MATHEMATICS OF COMPUTATION, VOLUME 26, NUMBER 118, APRIL 1972

Minimal Error Constant Numerical Differentiation (N.D.) Formulas

By A. Pelios and R. W. Klopfenstein

Abstract. In this paper, we consider a class of k-step linear multistep methods in the form (1.1) of numerical differentiation (N.D.) formulas. For each k, we have required the property of A-stability which implies at most second order for the associated operator. Among such second-order operators, the parameters of the formulas have been selected to minimize the error constant consistent with the A-stability property.

It is shown that the error constant approaches that of the trapezoidal rule as $k \to \infty$ and that significant reductions occur for quite modest k. Thus, these results have significance in practical applications.

I. Introduction. Consider the N.D. formula

$$hf_{n+1} = \sum_{m=1}^{k} a_{mk} \nabla^m y_{n+1}, \tag{1.1}$$

where $\nabla^m y_{n+1}$ is the backward difference operator. Henrici [4], Gear [3], and others select $a_{mk} = 1/m$ to obtain an operator of maximum order for a given number of back points. With this selection, (1.1) is of kth order with error constant

$$C_k = 1/(k + 1). \tag{1.2}$$

Such formulas are A-stable for $k = 1, 2$ and are stable on the negative real half-line for $k = 1, 2, \cdots, 6$.

In this note, we propose to retain the A-stability property by limiting the order to two [1] and use the additional degrees of freedom (from a_{mk}, $m > 2$) to obtain formulas of minimum error constant (for a given k) consistent with the A-stability property. This generalizes a result [5] previously obtained for the case $k = 3$. For every $k < \infty$, we must have

$$C_k > 1/12 = C_T, \tag{1.3}$$

where C_T is the error constant associated with the trapezoidal rule [1]. We shall exhibit minimal formulas for every k such that $C_k \to C_T^+$ as $k \to \infty$.

II. Stability Properties. To study the stability properties of (1.1), we apply it to the equation $y' = \nu y$ and obtain

$$qy_{n+1} = \sum_{m=1}^{k} a_{mk} \nabla^m y_{n+1}, \qquad q = \nu h. \tag{2.1}$$

By forming the characteristic equation through the substitution $y_{n+i} = \lambda^{k+i-1} y_{n+1-k}$,

Received June 30, 1971, revised October 29, 1971.

AMS 1969 *subject classifications.* Primary 6560, 6561; Secondary 6568.

Key words and phrases. Ordinary differential equations, numerical solution, multistep formulas, numerical stability.

we obtain

$$q\lambda^k = \sum_{m=1}^{k} a_{mk}\lambda^k(1 - \lambda^{-1})^m, \tag{2.2}$$

whence,

$$q = \sum_{m=1}^{k} a_{mk}\xi^m, \qquad \xi = 1 - \lambda^{-1}. \tag{2.3}$$

By Dahlquist's Lemma 2.1 [1], the N.D. algorithm (1.1) is A-stable if and only if

$$\mathrm{Re}(q) = \mathrm{Re} \sum_{m=1}^{k} a_{mk}\xi^m \geqq 0, \tag{2.4}$$

in the exterior of $|\lambda| = 1$, i.e., in the interior of the circle $|\xi - 1| = 1$. It is clear that it is necessary and sufficient for A-stability that (2.4) be satisfied on the circle $|\xi - 1| = 1$.

III. Minimal N.D. Formulas.

Definition 1. A N.D. formula (1.1) is a minimal N.D. formula for a given k, if it (a) is of second order, (b) is A-stable, and (c) has an error constant C_k equal to the lower bound of such error constants.

It is necessary and sufficient for a N.D. formula to be of second order that $a_{1k} = 1$, $a_{2k} = \frac{1}{2}$. Furthermore, the error constant $C_k = \frac{1}{3} - a_{3k}$. We therefore study polynomials of the form

$$f(x) = x + \tfrac{1}{2}x^2 + \tfrac{1}{6}x^3\Phi(x), \tag{3.1}$$

where Φ is zero or an arbitrary polynomial of degree $k - 3$. Now, we define a related polynomial

$$\begin{aligned} g(x) &= \frac{2xf(1 + x)}{(1 + x)^2} + \frac{1 - x}{1 + x}, \\ &= (1 + x)(1 + \tfrac{1}{3}x\Phi(1 + x)), \end{aligned} \tag{3.2}$$

and propose to examine its properties on the unit circle $|x| = 1$. The following lemma validates this procedure.

