Computational Techniques Based on the Lanczos Representation

By J. N. Lyness*

Abstract. In his book *Discourse on Fourier Series*, Lanczos deals in some detail with representations of $f(x)$ of the type $f(x) = h_{p-1}(x) + g_p(x)$ where $h_{p-1}(x)$ is a polynomial of degree $p - 1$ and $g_p(x)$ has the property that its full range Fourier coefficients converge at the rate r^{-p}.

In Part I, some properties of $h_p(x)$ and of the series $\{h_p(x)\}_1^\infty$ are described. These properties are used here to provide criteria for the convergence or divergence of the Euler-Maclaurin series, in the case when $f(x)$ is an analytic function. The similarities and differences between this series and the Lidstone and other two-point series are briefly mentioned.

In Part II, the Lanczos representation is employed to derive an approximate representation $F(x)$ for an analytic function $f(x)$ on the interval [0, 1] is derived. This has the form

$$F(x) = \sum_{q=1}^{p-1} \lambda_{q-1} B_q(x)/q! + 2 \sum_{r=0}^{m/2} (\mu_r \cos 2\pi rx + \nu_r \sin 2\pi rx)$$

and requires for its determination the values of the derivatives $f^{(q-1)}(1) - f^{(q-1)}(0)$ ($q = 1$, $2, \cdots p - 1$) and the regularly spaced function values $f(j/m)$ ($j = 0, 1, \cdots, m$). It involves replacing $g_p(x)$ by a discrete Fourier expansion based on trapezoidal rule approximations to its Fourier coefficients.

This representation is a powerful one. The drawback is that it requires derivatives. Most of Part II is devoted to the effect of using only approximate derivatives. It is shown that when these are successively less accurate with increasing order (the sort of behaviour encountered using finite difference formula), then the representation is still powerful and reliable. In a computational context the only penalty for using inaccurate derivatives is that a larger value of m may—or may not—be required to attain a specific accuracy.

PART I.** PROPERTIES OF THE SEQUENCE $h_p(x)$

1. The Lanczos Representation. In this section, we outline a derivation of what we term Lanczos' representation for a function $f(x)$. We suppose that $f(x)$ is an analytic function of x and is real valued when x is real. For convenience, we suppose that $f(x)$ is analytic in a region of the complex plane which contains the unit interval [0, 1], a restriction which we denote by

(1.1) $$f(x) \in AR[0, 1].$$

Most of the results require only that

81

(1.2) $$f(x) \in C^{(p)}[0, 1]$$

for some self-evident value of p.

A well-known representation over the unit interval is the Fourier series

(1.3) $$\bar{f}(x) = If + 2 \sum_{r=1}^{\infty} C^{(r)}f \cos 2\pi rx + S^{(r)}f \sin 2\pi rx.$$

The coefficients which occur in this series (Fourier coefficients) are given by

(1.4) $$If \equiv C^{(0)}f \equiv \int_0^1 f(x)\, dx$$

and by the real and imaginary parts of

(1.5) $$C^{(r)}f + iS^{(r)}f = \int_0^1 f(x)e^{2\pi irx}\, dx.$$

In cases in which $f(x)$ is a function of several variables, we indicate the integration variable explicitly by writing for example

(1.6) $$C_x^{(r)}(f(x)) = C^{(r)}f.$$

It is well known that within the open interval $(0, 1)$ the Fourier series coincides with $f(x)$. Specifically,

$$\bar{f}(x) = f(x), \qquad 0 < x < 1,$$
(1.7) $$\bar{f}(1) = \bar{f}(0) = \tfrac{1}{2}(f(1) + f(0)),$$
$$\bar{f}(x + 1) = \bar{f}(x) \quad \text{all} \quad x.$$

The Fourier series (1.3) is a particularly useful tool when $f(x)$ is a periodic function of period 1. In this case, the series converges reasonably rapidly. For example, if there are no singularities within a distance d of the real axis, the asymptotic behaviour of the Fourier coefficients is bounded by an inequality of the form

(1.8) $$|C^{(r)}f + iS^{(r)}f| < Ke^{-2\pi rd}.$$

However, if $f(x)$ is not periodic, the use of the Fourier series (1.3) has two related drawbacks of a computational nature. First is the circumstance that the series may converge very slowly. For example, when $f(1) \neq f(0)$, we have

(1.9) $$S^{(r)}f \simeq O(r^{-1}) \quad \text{as } r \to \infty$$

and, unless $f'(1) = f'(0)$,

(1.10) $$C^{(r)}f \simeq O(r^{-2}) \quad \text{as } r \to \infty.$$

If a truncated form of the Fourier series is used for numerical approximation, a large number of terms have to be retained to obtain a required accuracy. The second related drawback is the fact that methods for evaluating the integrals in (1.5) numerically are cumbersome. For the larger values of r, the integrands are more rapidly oscillatory and so more difficult to evaluate numerically.

