$

satisfies

$$\|y(t)\|_\infty \leqq K_1 e^{-\gamma_1 t} \quad \forall t \geqq 0 \tag{4.8}$$

where $\gamma_1 > 0$ *if* $\gamma_0 > 0$, *and* $\gamma_1 = 0$ *if* $\gamma_0 = 0$.

Using (4.6) as a first approximation to the solution of (4.3), we must choose $A_0(\tau)$, $B_0(\tau)$ to satisfy (4.5). With the use of (4.4c), equation (4.5) can be rewritten in the form

$$\begin{aligned} \frac{dA}{d\tau} &= \langle \phi_1, F(Y_0 + \varepsilon U)\rangle - \frac{1-w(\varepsilon)}{\varepsilon} B, \\ \frac{dB}{d\tau} &= \langle \phi_2, F(Y_0 + \varepsilon U)\rangle + \frac{1-w(\varepsilon)}{\varepsilon} A, \end{aligned} \tag{4.9}$$

where we have made use of

LEMMA 4.2. *Suppose* $f(t) \in \mathbb{R}^k$ *satisfies* (4.7) *with* $\gamma_0 > 0$. *Then* $\langle \phi_i, f\rangle = 0$.

LEMMA 4.3. *Suppose* $F\colon \mathbb{R}^k \to \mathbb{R}^k$ *is continuously differentiable and* $f(t)$ *satisfies* (4.7) *with* $\gamma_0 > 0$. *Then*

$$\langle \phi_i, F(u+f)\rangle = \langle \phi_i, F(u)\rangle, \qquad i = 1, 2.$$

Since Y_0 is composed of oscillatory and exponentially decaying terms, Lemma 4.3 applies, and for $\varepsilon = 0$, equation (4.9) reduces to

$$\begin{aligned} \frac{dA}{d\tau} &= \langle \phi_1, F(U_0, 0)\rangle + W_1 B, \\ \frac{dB}{d\tau} &= \langle \phi_2, F(U_0, 0)\rangle - W_1 A \end{aligned} \tag{4.10}$$

where

$$U_0 = A\phi_1(t^*) + B\phi_2(t^*).$$

The right hand side of (4.10) is exactly the expression found in (2.6), so that (4.10) can be simplified to

$$\begin{aligned} \frac{dA}{d\tau} &= \frac{1}{2}f(\rho)A + \left(W_1 - \frac{1}{2}g(\rho)\right)B \\ \frac{dB}{d\tau} &= \left(\frac{1}{2}g(\rho) - W_1\right)A + \frac{1}{2}f(\rho)B, \qquad \rho = A^2 + B^2. \end{aligned} \tag{4.11}$$

It follows from (4.11) that

$$\frac{d\rho}{d\tau} = \rho f(\rho), \qquad \rho = A^2 + B^2 \tag{4.12}$$

where $f(\rho)$ is the polynomial defined in (2.14). Furthermore, using the initial data $\rho(0) = A^2(0) + B^2(0) = y_{0_1}^2 + y_{0_2}^2$, solutions of (4.11) are given by

$$\begin{pmatrix} A_0(\tau) \\ B_0(\tau) \end{pmatrix} = \sqrt{\frac{\rho(\tau)}{\rho(0)}}\left(y_{0_2}\phi_1(G(\tau)) + y_{0_1}\phi_2(G(\tau))\right) \tag{4.13}$$

where

$$G(\tau) = \int_0^\tau \left(W_1 - \tfrac{1}{2}g(\rho)\right) d\tau$$

and where $\rho = \rho(\tau)$ satisfies (4.12) and $g(\rho)$ is defined in (2.14). Notice that the solution U_0 is now given by

$$U_0(t^*, \tau) = \sqrt{\frac{\rho(\tau)}{\rho(0)}}\left(y_{0_2}\phi_1(t^* - G(\tau)) + y_{0_1}\phi_2(t^* - G(\tau))\right) \tag{4.14}$$

so that the function $\rho(\tau)$ gives the slowly changing amplitude of the oscillations and $G(\tau)$ gives the slowly changing phase shift.

It is easy to see that $\rho(\tau)$ will approach a stable root of $f(\rho) = 0$. If we linearize (4.12) about the stable root ρ_0, it is also clear that $\rho(\varepsilon t)$ approaches ρ_0 exponentially with rate of decay μ_2, the Floquet exponent of (3.3). If $G(\tau)$ approaches a constant phase shift and $W_1 = \frac{1}{2}g(\rho_0)$ as in (2.14), then the first approximation (4.6a) does indeed approach a valid first approximation of the limit cycle (2.18). However, according to (4.13), requiring $G(\tau)$ to approach a constant phase shift is equivalent to requiring $W_1 = \frac{1}{2}g(\rho_0)$, since the polynomial $g(\rho)$ decays exponentially to $g(\rho_0)$.

A first approximation of U is provided by

$$LU_1 = F(Y_0, 0) - W_1\frac{\partial Y_0}{\partial t^*} - \frac{\partial Y_0}{\partial \tau}, \qquad \langle \phi_i, U_1 \rangle = 0 \tag{4.15}$$

where Y_0 is given by (4.6), (4.13). Since $Y_0(t^*, \tau)$ was shown to satisfy (4.4) the function $U_1(t^*, \tau)$ is clearly the sum of functions which are exponentially decaying or periodic in t^*. Furthermore, the function $\rho(\tau)$ decays exponentially to ρ_0 with decay rate $\lambda_0 = -\rho_0 f'(\rho_0)$, so that $|e^{\lambda\tau}\,\partial U_1/\partial\tau|$ is bounded for all t^* and τ, for any $0 < \lambda \leqq \lambda_1 < \lambda_0$.