LEMMA 1. *$f(1 + x)$ is a polynomial of degree $N \geqq 2$ with properties*

$$\begin{aligned} &\mathrm{Re}\, f(1 + e^{i\phi}) \geqq 0, \qquad -\pi \leqq \phi \leqq \pi, \\ &f(0) = 0, \qquad f'(0) = 1, \qquad f''(0) = 1, \end{aligned} \tag{3.3}$$

if and only if $g(x)$ is a polynomial of degree $N - 1 \geqq 1$ with properties

$$\begin{aligned} &\mathrm{Re}\, g(e^{i\phi}) \geqq 0, \qquad -\pi \leqq \phi \leqq \pi, \\ &g(-1) = 0, \qquad g(0) = 1. \end{aligned} \tag{3.4}$$

This follows from a direct consideration of the properties of f and g deduced from the definitions (3.1) and (3.2). As a corollary, we have that $f'''(0)$ is maximized if and only if $g'(-1) = 1 - \frac{1}{3}\Phi(0)$ is minimized.

We can state the following lemma valid for a general function of real type, analytic within and on the unit circle. The proof is given in Appendix 1. By a function of real type, we mean a function that is real-valued for real arguments.

LEMMA 2. *If $g(z)$ is a function of real type, analytic within and on the unit circle with the property $g(-1) = 0$, then*

$$(3.5)\qquad \text{(a)}\quad g(0) = \frac{1}{\pi}\int_0^{\pi} \operatorname{Re} g(e^{i\phi})\,d\phi, \quad \textit{and}$$

$$\text{(b)}\quad g'(-1) = \frac{1}{\pi}\int_0^{\pi} \frac{\operatorname{Re} g(e^{i\phi})\,d\phi}{1+\cos\phi}.$$

Now, we note that $\operatorname{Re} g(e^{i\phi})$ may be expressed as a polynomial in $\cos\phi$ and such polynomials satisfying $g(-1) = 0$ are obtained through

$$(3.6)\qquad \operatorname{Re}(g(e^{i\phi})) = (1+\cos\phi)Q,$$

and equivalently

$$\operatorname{Re}(f(1+e^{i\phi})) = (1+\cos\phi)^2 Q,$$

where Q is a real polynomial in $\cos\phi$ nonnegative on $-1 \leqq \cos\phi \leqq 1$. The problem of obtaining minimal N.D. formulas can now be expressed through the following constrained minimization problem:

For a given $k \geqq 3$, find the polynomial $Q(\cos\phi)$ of degree $k-2$ satisfying

$$(3.7)\qquad \frac{1}{\pi}\int_0^{\pi} (1+\cos\phi)Q(\cos\phi)\,d\phi = 1, \qquad Q(\cos\phi) \geqq 0,$$

which minimizes $\int_0^{\pi} Q(\cos\phi)\,d\phi$. This is equivalent to the problem of maximizing the ratio

$$(3.8)\qquad R = \frac{\int_{-1}^{1} w(x)xQ(x)\,dx}{\int_{-1}^{1} w(x)Q(x)\,dx},$$

subject to $Q(x) \geqq 0$, $-1 \leqq x \leqq 1$, with $w(x) = 1/(1-x^2)^{1/2}$. Having obtained such a Q, an appropriate normalization can be obtained through (3.7) and the corresponding function $f(x)$ through (3.6) and (3.2), thus leading to the minimal N.D. formulas sought.

We now establish the following lemmas relating to the location of the zeros of $Q(x)$:

LEMMA 3. *The zeros of a polynomial $Q(x)$ of degree equal to or less than M which maximizes* (3.8) *are contained in the closed interval* $[-1, 1]$.

Proof. Suppose $Q(x)$ to be factored into

$$Q(x) = Q_0(x)Q_1(x),$$

$$Q_0(x) \geqq 0, \qquad -1 \leqq x \leqq 1,$$

$$Q_1(x) > 0,$$

where $Q_0(x)$ has as zeros all of the zeros of $Q(x)$ contained in $[-1, 1]$ and $Q_1(x)$ has as zeros the remaining zeros of $Q(x)$. Assume that

$$\text{Degree}(Q_0) < M, \tag{3.9}$$

and that $Q(x)$ maximizes the ratio R of (3.8). We now construct a perturbation polynomial and show that the ratio R of (3.8) can be increased under this assumption, thus contradicting the assumption that Q maximizes (3.8).