A modified representation, based on the Fourier series, has been discussed by Lanczos [4]. This involves expressing $f(x)$ in the form

(1.11) $$f(x) = h_{p-1}(x) + g_p(x)$$

where $h_{p-1}(x)$ is a polynomial of degree $p - 1$ and $g_p(x)$ has Fourier coefficients which satisfy the order relation

(1.12) $$C^{(r)}g_p + iS^{(r)}g_p \simeq O(r^{-p}) \quad \text{as } r \to \infty.$$

A convenient method for obtaining this representation is by means of the Fourier Coefficient Asymptotic Expansion (FCAE). This may be derived using integration by parts and has the form

(1.13) $$C^{(r)}f + iS^{(r)}f = \sum_{q=1}^{p-1} \frac{(-1)^{q-1}(f^{(q-1)}(1) - f^{(q-1)}(0))}{(2\pi ir)^q} + C_p^{(r)}f + iS_p^{(r)}f,$$

where the remainder term has the integral representation

(1.14) $$C_p^{(r)}f + iS_p^{(r)}f = \frac{(-1)^p}{(2\pi ir)^p} \int_0^1 f^{(p)}(t)(e^{2\pi irt} - 1)\, dt.$$

Clearly,

(1.15) $$C_p^{(r)}f + iS_p^{(r)}f \simeq O(r^{-p}) \quad \text{as } r \to \infty.$$

The expressions (1.13) for $C^{(r)}f$ and $S^{(r)}f$ may be substituted into the Fourier series (1.3). The summations over index r may be expressed in terms of Bernoulli functions, defined by

(1.16) $$\frac{\bar{B}_q(x)}{q!} = -2 \sum_{r=1}^{\infty} \operatorname{Re} \frac{e^{2\pi irx - i\pi q/2}}{(2\pi r)^q}.$$

This follows the standard notation used for example in Abramowitz and Stegun [1, pp. 803 et seq.]. The result is

(1.17) $$\bar{f}(x) = If + \sum_{q=1}^{p-1} (f^{(q-1)}(1) - f^{(q-1)}(0)) \frac{\bar{B}_q(x)}{q!} + 2 \sum_{r=1}^{\infty} C_p^{(r)}f \cos 2\pi rx$$
$$+ 2 \sum_{r=1}^{\infty} S_p^{(r)}f \sin 2\pi rx.$$

The function $\bar{B}_q(x)$ coincides with the Bernoulli polynomial $B_q(x)$ in the open interval $(0, 1)$ for all values of q and in the closed interval $[0, 1]$ for all values of q other than $q = 1$. Since

(1.18) $$B_1(x) = x - \tfrac{1}{2}; \qquad \bar{B}_1(0) = \bar{B}_1(1) = 0,$$

it follows from (1.7) that

(1.19) $$f(x) - \bar{f}(x) = (f(1) - f(0))(B_1(x) - \bar{B}_1(x)), \qquad 0 \le x \le 1,$$

and, adding this into (1.17), we find

(1.20) $$f(x) = If + \sum_{q=1}^{p-1} (f^{(q-1)}(1) - f^{(q-1)}(0)) \frac{B_q(x)}{q!} + 2 \sum_{r=1}^{\infty} C_p^{(r)}f \cos 2\pi rx$$
$$+ 2 \sum_{r=1}^{\infty} S_p^{(r)}f \sin 2\pi rx, \qquad 0 \le x \le 1.$$

The sum over index q is clearly a polynomial of degree $p - 1$ in x. In many applica-

tions, it is convenient to define $h_{p-1}(x)$ in a manner which includes If. In this particular application, we include If in the definition of $g_p(x)$. Thus, we define

$$(1.21) \qquad h_{p-1}(x) = \sum_{q=1}^{p-1} (f^{(q-1)}(1) - f^{(q-1)}(0)) B_q(x)/q!$$

and

$$(1.22) \quad g_p(x) = f(x) - h_{p-1}(x) = If + 2 \sum_{r=1}^{\infty} C_p^{(r)} f \cos 2\pi r x + 2 \sum_{r=1}^{\infty} S_p^{(r)} f \sin 2\pi r x.$$

This gives the Fourier coefficients of $g_p(x)$ directly. Thus

$$(1.23) \qquad \begin{aligned} Ig_p &= If, \\ C^{(r)} g_p + i S^{(r)} g_p &= C_p^{(r)} f + i S_p^{(r)} f, \end{aligned}$$

and, in view of order relation (1.15), these Fourier coefficients are of order $O(r^{-p})$ as r becomes infinite. We state these results as a theorem.

THEOREM 1.24.   If $f(x) \in C^{(p)}[0, 1]$ then

$$f(x) = h_{p-1}(x) + g_p(x),$$

where $h_{p-1}(x)$ is the polynomial of degree $p - 1$ given by (1.21) above and $g_p(x)$ is a function whose Fourier coefficients satisfy the order relation

$$C^{(r)} g_p \simeq O(r^{-p}), \qquad S^{(r)} g_p \simeq O(r^{-p}) \quad \text{as } r \to \infty.$$

We conclude this section by stating some of the more obvious properties of $h_{p-1}(x)$ and $g_p(x)$. If we denote by $B_q^{(s)}(x)$ the sth derivative of the function $B_q(x)$ (and not a Bernoulli function of the sth kind), it is simple to show that

$$(1.25) \qquad \begin{aligned} B_q^{(s)}(1) - B_q^{(s)}(0) &= q! \quad \text{when } s = q - 1, \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