The behavior of higher order approximations is found by induction, using an iteration process. First of all, one must be able to solve (4.9) using higher order approximations of U and $F(Y_0+\varepsilon U)$. When U is nonzero, its effect on the coefficients A and B will be of order ε. Thus, we seek solutions of (4.9) of the form

$$\begin{aligned} A(\tau;\varepsilon) &= A_0(\tau)+\varepsilon A_1(\tau;\varepsilon), \\ B(\tau;\varepsilon) &= B_0(\tau)+\varepsilon B_1(\tau;\varepsilon), \\ w(\varepsilon) &= 1+\varepsilon W_1+\varepsilon^2 W_2(\varepsilon) \end{aligned} \tag{4.16}$$

where $A_0(\tau)$ and $B_0(\tau)$ are given by (4.13) and $W_1=\frac{1}{2}g(\rho_0)$ as in (2.14). Using (4.16) in (4.9) we find that

$$\begin{aligned} &\frac{d}{d\tau}\begin{pmatrix}A_1\\B_1\end{pmatrix}=\mathscr{L}\begin{pmatrix}A_1\\B_1\end{pmatrix}+W_2\begin{pmatrix}B_0\\-A_0\end{pmatrix}+\begin{pmatrix}h_1\\h_2\end{pmatrix}, \\ &A_1(0,\varepsilon)=-u_2(0,0;\varepsilon), \qquad B_1(0,\varepsilon)=-u_1(0,0;\varepsilon) \end{aligned} \tag{4.17a}$$

where

$$\mathscr{L}=\begin{pmatrix}\frac{1}{2}f(\rho)+f'(\rho)A_0^2-g'(\rho)A_0B_0 & W_1-\frac{1}{2}g(\rho)+f'(\rho)A_0B_0-g'(\rho)B_0^2\\ \frac{1}{2}g(\rho)-W_1+g'(\rho)A_0^2+f'(\rho)A_0B_0 & \frac{1}{2}f(\rho)+f'(\rho)B_0^2+g'(\rho)A_0B_0\end{pmatrix}. \tag{4.17b}$$

The functions h_1 and h_2 are remainder terms depending on U, τ, and ε, and on εA_1, εB_1 as well. Notice that the steady state version of (4.17) is exactly of the form (2.7), and, if it is meaningful to talk of the limit as $\tau\to\infty$ of h_1 and h_2, then the steady state limit of (4.9), and hence (4.17), gives solutions of (2.5) and (2.7) respectively. Thus, if the solutions of (4.17) have a steady state limit, the limit will automatically correspond to a known periodic solution of (1.1).

A necessary induction assumption is that for U_n, the nth approximation of U, $|e^{\lambda\tau}\,\partial U_n/\partial\tau|$ be bounded for all λ with $0<\lambda<\lambda_0$. Then the functions h_1 and h_2 have a steady state limit as $\tau\to\infty$. Equation (4.17) can be rewritten as an integral equation as follows: Define the matrices

$$\begin{aligned} K_0(\tau) &= \begin{pmatrix}\sin G(\tau) & \cos G(\tau)\\ \cos G(\tau) & -\sin G(\tau)\end{pmatrix}; \qquad G(\tau)=\int_0^\tau (W_1-\tfrac{1}{2}g(\rho(\sigma)))\,d\sigma; \\ K_1(\tau) &= f(\rho(\tau))\begin{pmatrix}\alpha^2 & \alpha\beta\\ \alpha\beta & \beta^2\end{pmatrix}-2G'(\tau)\begin{pmatrix}\alpha\beta & \beta^2\\ -\alpha^2 & -\alpha\beta\end{pmatrix}, \qquad \alpha=B_0(0), \\ & \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \beta=A_0(0); \\ K_2(\tau) &= \frac{2G'(\tau)}{f(\rho(\tau))}\begin{pmatrix}\alpha\beta & \beta^2\\ -\alpha^2 & -\alpha\beta\end{pmatrix}+\begin{pmatrix}\beta^2 & -\alpha\beta\\ -\alpha\beta & \alpha^2\end{pmatrix}, \\ K_3(\tau) &= \begin{pmatrix}\frac{1}{2}f(\rho(\tau)) & 0\\ -G'(\tau) & 1\end{pmatrix}. \end{aligned} \tag{4.18}$$

Then

$$\begin{pmatrix} A_1(\tau) \\ B_1(\tau) \end{pmatrix} = \sqrt{\frac{\rho(\tau)}{\rho(0)}} K_0(\tau) \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix} K_3(\tau) \begin{pmatrix} \gamma \\ \delta \end{pmatrix}$$

$$+ \sqrt{\frac{\rho(\tau)}{\rho(0)}} K_0(\tau) \int_0^\tau \left\{ W_2 \begin{pmatrix} -\beta \\ \alpha \end{pmatrix} + \frac{1}{\sqrt{\rho(0)\rho(\sigma)}} \right.$$

$$\left. \cdot \left[\left(\frac{1}{f(\rho(\sigma))} \right) K_1(\tau) + K_2(\sigma) \right] K_0(\sigma) \begin{pmatrix} h_1(\sigma) \\ h_2(\sigma) \end{pmatrix} \right\} d\sigma \tag{4.19}$$

where γ, δ are determined directly from the initial data on A_1, B_1. Since h_1 and h_2 depend on εA_1 and εB_1, equation (4.19) defines a nonlinear mapping for A_1 and B_1. We can study the properties of the solution of (4.19) by introducing the norm

$$\|u\|_\lambda = \|u\|_\infty + \|e^{\lambda\tau} du/d\tau\|_\infty, \qquad \lambda > 0 \tag{4.20}$$

where $\|\cdot\|_\infty$ is the usual vector maximum norm for $\tau \in (0, \infty)$. In this norm we have the following fact:

LEMMA 4.4. *Suppose* $\|u\|_\lambda < \infty$. *Then* $u_\infty = \lim_{\tau\to\infty} u(\tau)$ *exists and is unique.*

Proof. By definition

$$|u(\tau_1) - u(\tau_2)| \leqq \int_{\tau_1}^{\tau_2} \left| \frac{du}{d\tau} \right| d\tau \leqq \frac{\|u\|_\lambda}{\lambda} (e^{-\lambda\tau_1} - e^{-\lambda\tau_2}).$$

Thus, for an arbitrary increasing sequence $\{\tau_i\}$, the sequence $\{u(\tau_i)\}$ is a Cauchy sequence as $\tau_i \to \infty$.