Consider

$$Q(\epsilon) = Q + \epsilon(x - a)Q_0, \tag{3.10}$$

where a is selected such that

$$\int_{-1}^{1} w(x)(x - a)Q_0 \, dx = 0, \tag{3.11}$$

which is clearly possible from the fact that $w > 0$ and $Q_0 \geqq 0$. Furthermore, $-1 \leqq a \leqq 1$. The degree of $Q(\epsilon)$ is less than or equal to M since by assumption the degree of Q_0 is less than M.

We now show that there exists a nontrivial range for ϵ which includes $\epsilon = 0$ and such that $Q(\epsilon)$ satisfies the constraint $Q(\epsilon) \geqq 0$. For such ϵ, we must have by (3.10) that

$$Q_1(x) + \epsilon(x - a) \geqq 0. \tag{3.12}$$

Since, by assumption, Q_1 has no zeros in $[-1, 1]$, it has a minimum

$$Q_{\min} = \min_{x \in [-1,1]} (Q_1(x)) > 0. \tag{3.13}$$

It suffices for (3.12) to take

$$|\epsilon| \leqq Q_{\min}/(1 + |a|). \tag{3.14}$$

The perturbed polynomial $Q(\epsilon)$ leaves the denominator of R unchanged while the numerator becomes

$$\begin{aligned} N(\epsilon) &= N(0) + \epsilon \int_{-1}^{1} w(x)x(x - a)Q_0(x) \, dx, \\ &= N(0) + \epsilon I. \end{aligned} \tag{3.15}$$

Now $I \neq 0$, since $I = 0$, together with (3.11), implies that

$$\int_{-1}^{1} w(x)(x - a)^2 Q_0(x) \, dx = 0,$$

which is clearly impossible for nontrivial Q_0. Hence, ϵ may be selected to increase $N(\epsilon)$ and thus the ratio R as well. Q.E.D.

COROLLARY 1. *$Q(x)$ has full degree equal to M.*

The proof of Corollary 1 is obtained in an exactly similar way to that for Lemma 3.

LEMMA 4. *The polynomial $Q(x)$ which maximizes* (3.8) *does not have a zero at $x = +1$.*

Proof. If Q does have a zero at $x = +1$, it can be written as

$$Q(x) = (1 - x)Q_0(x), \quad \text{where Degree}(Q_0) = M - 1 < M. \tag{3.16}$$

Defining the scalar product $u \cdot v$ for some weight w by

$$u \cdot v = \int_{-1}^{1} uvw \, dx, \tag{3.17}$$

$w(x) > 0$ (almost everywhere), and defining a real function $r(x)$ by

$$r^2 = Q_0(x), \tag{3.18}$$

which is clearly possible due to the nonnegative character of Q_0, the ratio (3.8) can be expressed as

$$R_1 = \frac{xr \cdot r - xr \cdot xr}{r \cdot r - xr \cdot r}, \tag{3.19}$$

where here w is as in (3.8). We shall now show that $R_1 < R_{-1}$, where R_{-1} corresponds to

$$Q(x) = (1 + x)Q_0(x), \tag{3.20}$$

which is also an admissible Q with

$$R_{-1} = \frac{xr \cdot r + xr \cdot xr}{r \cdot r + xr \cdot r}. \tag{3.21}$$

Now, the proposition $R_1 < R_{-1}$ is implied by the proposition

$$(xr \cdot r)^2 < (xr \cdot xr)(r \cdot r), \tag{3.22}$$

which is Schwarz's inequality [7, pp. 381–382]. The inequality is strict since xr and r are linearly independent. Q.E.D.

Lemmas 3 and 4 taken together characterize Q as a real polynomial, nonnegative in $[-1, 1]$, and with all of its zeros in $[-1, 1)$. This implies that for odd degree, equal to $2l + 1$, Q can be expressed as

$$Q(x) = (1 + x)q_l^2, \tag{3.23}$$

and for even degree, equal to $2l$, as

$$Q(x) = q_l^2, \tag{3.24}$$

where q_l is a polynomial of degree l with zeros in $[-1, 1)$.