From this, it may be shown that

$$Ig_p = If, \qquad Ih_{p-1} = 0,$$

$$(1.26) \qquad \begin{aligned} g_p^{(r)}(1) - g_p^{(r)}(0) &= 0, & r &= 0, 1, \cdots, p - 2, \\ &= f^{(r)}(1) - f^{(r)}(0), & r &= p - 1, p, \cdots, \\ h_{p-1}^{(r)}(1) - h_{p-1}^{(r)}(0) &= f^{(r)}(1) - f^{(r)}(0), & r &= 0, 1, \cdots, p - 2, \\ &= 0, & r &= p - 1, p, \cdots. \end{aligned}$$

Given that $h_{p-1}(x)$ is a polynomial of degree $p - 1$, these relations provide enough information to completely specify $h_{p-1}(x)$. The relations which involve $g_p(x)$ may also be used in conjunction with the FCAE (1.13) applied to $g(x)$ to show the order relation (1.24) directly. It also follows from (1.14) and (1.22) that the function $g_p(x)$ has the integral representations

$$(1.27) \qquad \begin{aligned} g_p(x) &= If + \int_0^1 f^{(p)}(t) \frac{\bar{B}_p(x) - \bar{B}_p(x - t)}{p!} \, dt \\ &= If - \int_0^1 f^{(p-1)}(t) \frac{\bar{B}_{p-1}(x - t)}{(p - 1)!} \, dt. \end{aligned}$$

Certain special classes of functions give rise to particular cases. If $f(x)$ is a polynomial of degree $d \leq p - 1$, then

$$(1.28) \qquad\qquad\qquad g_p(x) = If$$

is a constant. If $f(x)$ is a periodic function of period 1 and so

$$(1.29) \qquad\qquad\qquad f(x + 1) = f(x) \quad \text{all } x,$$

then

$$(1.30) \qquad\qquad\qquad h_{p-1}(x) = 0, \qquad f(x) = g_p(x).$$

In this case, the theory given above is of course formally correct, but not of particular interest.

2. Convergence of the Sequence $h_p(x)$. In Part II, we discuss a method for the implementation of this representation. In this section, we consider the convergence properties of the sequence of polynomials

$$(2.1) \qquad\qquad\qquad If + h_p(x), \qquad p = 1, 2, 3, \cdots .$$

We have noted that, when $f(x)$ is a polynomial of degree d, all members of the sequence after the dth are identically equal to $f(x)$. Thus it is of interest to ask to what extent it is the case that the infinite sequence converges, and if it does converge, whether or not it converges to $f(x)$.

The question of whether or not the infinite series

$$(2.2) \qquad\qquad \sum_{q=1}^{\infty} (f^{(q-1)}(1) - f^{(q-1)}(0)) B_q(x)/q!, \qquad 0 \leq x \leq 1,$$

converges depends on the rate of growth of the Bernoulli polynomials and of the derivatives of $f(x)$. So far as the Bernoulli polynomials are concerned, we rely on the following well-known limits

$$(2.3) \qquad \lim_{q \to \infty} \left| \frac{\bar{B}_{2q}(x)}{(2q)!} (2\pi)^{2q} \right| = 2 |\cos 2\pi x|, \qquad 0 \leq x \leq 1,$$

$$(2.4) \qquad \lim_{q \to \infty} \left| \frac{\bar{B}_{2q}(x)}{(2q)!} (2\pi)^{4q} \right| = \tfrac{1}{2}, \qquad\qquad x = \tfrac{1}{4}, \tfrac{3}{4},$$

$$(2.5) \qquad \lim_{q \to \infty} \left| \frac{\bar{B}_{2q+1}(x)}{(2q+1)!} (2\pi)^{2q+1} \right| = 2 |\sin 2\pi x|, \qquad 0 \leq x \leq 1,$$

$$(2.6) \qquad\qquad \bar{B}_{2q+1}(x) = 0, \qquad\qquad x = 0, \tfrac{1}{2}, 1.$$

These are either given in Abramowitz and Stegun [1, pp. 803 et seq.], or are simple deductions from formulas given there. The coefficients involving the derivatives depend on $f(x)$ only through the derivatives of

$$(2.7) \qquad\qquad\qquad \varphi(x) = f(x + 1) - f(x)$$

evaluated at $x = 0$. Since $f(x)$ is analytic in a region containing the unit interval [0, 1] it follows that $\varphi(x)$ is analytic in a region which contains the origin. The rate of increase of the derivatives $\varphi^{(n)}(0)$ as n becomes infinite is closely connected with the behaviour of $\varphi(z)$ as z becomes large. If $\varphi(z)$ is not an entire function, and its Taylor

series about $z = 0$ has a finite radius of convergence ρ, then

$$(2.8) \qquad \limsup_{n \to \infty} |\varphi^{(n)}(0)/n!|^{1/n} = 1/\rho.$$

Using the limiting form of Stirling's formula

$$(2.9) \qquad \lim_{n \to \infty} (e/n)(n!)^{1/n} = 1,$$

it follows that

$$(2.10) \qquad \limsup_{n \to \infty} (e/n)|\varphi^{(n)}(0)|^{1/n} = 1/\rho.$$

On the other hand, if $\varphi(z)$ is an entire function, the rate of increase of the derivatives $\varphi^{(n)}(0)$ is less rapid than this and may be expressed in terms of the order μ and the type τ of the entire function $\varphi(z)$. The theory of entire functions is described in some detail in Boas [2]. The order-type classification of entire functions depends on the functional