In order to make the mapping (4.19) well defined, we must decide how W_2 is to be chosen. The choice which is consistent with our goal is to require that the integral (4.19) be convergent as $\tau \to \infty$. Then the limit $\tau \to \infty$ of $A_1(\tau)$, $B_1(\tau)$ will exist and will satisfy the steady state version of (4.17). Accordingly, we require

$$W_2 \begin{pmatrix} -\beta \\ \alpha \end{pmatrix} = -\lim_{\tau\to\infty} \frac{1}{\sqrt{\rho(0)\rho(\tau)}} K_2(\tau) K_0(\tau) \begin{pmatrix} h_1(\tau) \\ h_2(\tau) \end{pmatrix}. \tag{4.21a}$$

At first glance the condition (4.21a) appears to contain two requirements with only one unknown W_2. However, (4.21a) is equivalent to requiring

$$W_2 = \lim_{\tau\to\infty} \frac{1}{\sqrt{\rho(0)\rho(\tau)}} \left[\frac{2G'(\tau)}{f(\rho(\tau))} (\alpha H_1 + \beta H_2) + (\beta H_1 - \alpha H_2) \right] \tag{4.21b}$$

where $G'(\tau) = W_1 - \frac{1}{2} g(\rho(\tau))$,

$$\begin{pmatrix} H_1(\tau) \\ H_2(\tau) \end{pmatrix} = K_0(\tau) \begin{pmatrix} h_1(\tau) \\ h_2(\tau) \end{pmatrix}. \tag{4.21c}$$

Notice that

$$\lim_{\tau\to\infty} \frac{2G'(\tau)}{f(\rho(\tau))} = \frac{g'(\rho_0)}{f'(\rho_0)}, \quad \text{where } f(\rho_0) = 0$$

and by assumption (inductive hypothesis) the limit of $H_i(\tau)$ exists. Hence the value for W_2 is uniquely defined by (4.21b). It is of interest to our later discussion to notice that by taking the limit as $\tau \to \infty$ in (4.17) one finds the same result. That is, in the limit as $\tau \to \infty$ of (4.17), the null space of $\mathscr{L}(\tau)$ and $\mathscr{L}^T(\tau)$ are spanned by

$$\text{(4.22a)} \quad \Gamma_\infty = \lim_{\tau\to\infty} K_0(\tau)\begin{pmatrix} -\beta \\ \alpha \end{pmatrix}, \qquad \Gamma_\infty^* = \lim_{\tau\to\infty} \begin{pmatrix} g'(\rho(\tau)) & -f'(\rho(\tau)) \\ f'(\rho(\tau)) & g'(\rho(\tau)) \end{pmatrix} K_0(\tau)\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

respectively, so that in the limit as $\tau \to \infty$ one should find that

$$\text{(4.22b)} \qquad \lim_{\tau\to\infty} \left(\Gamma_\infty^*, W_2\begin{pmatrix} B_0(\tau) \\ -A_0(\tau) \end{pmatrix} + \begin{pmatrix} h_1(\tau) \\ h_2(\tau) \end{pmatrix} \right) = 0.$$

It is not surprising to find that (4.22a) applied to (4.22b) reduces to the requirement (4.21b).

Using (4.21), the mapping (4.19) can be rewritten as

$$\begin{aligned} \begin{pmatrix} A_1(\tau) \\ B_1(\tau) \end{pmatrix} &= \sqrt{\frac{\rho(\tau)}{\rho(0)}} K_0(\tau)\begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix} K_3(\tau)\begin{pmatrix} \gamma \\ \delta \end{pmatrix} \\ &\quad + \frac{1}{\rho(0)} K_0(\tau) \int_0^\tau \sqrt{\frac{\rho(\tau)}{\rho(\sigma)}} \frac{1}{f(\rho(\sigma))} K_1(\tau)\begin{pmatrix} h_1(\sigma) \\ h_2(\sigma) \end{pmatrix} d\sigma \\ &\quad + \frac{1}{\rho(0)} \sqrt{\rho(\tau)}\, K_0(\tau) \int_0^\tau \left\{ \frac{1}{\rho(\sigma)} K_2(\sigma) K_0(\sigma)\begin{pmatrix} h_1(\sigma) \\ h_2(\sigma) \end{pmatrix} \right. \\ &\qquad \left. - \left[\frac{1}{\rho(\eta)} K_2(\eta) K_0(\eta)\begin{pmatrix} h_1(\eta) \\ h_2(\eta) \end{pmatrix} \right]_{\eta=\infty} \right\} d\sigma. \end{aligned} \tag{4.23}$$

The two integrals of (4.23) suggest that we consider two mappings described in the following lemmas.

LEMMA 4.5. *Suppose that the function $F(t, u)$ satisfies*

$$\text{(4.24a)} \qquad \|F(t, u)\|_\lambda \leqq C_1 \quad \textit{for each fixed } u, \qquad |u| \leqq C_0$$

$$\text{(4.24b)} \qquad \begin{aligned} &\frac{\partial F}{\partial u} \textit{ exists and is uniformly Lipschitz continuous in } u \textit{ for } |u| \leqq C_0, \\ &\textit{and } \left\| \frac{\partial F}{\partial u}(t, u) \right\|_\lambda \leqq C_2 \textit{ for all } |u| \leqq C_0. \end{aligned}$$

If the function $G(t)$ satisfies $\|G(t)\|_\lambda < \infty$, then the mapping $u \to Tu$ defined for each $u(t)$ satisfying $\|u\|_\lambda < C_0$ by

$$\text{(4.25)} \qquad Tu = G(\tau) \int_0^\tau [F(\sigma, u(\sigma)) - F(\infty, u(\infty))]\, d\sigma$$

is bounded and continuous in the norm $\|\cdot\|_\lambda$ of (4.20).