The following discussion relates to both the even and odd degree cases and will be made specific in this regard later. Through the definitions of (3.23) and (3.24) and an appropriate choice of a weight function the optimization problem (3.8) is equivalent to an unconstrained maximization of the form

$$R = \frac{xq_l \cdot q_l}{q_l \cdot q_l}, \tag{3.25}$$

with respect to polynomials q_l of degree equal to or less than l. To this end, we define an orthogonal set of polynomials by the three-term recursion [2], [6, pp. 42–44]

$$P_{n+1} = (x - a_n)P_n - b_n P_{n-1}, \qquad n \geqq 1, \tag{3.26}$$

with

$$P_0 = 1, \qquad P_1 = x - a_0,$$

and the a_n, b_n defined through

$$xP_n \cdot P_n = a_n P_n \cdot P_n, \qquad P_n \cdot P_n = b_n P_{n-1} \cdot P_{n-1}.$$

Now, we define a linear operator T on the space E of all polynomials of degree equal to or less than l with the inner product of (3.17) by

$$Tu = \text{the projection of } xu \text{ on } E. \tag{3.27}$$

Now, let λ be a (real) root of P_{l+1}, and u given by

$$u = P_{l+1}/(x - \lambda) \in E. \tag{3.28}$$

Then,

$$\begin{aligned} Tu &= \mathrm{Proj}\left(\frac{xP_{l+1}}{x - \lambda}\right), \\ &= \mathrm{Proj}\left(\frac{\lambda}{x - \lambda} + 1\right)P_{l+1}, \\ &= \lambda u + \mathrm{Proj}(P_{l+1}), \\ &= \lambda u. \end{aligned} \tag{3.29}$$

Thus, the eigenvalues λ_k of T are the zeros of P_{l+1}, and q_l may be expanded in terms of the corresponding eigenvectors u_k to give

$$R = \frac{\sum \lambda_k a_k \, ||u_k||^2}{\sum a_k \, ||u_k||^2}, \tag{3.30}$$

which is clearly maximized by taking

$$q_l = P_{l+1}/(x - \lambda_{\max}), \tag{3.31}$$

where $\lambda_{\max}$ is the largest eigenvalue of T (zero of P_{l+1}). The theory of orthogonal polynomials guarantees that this choice is unique [6, pp. 44–47].

This discussion can be made specific to the odd (even) degree cases of (3.23) and (3.24) by taking

$$w(x) = \frac{1 + x}{\pi(1 - x^2)^{1/2}} \quad (\text{odd}), \quad \text{or} \quad w(x) = \frac{1}{\pi(1 - x^2)^{1/2}} \quad (\text{even}). \tag{3.32}$$

The corresponding orthogonal polynomials are Jacobi polynomials [6, p. 3]

$$\begin{aligned} S_n(\cos\phi) &= \frac{\cos(n + \frac{1}{2})\phi}{\cos(\frac{1}{2}\phi)} \quad (\text{odd}), \quad \text{or} \\ T_n(\cos\phi) &= \cos(n\phi) \quad (\text{even}). \end{aligned} \tag{3.33}$$

In both cases, these lead to the polynomial Q given by

$$Q(\cos\phi) = \frac{1 + \cos(k\phi)}{(\cos\phi - \cos(\pi/k))^2}, \tag{3.34}$$

within a positive multiplicative factor.

By (3.6),

$$\mathrm{Re}\, f(1 + e^{i\phi}) = \alpha\left[\frac{1 + \cos\phi}{\cos\phi - \cos(\pi/k)}\right]^2 (1 + \cos(k\phi)), \qquad \alpha > 0. \tag{3.35}$$

Now, specification of the real part of an analytic function of real type on a circle specifies the function [7, p. 124]. We can construct this function by using the substitution $\cos \phi = z + 1/z$ in the first factor of (3.35) and writing

$$f(1 + z) = \alpha\left\{\frac{(z + 1)^4(1 + z^k)}{(z^2 - 2\cos(\pi/k)z + 1)^2} + \beta\frac{z^2 - 1}{z^2 - 2\cos(\pi/k)z + 1}\right\}, \tag{3.36}$$

where the contribution of the second term is purely imaginary on the unit circle and β may be selected to remove the two simple poles of the first term. When this is done and α is selected by the requirement $f'(0) = 1$, we obtain

$$\begin{aligned} f(1 + z) &= (1 + \cos(\pi/k))\frac{1 - z^2}{1 - 2\cos(\pi/k)z + z^2} \\ &\quad + \frac{1 - \cos(\pi/k)}{k}\frac{(1 + z)^4(1 + z^k)}{(1 - 2\cos(\pi/k)z + z^2)^2}, \qquad k \geqq 2, \end{aligned} \tag{3.37}$$

which is a polynomial of degree k in z which holds for $k \geqq 2$, even though the preceding development was carried out for $k \geqq 3$.