$$(2.11) \qquad M(r; \varphi) = \max_{|z| = r} |\varphi(z)|.$$

The entire function $\varphi(z)$ is of order μ if

$$(2.12) \qquad \limsup_{r \to \infty} \frac{\log \log M(r; \varphi)}{\log r} = \mu,$$

and if $0 < \mu < \infty$, it is of type τ if

$$(2.13) \qquad \limsup_{r \to \infty} r^{-\mu} \log M(r; \varphi) = \tau.$$

Thus the function

$$(2.14) \qquad f(z) = p(z)e^{\tau z^{\mu}},$$

where $p(z)$ is any polynomial and μ is an integer of order μ and type τ. The three functions

$$p(z), \quad p(z)p(e^{z})e^{\cos 2\pi z}, \quad e^{3\sqrt{z}} + e^{-3\sqrt{z}}$$

are of orders $\mu = 0$, $\mu = \infty$ and $\mu = \frac{1}{2}$, respectively. In the first two cases, the type is not defined. In the third case, $\tau = 3$. Finally, a class of functions which frequently occurs is defined as follows: An entire function is of *exponential type* τ if it is of order $\mu < 1$ or if it is of order $\mu = 1$ and type $\leq \tau$.

The rate of growth of derivatives $\varphi^{(n)}(0)$ is related to the order and the type of $\varphi(z)$ in a natural manner. In comparing this growth for two different functions, the one with highest order has the highest ultimate rate of growth. If the two functions have the same order, the one of highest type has the highest rate of growth. These remarks apply to zero and infinite orders and types when they are defined. If μ and τ are both finite, the results which correspond to (2.8) and (2.10) are

$$(2.15) \qquad \limsup_{n \to \infty} n \, |\varphi^{(n)}(0)/n!|^{\mu/n} = e\mu\tau$$

and

(2.16)
$$\limsup_{n\to\infty} (e/n)^{\mu-1} |\varphi^{(n)}(0)|^{\mu/n} = \mu\tau.$$

We now return to the series (2.2) in which we are interested. It is convenient to decompose $\varphi(x)$ into its even and odd parts, defined by

(2.17)
$$\varphi_{\pm}(x) = \tfrac{1}{2}(\varphi(x) \pm \varphi(-x)).$$

Then

(2.18)
$$\varphi_{+}^{(2q)}(0) = \varphi^{(2q)}(0), \qquad \varphi_{+}^{(2q+1)}(0) = 0,$$
$$\varphi_{-}^{(2q)}(0) = 0, \qquad\qquad \varphi_{-}^{(2q+1)}(0) = \varphi^{(2q+1)}(0),$$

and the infinite series (2.2) may be written

(2.19)
$$\sum_{n=0}^{\infty} \varphi_{-}^{(n)}(0) \frac{B_{n+1}(x)}{(n+1)!} + \sum_{n=0}^{\infty} \varphi_{+}^{(n)}(0) \frac{B_{n+1}(x)}{(n+1)!},$$

displaying the symmetric (about $x = \tfrac{1}{2}$) and the antisymmetric parts separately. (Alternate terms in each of these two summations are zero in view of (2.18).) We apply the Cauchy convergence criterion separately to each sum. We set

(2.20)
$$u_n = \varphi_{-}^{(n)}(0) \frac{B_{n+1}(x)}{(n+1)!}, \qquad v_n = \varphi_{+}^{(n)}(0) \frac{B_{n+1}(x)}{(n+1)!}$$

and evaluate $\limsup |u_n|^{1/n}$ and $\limsup |v_n|^{1/n}$. It is necessary to make several evaluations of this nature to cover the cases in which both $\varphi_{+}(x)$ and $\varphi_{-}(x)$ are not entire or are entire of order μ and type τ. The series u_n involves Bernoulli functions of even order and a different result is obtained when $x = \tfrac{1}{4}$ or $\tfrac{3}{4}$ because (2.3) leads to indeterminate results and (2.4) has to be used instead. We give details below of one limit evaluation, the case of u_n with $x \neq \tfrac{1}{4}$ or $\tfrac{3}{4}$.

If $\varphi_{-}(z)$ has a finite radius of convergence ρ we rewrite (2.20) in the form

(2.21)
$$|u_n|^{1/n} = \left[\frac{e}{n} |\varphi_{-}^{(n)}(0)|^{1/n}\right]\left[\left\{\frac{|B_{n+1}(x)|}{(n+1)!}(2\pi)^{n+1}\right\}^{1/n}\right]\frac{n}{e}\frac{1}{(2\pi)^{1+1/n}}.$$

In view of (2.10), the superior limit of the first term in square brackets is $1/\rho$. In view of (2.3), the limit of the second term in square brackets is 1. The remaining terms become infinite with increasing n. Thus

(2.22)
$$\limsup_{n\to\infty} |u_n|^{1/n} = \infty$$

and the Cauchy convergence criterion indicates that the series u_n diverges.