LEMMA 4.6. *Suppose that the function $F(t, u)$ satisfies*

$$\|F(t, u)\|_\lambda < C_1 \quad \text{for each fixed } u, \quad |u| \leqq C_0 \tag{4.26a}$$

$$\frac{\partial F}{\partial u} \text{ exists, is Lipschitz continuous in } u, \text{ and} \tag{4.26b}$$

$$\left\|\frac{\partial F}{\partial u}(t, u)\right\|_\lambda \leqq C_2 \text{ for each fixed } u, \quad |u| \leqq C_0.$$

If the functions $G(t)$ and $H(t)$ are differentiable and satisfy

$$\|G\|_{\lambda_0} \leqq C_3, \qquad |G(t)| \leqq C_3 e^{-\lambda_0 t} \tag{4.26c}$$

$$C_4 e^{-\lambda_0 \tau} \leqq |H(t)| \leqq C_5 e^{-\lambda_0 \tau}, \qquad \lambda_0 > 0,$$

$$\left\| \int_0^t \frac{G(t)}{H(\sigma)}\, d\sigma \right\|_{\lambda_0} < \infty, \tag{4.26d}$$

then the mapping $u \to Su$ defined for each $u(t)$ satisfying $\|u\|_\lambda \leqq C_0$ by

$$Su = \int_0^t \frac{G(t)}{H(\sigma)} F(\sigma, u(\sigma))\, d\sigma \tag{4.27}$$

is a bounded continuous mapping in the norm (4.20) *for any $\lambda < \lambda_0$.*

The proofs of the preceding lemmas can be found in the Appendix B.

The mapping (4.23) consists of mappings of εA_1, εB_1 of the form (4.25) and (4.27). Thus, to show that εA_1 and εB_1 exist and are bounded in the norm (4.20), we need to verify that the hypotheses (4.24) and (4.26) hold. In particular we must verify that $f(\rho(\tau))$ and $K_1(\tau)$ satisfy the hypothesis of Lemma 4.6, and that the functions h_1, h_2 satisfy the hypothesis (4.24), (4.26) whenever εA_1 and εB_1 are bounded in the norm (4.20).

Notice that when the root ρ_0 of $f(\rho)$ satisfies $f'(\rho_0) \neq 0$, then solutions $\rho(\tau)$ of (4.12) decay to ρ_0 with linear decay rate $\lambda_0 = -\rho f'(\rho_0)$. Furthermore, the polynomial character of $F(y; \varepsilon)$ maintains this decay. Exponential decay for $U_n(t^*, \varepsilon)$ follows inductively from the fact that when $LU = g$, $\langle \phi_i, g \rangle = 0$, $i = 1, 2$, and $|e^{\lambda\tau}\, \partial g/\partial \tau|$ is uniformly bounded for all t^* and τ, then $|e^{\lambda\tau}\, \partial U/\partial \tau|$ is also uniformly bounded for all t^* and τ. It follows from using Lemma 1.3 that the nth order approximate solution of (4.3) exists, is uniformly bounded in t^* and exponentially decaying in τ.

THEOREM 4.7. *Suppose $\rho_0 \neq 0$ is a root of $f(\rho_0) = 0$ and $f'(\rho_0) < 0$. Suppose that $Y_n(t^*, \tau; \varepsilon)$, $w_n(\varepsilon)$ is an iterative approximate solution of* (4.3) *satisfying*

$$LY_n - \varepsilon F(Y_n, \varepsilon) - (1 - w_n)\frac{\partial Y_n}{\partial t^*} + \varepsilon \frac{\partial Y_n}{\partial \tau} = O(\varepsilon^n), \tag{4.28}$$

$$Y(0, 0, \varepsilon) = Y_0$$

uniformly for all positive t^*, τ and for ε sufficiently small. Then there exists a function $U_n(t^*, \varepsilon)$, periodic in t^*, independent of τ, satisfying

$$LU_n - \varepsilon F(U_n, G) - (1 - w_n)\frac{dU_n}{dt^*} = O(\varepsilon^n) \tag{4.29}$$

for which

$$\|U_n(t^*;\varepsilon)-Y_n(t^*,\tau;\varepsilon)\|_\infty \leqq K_1 e^{-\nu_1\tau}+K_2 e^{-\nu_2 t^*}. \tag{4.30}$$

Since $Y_n(t^*,\tau;\varepsilon)=P_n(t^*,\tau;\varepsilon)+E_n(t^*,\tau;\varepsilon)$ (see (4.4a)), define $U_n(t^*,\varepsilon)=\lim_{\tau\to\infty} P_n(t^*,\tau;\varepsilon)$. Then the exponential decay in τ and the decay of $E_n(t^*,\tau;\varepsilon)$ guarantee that (4.29) holds. Furthermore

$$\|U_n(t^*,\varepsilon)-Y_n(t^*,\tau;\varepsilon)\| \leqq \|P_n(t^*,\infty;\varepsilon)-P_n(t^*,\tau;\varepsilon)\|+\|E_n(t^*,\tau;\varepsilon)\|,$$

from which (4.30) follows.

THEOREM 4.8. *Suppose $\rho_0\neq 0$ is a root $f(\rho)=0$ with $f'(\rho_0)<0$. Suppose that $Y_n(t^*,\tau,\varepsilon)$ satisfies* (4.28). *Then for each fixed $T>0$, $\delta>0$, there exists a stable periodic solution $U(t,\varepsilon)$ of* (1.1) *and a phase shift $\phi(\varepsilon)$ for which*

$$\max_{(T,T+\delta)} \|U(t-\phi(\varepsilon),\varepsilon)-Y_n(w_n(\varepsilon)t,\varepsilon t;\varepsilon)\| \leqq O(\varepsilon^n)+K_1 e^{-\varepsilon\nu_1 t}+K_2 e^{-\nu_2 T}. \tag{4.31}$$