We now have sufficient information to state the following:

THEOREM. *The minimal N.D. formulas of Definition* 1 *are specified by the kth degree polynomials* (3.37) *for* $k \geqq 2$. *These polynomials exhibit directly the coefficients* a_{mk} *through the substitution* $z = \xi - 1$ *and comparison with* (2.3). *The error constant of the minimal N.D. formulas is given by*

$$C_k = \frac{2 - \cos(\pi/k)}{6[1 + \cos(\pi/k)]}, \qquad k \geqq 2. \tag{3.38}$$

IV. Discussion. In Table 1, we show the coefficients of some minimal N.D. formulas together with their associated error constants. It is seen that significant reductions in the error constant are obtained by taking k greater than two. In particular, taking k equal to 3, results in a reduction of the error constant by a factor of two while requiring no additional storage if a second-order extrapolative predictor is used in conjunction with (1.1).

TABLE 1

Minimal N.D. Formulas and Associated Error Constants

k	$12C_k$	a_{3k}	a_{4k}	a_{5k}
2	4	0	0	0
3	2	$\frac{1}{6}$	0	0
4	1.5148. . .	$\frac{1}{2}(\sqrt{2} - 1)$	$\frac{1}{8}(2 - \sqrt{2})$	0
5	1.3167. . .	$\frac{c}{2(1 + c)}$	$\frac{2c - 1}{4(1 + c)}$	$\frac{1}{5}(1 - c)$

Note: $a_{1k} = 1$ *and* $a_{2k} = \frac{1}{2}$ *for all* $k \geqq 2$. $c = \cos(\pi/5)$.

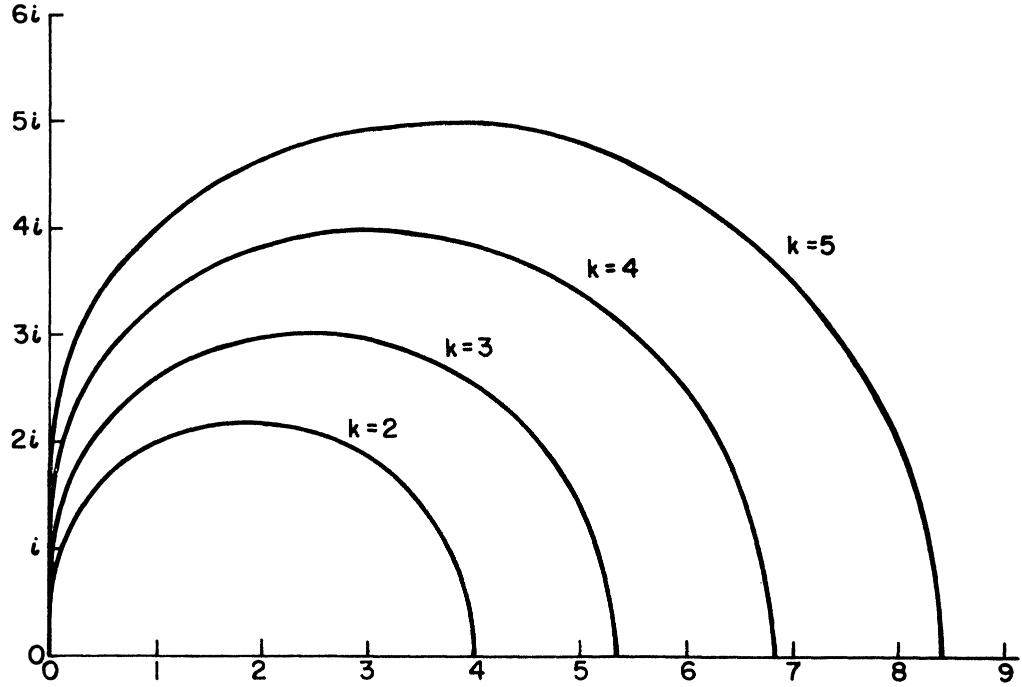

FIGURE 1. *Regions of Stability for Minimal N.D. Formulas*

The boundary for the stable regions of the algorithms tabulated in Table 1 are displayed in Fig. 1. As $k \to \infty$, the error constant approaches that of the trapezoidal rule from above with