If $\varphi_{-}(x)$ is an entire function of order μ and type τ (both being finite), we express $|u_n|^{1/n}$ in the form

(2.23)
$$|u_n|^{1/n} = \left[\left\{\left(\frac{e}{n}\right)^{\mu-1} |\varphi^{(n)}(0)|^{\mu/n}\right\}^{1/\mu}\right]\left[\left\{\frac{|B_{n+1}(x)|}{(n+1)!}(2\pi)^{n+1}\right\}^{1/n}\right]$$
$$\times \frac{1}{(2\pi)^{1+1/n}}\left(\frac{n}{e}\right)^{1-1/\mu}.$$

Again, the limit superior of the first term and the limit of the second term are finite and their product is $\mu\tau$. Thus

$$\lim_{n\to\infty} \sup |u_n|^{1/n} = 0, \qquad \mu < 1,$$

$$(2.24) \qquad\qquad\qquad = \tau/2\pi, \qquad \mu = 1,$$

$$\qquad\qquad\qquad = \infty, \qquad \mu > 1.$$

The series $\sum u_n$ converges if this limit is less than 1 and diverges if this limit is greater than 1. We note that $\sum u_n$ is only the symmetric part of the series (2.2) and that this particular result is invalid when $x = \frac{1}{4}, \frac{3}{4}$, since then the second term in square brackets in (2.23) does not have the limit 1. However, we may state that, if $\varphi_-(x)$ is of exponential type $\tau < 2\pi$ and $x \neq \frac{1}{4}, \frac{3}{4}$, then the symmetric part converges while, if $\varphi_-(x)$ is not of exponential type $\tau \leqq 2\pi$ and $x \neq \frac{1}{4}, \frac{3}{4}$, then the symmetric part diverges.

The results of carrying out a similar investigation in each of the several cases involved lead to the following theorem.

THEOREM 2.25. *If $f(x)$ is analytic in a region containing the interval* [0, 1], *the convergence or divergence of the series*

$$If + \sum_{q=1}^{\infty} (f^{(q-1)}(1) - f^{(q-1)}(0))B_q(x)/q!$$

depends on the nature of the even and odd parts $\varphi_+(x)$ and $\varphi_-(x)$ defined by (2.17) of the function $\varphi(x) = f(x + 1) - f(x)$ in the following manner. Following Table 2.26, for $x \in X$, the series converges if $F(x)$ is an entire function of exponential type $\tau < T$ and diverges if $F(x)$ is not an entire function of exponential type $\tau \leqq T$.

TABLE 2.26

X	$F(x)$	T
$4x \neq$ integer	$\varphi(x)$	2π
$4x =$ even integer	$\varphi_-(x)$	2π
$4x =$ odd integer	$\begin{cases}\varphi_-(x) \\ \varphi_+(x)\end{cases}$	$\begin{matrix}2\pi \\ 4\pi\end{matrix}$

In the final case, both conditions must be satisfied for convergence but either condition gives rise to divergence.

The proof, given in part above, does not cover the cases in which μ or τ is infinite. However, it is clear that the condition for convergence depends on the rate of growth of $\varphi^{(n)}(0)$ with increasing n, which in turn depends on the value of μ (or if μ is kept fixed the value of τ) in a monotonic manner. Thus, the only functions for which the theorem above does not give a specific answer to the question of whether or not the series converges are those for which $F(x)$ is precisely of order 1 and type T. The foregoing discussion determines the conditions under which the series obtained by letting p become infinite in

$$If + \sum_{q=1}^{p-1} (f^{(q-1)}(1) - f^{(q-1)}(0))B_q(x)/q!$$

converges or diverges. However, it does not provide the information that, when there is convergence, the limit is $f(x)$. In fact, this need not be the case. If $f(x)$ is a periodic function, i.e., one for which

$$(2.27) \qquad\qquad\qquad f(x + 1) = f(x),$$

then $\varphi(x)$ is identically zero and so each term in the series is zero. The relation for finite p is of course satisfied, that is

$$(2.28) \qquad\qquad f(x) = h_{p-1}(x) + g_p(x),$$

but

$$(2.29) \qquad\qquad \lim_{p\to\infty} (If + h_{p-1}(x)) = If$$

which is not identical with $f(x)$ unless $f(x)$ is constant. We now prove a theorem which gives a sufficient (but not necessary) condition for convergence to the appropriate result.