Let $U(t;\varepsilon)$ be the periodic solution of (1.1) found using the $u_n(t^*;\varepsilon)$ of Theorem 4.7 and Corollary 2.4. Then by a standard change of variables, define $U(t;\varepsilon)=V(w(\varepsilon)t;\varepsilon)$ where V is 2π periodic in its first argument. Then we have

$$\begin{aligned}\|U(t-\phi;\varepsilon)-Y_n(w_n t,\varepsilon t;\varepsilon)\| &= \|V(w(t-\phi),\varepsilon)-Y_n(w_n t,\varepsilon t;\varepsilon)\| \\ &\leqq \|V(w(t-\phi);\varepsilon)-U_n(w(t-\phi);\varepsilon)\| \\ &\quad +\|U_n(w(t-\phi);\varepsilon)-U_n(w_n(t-\varepsilon);\varepsilon)\| \\ &\quad +\|U_n(w_n t;\varepsilon)-Y_n(w_n t,\varepsilon t;\varepsilon)\|\end{aligned}$$

where we require that $\phi(\varepsilon)=2\pi k/(w_n(\varepsilon))$ for some integer k. It follows from (4.29), (2.20) and (4.30) that

$$\|u(t-\phi;\varepsilon)-Y_n(w_n t,\varepsilon t;\varepsilon t)\| \leqq \varepsilon^n(K_3+K_4|t-\phi|)+K_1 e^{-\varepsilon\nu_1 t}+K_2 e^{-\nu_2 w_n t}.$$

The estimate (4.31) now follows when $|t-\phi|$ is restricted in size. Thus, for an estimate where T is "large", the integer k is chosen to keep $|T-\phi|$ "small".

5. Two-timing and domains of attraction. In this section we discuss briefly the question of finding the domain of attraction for a given stable periodic solution of (1.1). In fact, in certain situations, the two-timing solution predicts incorrect limiting behavior, even though we know that the limiting limit cycle is always a stable periodic solution of the original problem.

THEOREM 5.1. *Let $k=2$ and suppose the polynomial $f(\rho)$ has at least two stable positive roots $\rho_1<\rho_3$ (i.e. $f'(\rho_i)<0$) separated by one unstable root ρ_2. Then there exists a region $\mathscr{D}(\varepsilon)$ in $\mathbb{R}^2$ such that the solution of* (1.1) *with initial data $Y_0\in\mathscr{D}(\varepsilon)$ approaches the limit cycle associated with the root $\rho_1(\rho_3)$, while the two-timing method predicts that the solution of the same initial value problem approaches the limit cycle associated with $\rho_3(\rho_1)$.*

Proof. Let $Y_i(t,\varepsilon)$, $i=1,2,3$, be the periodic solution of (1.1) associated with the root ρ_i. In $\mathbb{R}^2$ define $I(C)$ and $E(C)$ to be the interior and exterior, respectively, of any closed curve C. From phase plane considerations, the domain of attraction of $Y_1(t,\varepsilon)$ lies entirely inside $I(Y_2(t,\varepsilon))$, while the two-timing method (see (4.12), (4.13)) predicts that the domain of attraction of $Y_3(t,\varepsilon)$ lies in $E(Y_2(t,0))$. At least

one of

$$\mathscr{D}_1(\varepsilon)=I(Y_2(t,\varepsilon))\cap E(Y_2(t,0)),$$
$$\mathscr{D}_2(\varepsilon)=I(Y_2(t,0))\cap E(Y_2(t,\varepsilon))$$

is nonempty, so that two-timing predicts a completely incorrect asymptotic limit cycle for initial data $Y_0\in\mathscr{D}_1(\varepsilon)\cup\mathscr{D}_2(\varepsilon)=\mathscr{D}(\varepsilon)$.

It would be illustrative to give an example of a system to which Theorem 5.1 applies and to show what goes wrong in the expansion procedure. However, for such a system $f(\rho)$ is at least a cubic polynomial and the actual calculations are far too cumbersome to present here. Instead we consider a less complicated example not satisfying the hypotheses of Theorem 5.1 but which nonetheless has the same failings mentioned there.

Consider the nonlinear system

$$\begin{aligned} dy_1/dt &= y_2+\varepsilon y_1(y_1^2+y_2^2-1+\varepsilon),\\ dy_2/dt &= -y_1+\varepsilon y_2(y_1^2+y_2^2-1+\varepsilon). \end{aligned} \tag{5.1}$$

Obviously, this example has been rigged to have an easily found exact solution. In fact, multiplying the two equations by y_1 and y_2 respectively and adding gives

$$dr/dt=2\varepsilon r(r-1+\varepsilon),\qquad r=y_1^2+y_2^2, \tag{5.2a}$$

and setting $\theta=\tan^{-1}(y_1/y_2)$ gives

$$d\theta/dt=1. \tag{5.2b}$$

Thus, the explicit solution of (5.1) can be written as

$$\begin{aligned} y_1(t,\varepsilon)&=r^{1/2}(t,\varepsilon)\sin(t+\phi_0),\\ y_2(t,\varepsilon)&=r^{1/2}(t,\varepsilon)\cos(t+\phi_0),\\ r(t,\varepsilon)&=\frac{r_0(1-\varepsilon)}{r_0-(r_0-1+\varepsilon)\exp(2\varepsilon t/(1+\varepsilon))},\\ r_0&=y_1^2(0)+y_2^2(0). \end{aligned} \tag{5.3}$$

Notice that the solution (5.3) blows up in a finite amount of time if $r_0>1-\varepsilon$, and the solution decays to zero otherwise. There are no stable periodic orbits for this problem, although the orbit with $y_1^2+y_2^2=1-\varepsilon$ is unstable.