$$C_k \sim \frac{1}{12} + \frac{1}{16}\left(\frac{\pi}{k}\right)^2 + O\left(\frac{1}{k^4}\right), \tag{4.1}$$

and the boundary of the stable region approaches the imaginary axis from the right.

Appendix 1. *Proof of Lemma* 2. Part (a) of Lemma 2 follows directly from an application of Cauchy's integral formula

$$g(0) = \frac{1}{2\pi i}\oint \frac{g(\zeta)}{\zeta}\, d\zeta, \tag{A1.1}$$

specialized to the unit circle, $\zeta = e^{i\phi}$, as an integration path leading to

$$g(0) = \frac{1}{2\pi}\int_0^{2\pi} g(e^{i\phi})\, d\phi, \tag{A1.2}$$

whence,

$$g(0) = \frac{1}{\pi}\int_0^{\pi} \mathrm{Re}\, g(e^{i\phi})\, d\phi. \tag{A1.3}$$

To establish part (b) of Lemma 2, we consider the function

$$h(\zeta) = \frac{g(\zeta) - g'(-1)(1 + \zeta)}{(1 + \zeta)^2}, \tag{A1.4}$$

and note that $h(\zeta)$ is analytic within and on the unit circle by assumption on the properties of $g(\zeta)$ and since the numerator of $h(\zeta)$ has a double zero at $\zeta = -1$.

Cauchy's integral theorem states that

$$\oint h(\zeta)\, d\zeta = 0, \tag{A1.5}$$

which when specialized to the unit circle, $\zeta = e^{i\phi}$, leads to

$$\int_0^{2\pi} \frac{\operatorname{Re} g(e^{i\phi}) - g'(-1)(1 + \cos\phi)}{2(1 + \cos\phi)}\, d\phi + i \int_0^{2\pi} \frac{\operatorname{Im} g(e^{i\phi}) - g'(-1)\sin\phi}{2(1 + \cos\phi)}\, d\phi = 0. \tag{A1.6}$$

Since $g(z)$ is of real type, $g(\bar{z}) = \bar{g}(z)$, and the second term of (A1.6) vanishes, so that (A1.6) may be written as

$$g'(-1) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\operatorname{Re} g(e^{i\phi})}{1 + \cos\phi}\, d\phi, \tag{A1.7}$$

which is equivalent to part (b) of Lemma 2.

RCA Laboratories
David Sarnoff Research Center
Princeton, New Jersey 08540

1. G. Dahlquist, "A special stability problem for linear multistep methods," *Nordisk Tidskr. Informationsbehandling,* v. 3, 1963, pp. 27–43. MR **30** #715.
2. G. E. Forsythe, "Generation and use of orthogonal polynomials for data-fitting with a digital computer," *J. Soc. Indust. Appl. Math.,* v. 5, 1957, pp. 74-88. MR **19,** 1079.
3. C. W. Gear, *The Automatic Integration of Stiff Ordinary Differential Equations. (With Discussion),* Proc. IFIP Congress, Information Processing (Edinburgh, 1968), Vol. 1: Mathematics, Software, North-Holland, Amsterdam, 1969, pp. 187–193. MR **41** #4808.
4. P. Henrici, *Discrete Variable Methods in Ordinary Differential Equations,* Wiley, New York, 1962, pp. 206-209. MR **24** #B1772.
5. R. W. Klopfenstein, "Numerical differentiation formulas for stiff systems of ordinary differential equations," *RCA Review,* v. 32, 1971, pp. 447–462.
6. G. Szegö, *Orthogonal Polynomials,* 2nd rev. ed., Amer. Math. Soc. Colloq. Publ., vol. 23, Amer. Math. Soc., Providence, R.I., 1959. MR **21** #5029.
7. E. C. Titchmarsh, *The Theory of Functions,* 2nd ed., Oxford Univ. Press, Oxford, 1939.