THEOREM 2.30.   *If $f(x)$ is an entire function of exponential type $\tau < 2\pi$, then*

$$\lim_{p\to\infty} (If + h_p(x)) = f(x), \qquad 0 < x < 1.$$

Proof.   Since

$$(2.31) \qquad\qquad f(x) = (If + h_{p-1}(x)) + (g_p(x) - If),$$

it is sufficient to show that the final term approaches zero as p becomes infinite. Using integral representation (1.27), we find

$$(2.32) \qquad \begin{aligned} |g_{p+1}(x) - If| &= \left| -\int_0^1 f^{(p)}(t)\, \frac{\bar{B}_p(x-t)}{p!}\, dt \right| \\ &< \max_{0\leq x\leq 1} |\bar{B}_p(x)/p!|\, M_p, \end{aligned}$$

where M_p is the maximum of $|f^{(p)}(t)|$ in the range $0 \leq t \leq 1$. In view of Eqs. (2.3)–(2.6), there exists a number K, independent of p, for which

$$(2.33) \qquad\qquad |\bar{B}_p(x)/p!| < K/(2\pi)^p.$$

If $f(x)$ is an entire function of order μ and type τ, then $g(x) = f(x + c)$ is also such a function. In view of this, Eq. (2.16) is satisfied with $\varphi^{(n)}(0)$ replaced by any $f^{(n)}(x)$ with $0 \leq x \leq 1$ and hence also with $\varphi^{(n)}(0)$ replaced by

$$(2.34) \qquad\qquad M_n = \max_{0\leq x\leq 1} |f^{(n)}(x)|.$$

Thus, in view of (2.16),

$$(2.35) \qquad\qquad \limsup_{n\to\infty} (e/n)^{\mu-1} M_n^{\mu/n} = \mu\tau.$$

Consequently, for all $\epsilon > 0$, there exists a number $K(\epsilon)$ which satisfies the inequality

$$(2.36) \qquad M_n < K(\epsilon)[(\mu\tau + \epsilon)(n/e)^{\mu-1}]^{n/\mu}, \qquad \epsilon > 0.$$

Substituting this into (2.32) gives

$$(2.37) \qquad |g_{p+1}(x) - If| < K \cdot K(\epsilon)((\tau + \epsilon)/2\pi)^p, \qquad \mu = 1,$$

$$(2.38) \qquad |g_{p+1}(x) - If| < K(K'(\epsilon, \tau, \mu))^p (p^p)^{1-1/\mu}, \qquad 0 < \mu < 1.$$

In the second case, the limit, as p becomes infinite, is zero, independently of the value of τ. In the first case, with $\mu = 1$ and $\tau < 2\pi$, we may choose $\epsilon = (\tau + 2\pi)/2$ and so the limit, as p becomes infinite, is also zero. In view of (2.31), this establishes Theorem 2.30.

3. The Convergence of the Euler-Maclaurin Series. The Lanczos representation is in fact a special case of the Euler-Maclaurin summation formula (3.5) below. Equation (1.20) may be thought of as expressing the difference between the one-point quadrature rule $f(x)$ and the integral If in the form of a series $h_{p-1}(x)$ and a remainder term $g_p(x) - If$. The Euler-Maclaurin summation formula may be derived from the Lanczos representation as follows: We denote the trapezoidal rule operator by

$$(3.1) \qquad R_x^{[m,\alpha]}(f(x)) = \frac{1}{m} \sum_{j=0}^{m-1} \bar{f}((j + t_\alpha)/m), \qquad t_\alpha = (1 + \alpha)/2.$$

This is abbreviated to $R^{[m,\alpha]}f$ in cases in which no confusion is likely to arise. It is trivial to verify that

$$(3.2) \qquad R_x^{[m,\alpha]}(e^{2\pi ikx}) = e^{2\pi i(k/m)t_\alpha}, \qquad k/m = \text{integer or zero},$$

$$= 0, \qquad\qquad\qquad \text{otherwise}.$$

The trapezoidal rule, applied to a Bernoulli function, generates another Bernoulli function. Applying (3.2) to (1.16) gives

$$(3.3) \qquad R_x^{[m,\alpha]}(\bar{B}_q(x - t)) = \bar{B}_q(t_\alpha - mt)/m^q.$$

Applying this operator to the Lanczos representation

$$(3.4) \qquad R^{[m,\alpha]}f = R^{[m,\alpha]}h_{p-1} + R_x^{[m,\alpha]}(g_p(x) - If) + R_x^{[m,\alpha]}(If),$$

we find using (3.3) that

$$(3.5) \qquad R^{[m,\alpha]}f = If + \sum_{q=1}^{p-1} (f^{(q-1)}(1) - f^{(q-1)}(0)) \frac{\bar{B}_q(t_\alpha)}{m^q q!} + E_p^{[m,\alpha]}f,$$

where the remainder term, denoted here by $E_p^{[m,\alpha]}f$, is

$$(3.6) \qquad E_p^{[m,\alpha]}f = R_x^{[m,\alpha]}(g_p(x) - If).$$

The standard integral representations for this remainder term may be obtained directly from the integral representations (1.27) for $g_p(x)$. Using (3.6) again, we find

$$E_p^{[m,\alpha]}f = \frac{1}{m^p} \int_0^1 f^{(p)}(t) \frac{\bar{B}_p(t_\alpha) - \bar{B}_p(t_\alpha - mt)}{p!} dt$$

(3.7)

$$= -\frac{1}{m^{p-1}} \int_0^1 f^{(p-1)}(t) \frac{\bar{B}_{p-1}(t_\alpha - mt)}{(p - 1)!} dt.$$

This particular derivation of this well-known result suggests that there should be a close connection between the convergence properties of the infinite series known as the Euler-Maclaurin series obtained by allowing p to become infinite in (3.5) and the infinite series discussed at length in Section 2. In fact, these series differ only in so far as there is an additional m^{-q} factor in the qth term and that x is replaced by t_α. Thus, Theorem 2.25 applies here, with these two alterations; i.e., t_α replaces x and the conditions $\tau > T$ and $\tau \leq T$ are replaced by $\tau > Tm$ and $\tau \leq Tm$, respectively.