The computation of the two-timing expansion for this problem is straightforward. Using (2.11), (2.12) and Appendix C, we find the first order amplitude equation

$$d\rho/d\tau=2\rho(\rho-1) \tag{5.4}$$

for which

$$\rho(\tau)=\frac{\rho_0}{\rho_0-(\rho_0-1)e^{2\tau}},\qquad \rho(0)=\rho_0.$$

The second order equations for εA_1, εB, reduce to

$$A^2+B^2=\rho(\tau)+\varepsilon\eta(\tau),\qquad \frac{d\eta}{d\tau}=2(2\rho(\tau)-1)\eta+2\rho(\tau), \tag{5.5}$$

whose solution we find to be

$$\eta(\tau)=\left[\frac{\rho_0}{\rho_0-1}(e^{-2\tau}-1)+2\tau\right](\rho(\tau)-\rho^2(\tau)).$$

Thus, we can write the approximate solution of (5.1) as

$$y_1(t;\varepsilon)=A^{1/2}(t;\varepsilon)\sin(t+\phi_0),\qquad y_2(t;\varepsilon)=A^{1/2}(t;\varepsilon)\cos(t+\phi_0)$$

$$A(t;\varepsilon)=\rho(\tau)\left\{1+(1-\rho(\tau))\left[\frac{\varepsilon\rho_0}{\rho_0-1}(e^{-2\varepsilon t}-1)+2\varepsilon^2 t\right]\right\} \tag{5.6}$$

$$\rho(\tau)=\frac{\rho_0}{\rho_0-(\rho_0-1)e^{2\varepsilon t}}.$$

For $\rho_0<1$, $\rho(\tau)$ decays exponentially to zero, and therefore $A(t,\varepsilon)$ also decays exponentially to zero, regardless of how close ρ_0 is to 1. Thus, we see that, for ε fixed, the two timing approximation fails to predict the correct limiting behavior for all initial data with $1-\varepsilon<y_1^2(0)+y_2^2(0)<1$. It is interesting to note that this failure is in some sense predicted by the expansion itself. Notice that the correction term $\eta(\tau)$ contains the factor $\rho_0/(\rho_0-1)$ which for $1>\rho_0>1-\varepsilon$ satisfies $|\rho_0/(\rho_0-1)|>1/\varepsilon$. It could therefore be argued that the correction term $\eta(\tau)$ is not uniformly small in ρ_0 for fixed ε, suggesting that the expansion has failed. Of course, for fixed ρ_0 one can always find an ε sufficiently small so that the expansion is again correct, but the expansion itself does not tell us how small this ε must be.

Notice that this result does not contradict the pointwise validity of the two-timing method for times of order $O(1/\varepsilon)$, but it does indicate how drastically incorrect the long time behavior can be. As seen in § 4, higher order corrections are uniformly bounded and therefore cannot overcome the initial defects of $\rho(\tau)$.

The region $\mathcal{D}(\varepsilon)$ in Theorem 5.1 has an area whose size is order $O(\varepsilon)$. In effect, the above theorem suggests that the two-timing method can predict the boundary of a given domain of attraction only to within regions of order $O(\varepsilon)$. It seems reasonable, however, that exclusive of some boundary region of order $O(\varepsilon)$, the two-timing method predicts the correct asymptotic limit cycle, but this conjecture remains to be proven in general. Unfortunately, the correct domain of attraction can be ascertained only from some a priori information which is independent of the two-timing approximation itself, as for example, with the Van der Pol oscillator, for which there is a unique stable periodic solution.

Appendix A. In this Appendix we give a derivation of equations (4.18), (4.19) from (4.17). First, we want to find the fundamental solution matrix $Y_0(\tau)$ for the homogeneous equation

$$\frac{d}{d\tau}\begin{pmatrix}u\\v\end{pmatrix}=\mathscr{L}(\tau)\begin{pmatrix}u\\v\end{pmatrix} \tag{A.1}$$

where $\mathscr{L}$ is given by (4.17b). Since (A.1) is a linearization of (4.11), one solution of (A.1) is given by $\begin{pmatrix}u_1\\v_1\end{pmatrix}=\begin{pmatrix}A_0'\\B_0'\end{pmatrix}$ where $A_0(\tau)$, $B_0(\tau)$, the solutions of (4.11) are given

by

$$\text{(A.2)}\qquad \begin{pmatrix} A_0(\tau) \\ B_0(\tau) \end{pmatrix} = \sqrt{\frac{\rho(\tau)}{\rho(0)}}\Phi(\tau)\begin{pmatrix} \alpha \\ \beta \end{pmatrix}; \qquad \Phi(\tau) = \begin{pmatrix} \sin G(\tau) & \cos G(\tau) \\ \cos G(\tau) & -\sin G(\tau) \end{pmatrix}$$
$$G(\tau) = \int_0^\tau (W_1 - \tfrac{1}{2}g(\rho(\sigma)))d\sigma, \qquad \alpha^2 + \beta^2 = \rho(0).$$

Thus, one solution of (A.1) is given by

$$\text{(A.3)}\qquad \begin{pmatrix} U_1 \\ V_1 \end{pmatrix} = \sqrt{\frac{\rho(\tau)}{\rho(0)}}\begin{pmatrix} \frac{1}{2}f(\rho(\tau)) & G'(\tau) \\ -G'(\tau) & \frac{1}{2}f(\rho(\tau)) \end{pmatrix}\Phi(\tau)\begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

To find a second linearly independent solution, notice that $W = \det Y_0(\tau)$ must satisfy

$$\text{(A.4)}\qquad dW/d\tau = (\operatorname{trace} \mathscr{L}(\tau))W = (f(\rho) + \rho f'(\rho))W.$$

Now,

$$\frac{d}{d\tau}\left(\ln \frac{d\rho}{d\tau}\right) = f(\rho(\tau)) + \rho f'(\rho(\tau))$$

so that

$$\text{(A.5)}\qquad W = W_0\rho(\tau)f(\rho(\tau)).$$

Furthermore,

$$\frac{1}{2}\frac{d\rho}{d\tau} = \frac{1}{2}\rho f(\rho) = A_0A_0' + B_0B_0' = A_0u_1 + B_0v_1.$$