THEOREM 3.8. *Under the conditions of Theorem 2.25 and following the notation of Theorem 2.25, the Euler-Maclaurin series*

$$\sum_{q=1}^{\infty} (f^{(q-1)}(1) - f^{(q-1)}(0))B_q(t_\alpha)/m^q q!$$

converges for $t_\alpha \in X$ *if* $F(x)$ *is an entire function of exponential type* $\tau < mT$ *and diverges for* $t_\alpha \in X$ *if* $F(x)$ *is not an entire function of exponential type* $\tau \leqq mT$. $X, F(x)$ *and* T *are given in Table 2.26.*

There is a direct analogue of Theorem 2.30, which may be proved in just the same way. This is

THEOREM 3.9. *If* $f(x)$ *is an entire function of exponential type* $\tau < 2\pi m$, *then the Euler–Maclaurin series converges to the expected result, that is*

$$R^{[m,\alpha]}f = If + \sum_{q=1}^{\infty} (f^{(q-1)}(1) - f^{(q-1)}(0))B_q(t_\alpha)/m^q q!.$$

This is only a sufficient condition. Theorems 3.8 and 3.9 are illustrated by the following example: Let

(3.10) $$f(x) = e^{2\pi\beta x} + \cos 2\pi nx, \qquad \beta = 3\tfrac{1}{2}, \quad n = 6.$$

Then $f(x)$ is an entire function of exponential type $\tau = \max(|n|, |\beta|) = 6$ while

(3.11) $$\varphi(x) = (e^{2\pi\beta} - 1)e^{2\pi\beta x}$$

is an entire function of exponential type $\tau = |\beta| = 3\tfrac{1}{2}$.

Thus, according to Theorem 3.8, the Euler–Maclaurin series diverges for $m = 1$, 2, 3 and converges for $m > |\beta| = 3\tfrac{1}{2}$. According to Theorem 3.9, the Euler-Maclaurin series converges to the proper result for $m > n = 6$. For the intervening values of m, i.e., $m = 4, 5, 6$, the series converges, but so far as Theorem 3.9 is concerned it may converge to a different result. Inspection of this particular function shows that in fact it converges correctly for $m = 4, 5$ but incorrectly for $m = 6$. This example establishes that the converse of Theorem 3.9 is not generally valid.

To complete the picture, we note a well-known result, namely that there are cases in which the Euler-Maclaurin expansion is semiconvergent. This happens for the standard cases of the trapezoidal endpoint rule ($\alpha = 1$, $t_\alpha = 1$) and the trapezoidal midpoint rule ($\alpha = 0$; $t_\alpha = \tfrac{1}{2}$) when the derivatives of $f(x)$ satisfy one of the following inequalities:

(3.12a) $$f^{(2p)}(x) \geqq 0 \quad \text{all } p \geqq p_0, \qquad 0 \leqq x \leqq 1,$$

(3.12b) $$f^{(2p)}(x) \leqq 0 \quad \text{all } p \geqq p_0, \qquad 0 \leqq x \leqq 1.$$

This follows quite simply using the inequalities

(3.13)
$$(-1)^p(B_{2p}(1) - \bar{B}_{2p}(1 - mt)) \geqq 0, \qquad 0 \leqq t \leqq 1,$$

$$(-1)^p(\bar{B}_{2p}(\tfrac{1}{2}) - \bar{B}_{2p}(\tfrac{1}{2} - mt)) \leqq 0, \qquad 0 \leqq t \leqq 1.$$

Thus when $\alpha = 1$ or 0, the remainder term $E_{2p}^{[m,\alpha]}f$ given by (3.7) alternates in sign with p when $p > p_0$ so long as the even order derivatives of $f(x)$ satisfy one of (3.12a) or (3.12b).

The corresponding result about the series $(If + h_p(x))$ is rather uninteresting. This is that, if one of conditions (3.12a) or (3.12b) is satisfied, the sequence is semiconvergent at the values $x = \tfrac{1}{2}$ and $x = 1$.

To sum up, the convergence properties of the Euler-Maclaurin series are most unsatisfactory. For the rather narrow class of entire functions of exponential type $\tau < 2\pi m$, the series is known to converge to the correct result. For functions whose derivatives satisfy relations (3.12), the standard series (with $\alpha = 1$ and $\alpha = 0$) are semiconvergent. In the absence of this sort of information about $f(x)$, the series may diverge or may converge and if it converges this may be to a correct or to an incorrect result.