Thus, a second linearly independent solution of A.1 is

$$\text{(A.6)}\qquad \begin{pmatrix} u_2 \\ v_2 \end{pmatrix} = \begin{pmatrix} -B_0 \\ A_0 \end{pmatrix} = \sqrt{\frac{\rho(\tau)}{\rho(0)}}\Phi(\tau)\begin{pmatrix} \beta \\ -\alpha \end{pmatrix}$$

so that

$$\text{(A.7a)}\qquad Y_0(\tau) = \sqrt{\frac{\rho(\tau)}{\rho(0)}}\Phi(\tau)\begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix}\begin{pmatrix} \frac{1}{2}f(\rho(\tau)) & 0 \\ -G'(\tau) & 1 \end{pmatrix}$$

and

$$\text{(A.7b)}\qquad Y_0^{-1}(\tau) = \frac{1}{\sqrt{\rho(0)\rho(\tau)}}\begin{pmatrix} \dfrac{2}{f(\rho(\tau))} & 0 \\ \dfrac{2G'(\tau)}{f(\rho(\tau))} & 1 \end{pmatrix}\begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix}\Phi(\tau).$$

Finally, we note that the solution of (4.17) is given by

$$\text{(A.8)}\qquad \begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = Y_0(\tau)\begin{pmatrix} A_1(0) \\ B_1(0) \end{pmatrix} + Y_0(\tau)\int_0^\tau Y_0^{-1}(\sigma)\begin{pmatrix} h_1(\sigma) \\ h_2(\sigma) \end{pmatrix} d\sigma,$$

and we easily calculate that

$$\text{(A.9a)}\quad Y_0(\tau)Y_0^{-1}(\sigma)=\frac{1}{\rho(0)}\sqrt{\frac{\rho(\tau)}{\rho(\sigma)}}\Phi(\tau)\left[\frac{1}{f(\rho(\sigma))}K_1(\tau)+K_2(\sigma)\right]\Phi(\sigma)$$

and that

$$\text{(A.9b)}\qquad Y_0(\tau)Y_0^{-1}(\sigma)\begin{pmatrix}B_0(\sigma)\\-A_0(\sigma)\end{pmatrix}=\sqrt{\frac{\rho(\tau)}{\rho(0)}}\Phi(\tau)\begin{pmatrix}-\beta\\\alpha\end{pmatrix},$$

from which (4.19) naturally follows.

Appendix B.

Proof of Lemma 4.6. To show that the mapping Tu, defined by (4.25), is bounded in the norm $\|\cdot\|_\gamma$, we must show that $|Tu|$ and $|e^{\lambda\tau}\,dTu/d\tau|$ are bounded for some λ. We make frequent use of the inequality

$$\text{(B.1)}\qquad \begin{aligned}|u(\sigma)-u(\infty)|&\leqq\int_\tau^\infty\left|\frac{du}{d\sigma}\right|d\sigma\leqq\int_\tau^\infty e^{-\lambda\sigma}\left|e^{\lambda\sigma}\frac{du}{d\sigma}\right|d\sigma\\&\leqq\|u\|_\lambda\int_\tau^\infty e^{-\lambda\sigma}\,d\sigma\leqq\frac{\|u\|_\lambda}{\lambda}e^{-\lambda\tau}.\end{aligned}$$

To show that Tu is bounded, note by the triangle inequality and (4.24a,b) that

$$\text{(B.2)}\qquad \begin{aligned}|F(\sigma,u(\sigma))-F(\infty,u_\infty)|&\leqq|F(\sigma,u(\sigma))-F(\sigma,u_\infty)|+|F(\sigma,u_\infty)-F(\infty,u_\infty)|\\&\leqq C_2|u-u_\infty|+\|F(\cdot\,,u_\infty)\|\frac{e^{-\lambda\sigma}}{\lambda}\\&\leqq\frac{1}{\lambda}(C_2\|u\|_\lambda+C_1)\,e^{-\lambda\sigma}.\end{aligned}$$

Thus, from (4.25),

$$|Tu|\leqq|G(\tau)|\cdot\frac{1}{\lambda^2}(C_2\|u\|_\lambda+C_1)(1+e^{-\lambda\tau})\leqq\frac{1}{\lambda^2}\|G\|_\lambda(C_2\|u\|_\lambda+C_1).$$

From (4.25), it is obvious that

$$\frac{dTu}{d\tau}=G'(\tau)\int_0^\tau(F(\sigma,u(\sigma))-F(\infty,u_\infty))\,d\sigma+G(\tau)(F(\tau,u(\tau))-F(\infty,u_\infty))$$

so that, from the boundedness of $G(\tau)$ and (B.2)

$$\text{(B.3)}\qquad \left|\frac{dTu}{d\tau}\right|<\|G\|_\lambda(C_2\|u\|_\lambda+C_1)\left(\frac{1}{\lambda}+\frac{1}{\lambda^2}\right)e^{-\lambda\tau},$$

from which it follows that $|e^{\lambda\tau}\,dTu/d\tau|$ is bounded.

To show the continuity of Tu we must show that $|Tu-Tv|$ and $|e^{\lambda\tau}(dTu/d\tau-dTv/d\tau)|$ are bounded for positive τ. To do so requires that we

make use of the following identity:

$$
\begin{aligned}
&(f(u)-f(u_\infty))-(f(v)-f(v_\infty)) \\
&\quad=((u-u_\infty)-(v-v_\infty))\int_0^1 f'(su+(1-s)u_\infty)\,ds \\
&\quad\quad+(v-v_\infty)\int_0^1 [f'(su+(1-s)u)_\infty - f'(sv+(1-s)v_\infty)]\,ds.
\end{aligned}
\tag{B.4}
$$

If df/du is assumed to be uniformly Lipschitz continuous, then

$$
|(f(u)-f(u_\infty))-(f(v)-f(v_\infty))| \leqq (K_1+K_2|v-v_\infty|)|(u-u_\infty)-(v-v_\infty)|. \tag{B.5}
$$

Furthermore, from (B.1) we have that

$$
|(u-u_\infty)-(v-v_\infty)| \leqq \frac{1}{\lambda}\|u-v\|_\lambda\, e^{-\lambda\tau}. \tag{B.6}
$$

Thus, with the use of (B.5) and (B.6),

$$
\begin{aligned}
&|(F(\sigma,u)-F(\infty,u_\infty))-(F(\sigma,v)-F(\infty,v_\infty))| \\
&\quad\leqq |(F(\sigma,u(\sigma))-F(\sigma,u_\infty))-(F(\sigma,v(\sigma))-F(\sigma,v_\infty))| \\
&\quad\quad+|(F(\sigma,u_\infty)-F(\sigma,v_\infty))-(F(\infty,u_\infty)-F(\infty,v_\infty))| \\
&\quad\leqq \left(C_2+C_3\frac{\|v\|_\lambda}{\lambda}e^{-\lambda\sigma}\right)\|u-v\|_\lambda\frac{e^{-\lambda\sigma}}{\lambda}+C_2\,e^{-\lambda\sigma}|u_\infty-v_\infty| \\
&\quad\leqq K\frac{e^{-\lambda\sigma}}{\lambda}\|u-v\|_\lambda.
\end{aligned}
\tag{B.7}
$$

The continuity of Tu follows directly from (B.7).

Proof of Lemma 4.7. To show that the mapping Su, defined by (4.27), is bounded and continuous in $\|\cdot\|_\lambda$ for $\lambda<\lambda_0$ we define two new mappings S_1u and S_2u by

$$
\begin{aligned}
S_1u &= \int_0^\tau \frac{G(\tau)}{H(\sigma)}[F(\sigma,u(\sigma))-F(\infty,u(\infty))]\,d\sigma, \\
S_2u &= F(\infty,u_\infty)\int_0^\tau \frac{G(\tau)}{H(\sigma)}\,d\sigma.
\end{aligned}
\tag{B.8}
$$

Clearly the mapping Su is the sum of S_1u and S_2u, and the boundedness and continuity of S_2u follow directly from the assumption (4.26d).

For S_1u, notice that from (4.26c) and (B.2),

$$
|S_1u| \leqq C_3\,e^{-\lambda_0\tau}\int_0^\tau C_4\,e^{\lambda_0\sigma}\cdot\frac{1}{\lambda}(C_2\|u\|_\lambda+C_1)\,e^{-\lambda\sigma}\,d\sigma, \tag{B.9}
$$

which approaches zero for all positive values of λ and λ_0. Furthermore,

$$
\frac{dS_1u}{d\tau}=G'(\tau)\int_0^\tau\frac{F(\sigma,u(\sigma))-F(\infty,u_\infty)}{H(\sigma)}\,d\sigma+\frac{G(\tau)}{H(\tau)}(F(\tau,u(\tau))-F(\infty,u_\infty))
$$

so that

$$\text{(B.10)}\qquad \left|\frac{dS_1u}{d\tau}\right| \leqq \frac{1}{\lambda}\,\frac{(C_2\|u\|_\lambda + C_1)}{C_4}\left(\left(\frac{\|G\|_{\lambda_0}}{\lambda_0-\lambda}+C_3\right)e^{-\lambda\tau} - \frac{\|G\|_{\lambda_0}}{\lambda_0-\lambda}\,e^{-\lambda_0\tau}\right), \qquad \lambda \neq \lambda_0.$$

Thus $|e^{\lambda\tau}\, dS_1u/d\tau|$ is bounded for all $\lambda < \lambda_0$.

The continuity of S_1u follows in the same way with the single change that the inequality (B.7) replaces the inequality (B.2).

Appendix C.

TABLE 1
Tabulation of α_{pq} *(see* (2.11)*) where* $\alpha_{pq} = \alpha_{qp}$

p/q	0	1	2	3	4	5	6	7	8	9
0	0									
1	1	0								
2	0	$\frac{1}{4}$	0							
3	$\frac{3}{4}$	0	$\frac{1}{8}$	0						
4	0	$\frac{1}{8}$	0	$\frac{3}{64}$	0					
5	$\frac{5}{8}$	0	$\frac{5}{64}$	0	$\frac{3}{128}$	0				
6	0	$\frac{5}{64}$	0	$\frac{3}{128}$	0	$\frac{5}{512}$	0			
7	$\frac{35}{64}$	0	$\frac{7}{128}$	0	$\frac{7}{512}$	0	$\frac{5}{1024}$	0		
8	0	$\frac{7}{128}$	0	$\frac{7}{512}$	0	$\frac{5}{1024}$	0	$\frac{35}{16384}$	0	
9	$\frac{63}{128}$	0	$\frac{21}{512}$	0	$\frac{9}{1024}$	0	$\frac{45}{16384}$	0	$\frac{35}{32768}$	0

REFERENCES

[1] N. N. BOGOLIUBOV AND Y. A. MITROPOLSKY, *Asymptotic Methods in the Theory of Non-linear Oscillations*, Hindustan Publishing Corp., Delhi, India, 1961.

[2] D. S. COHEN AND J. P. KEENER, *Oscillatory processes in the theory of particulate formation in supersaturated chemical solutions*, SIAM J. Appl. Math., 28 (1975), pp. 307–318.

[3] ———, *Multiplicity and stability of oscillatory states in a continuous stirred tank reactor with exothermic consecutive reactions* $A \to B \to C$, Chem. Eng. Sci., 31 (1976), pp. 115–122.

[4] J. D. COLE, *Perturbation Methods in Applied Mathematics*, Blaisdell, Waltham, MA, 1968.

[5] E. L. INCE, *Ordinary Differential Equations*, Dover, New York, 1956.

[6] J. P. KEENER, *Perturbed bifurcation theory at multiple eigenvalues*, Arch. Rational Mech. Anal., 56 (1974), pp. 348–366.

[7] H. B. KELLER AND W. F. LANGFORD, *Iterations, perturbations and multiplicities in nonlinear bifurcation problems*, Ibid., 48 (1972), pp. 83–108.

[8] F. W. KOLLETT, *Two-timing methods valid on expanding intervals*, this Journal, 5 (1974), pp. 613–624.

[9] N. MINORSKY, *Introduction to Nonlinear Mechanics*, Edwards Brothers, Ann Arbor, MI, 1947.

[10] W. M. GREENLEE AND R. E. SNOW, *Two-timing on the half line for damped oscillation equations*, J. Math. Anal. Appl., 51 (1975), pp. 394–428.