4. The Euler-Maclaurin Quadrature Rule. The rather unsatisfactory convergence properties of the Euler-Maclaurin series are not well known. For example, in elementary numerical analysis textbooks, a quadrature method called the Euler-Maclaurin rule is occasionally presented in the form

$$\frac{1}{h} \int_{x_0}^{x_n} f(x) \, dx = \tfrac{1}{2}f(x_0) + f(x_1) + \cdots + f(x_{n-1}) + \tfrac{1}{2}f(x_n)$$

$$(4.1) \qquad\qquad - \frac{h}{12} (f'(x_n) - f'(x_0)) + \frac{h^3}{720} (f'''(x_n) - f'''(x_0))$$

$$+ \frac{h^5}{30240} (f^{(5)}(x_n) - f^{(5)}(0)) \cdots$$

where

$$(4.2) \qquad\qquad x_j = x_0 + jh$$

together with a statement that the remainder term is

$$(4.3) \qquad\qquad \frac{n B_{2m} h^{2m}}{(2m)!} f^{(2m)}(\xi), \qquad x_0 < \xi < x_n.$$

and possibly a statement about the semiconvergent nature of the expansion. While perfectly correct, such a presentation may be very misleading, as it appears to be an invitation to the user to take a fixed value of n and to carry on calculating the terms in the expansion until they appear to be negligible. If the function in question happened to be periodic with period $x_n - x_0$, each term in the expansion is zero. The user is left with the usually incorrect result

$$(4.4) \qquad\qquad \int_0^1 f_p(x) \, dx = R^{[n,1]}f_p.$$

In this case, the appearance of a number of zero terms might act as a warning. But for the function

$$(4.5) \qquad\qquad f(x) = f_p(x) + g(x)$$

in the case where the Euler-Maclaurin expansion converged for $g(x)$, he would find the expansion for $f(x)$ converged to

$$(4.6) \qquad\qquad R^{[n,1]}f_p + Ig,$$

leaving an error $If - R^{[n,1]}f_p$. While this error may be small, it may be very much larger than the user imagined. His only fault in such a case is that he did not check the magnitude of the remainder term. However, in elementary computational pro-

cedure this step is habitually omitted as reference to most elementary textbooks will confirm.

In cases where the derivatives at the endpoints of the integration interval are known, the Euler-Maclaurin rule is a powerful tool, but it should be used in a manner in which theoretical convergence to the result is assured. For example, one could take a fixed value of p and evaluate approximations

$$I_{m,2p} = R^{[m,1]}f - \frac{B_2}{m^2}(f'(1) - f'(0)) - \frac{B_4}{m^4}(f'''(1) - f'''(0))$$

(4.7)

$$- \cdots - \frac{B_{2p}}{m^{2p}}(f^{(2p-1)}(1) - f^{(2p-1)}(0))$$

for a sequence of values of m. In this case,

(4.8)
$$I_{m,2p} = If - E_{2p+2}^{[m,1]}f$$

and the discretisation error goes to zero, satisfying the order relation

(4.9)
$$E_{2p+2}^{[m,1]}f \simeq O(m^{-2p-2}) \quad \text{as } m \to \infty.$$

Even if the derivatives in (4.7) are incorrect, the discretisation error approaches zero but at a different rate. While the user can be misled into thinking that the sequence has converged to within his required accuracy, at least he is constructing a sequence which does converge to the correct result.

Another quadrature rule which is open to the same sort of misuse is the Gregory rule. This may be derived from the Euler-Maclaurin series by reexpressing the derivatives in terms of finite differences. This is often presented in the form

$$\int_{x_0}^{x_n} f(x)\,dx = h(\tfrac{1}{2}f(x_0) + f(x_1) + \cdots + f(x_{n-1}) + \tfrac{1}{2}f(x_n))$$

(4.10)
$$- \frac{1}{12}(\Delta y_n - \Delta y_0) - \frac{1}{24}(\nabla^2 y_n + \Delta^2 y_0)$$

$$- \frac{19}{720}(\nabla^3 y_n - \Delta^3 y_0) - \frac{3}{160}(\nabla^4 y_n + \Delta^4 y_0)\cdots.$$

Again, using a fixed value of n and proceeding to evaluate the finite differences may be quite unreliable. But the theory is rather complicated since the errors incurred by approximating the derivatives by finite differences sometimes tend to compensate for an error of the form $R^{[m,1]}f_p - If_p$. When precisely n terms are retained, the formula is identical with the Newton-Cotes $(n + 1)$-point formula of closed type. A satisfactory procedure for using Gregory's formula is to fix the number of terms to be retained and to form a sequence of approximations in a manner analogous to that described above for the Euler-Maclaurin rule.

5 . Other Two-Point Series. There are naturally many other methods of decomposing a function $f(x)$ into the sum of a polynomial $h_{p-1}(x)$ and a function $g_p(x)$ having specified properties. The Lanczos representation is simply one which has a useful application which is described in Part II. There are in particular two other representations which have been described in the literature (Jones and Hardy [3])

which are extremely close to the Lanczos representation. One of these is the Lidstone series. We describe these two briefly in this section.

The Lanczos representation is derived in Section 1 by substituting the Fourier coefficient asymptotic expansion (1.13) into the Fourier series (1.3). The two similar representations are derived in an analogous manner using the cosine and the sine half-range Fourier series respectively. We set

$$(5.1) \qquad f(x) = a_0 + 2 \sum_{r=1}^{\infty} a_r \cos \pi r x, \qquad 0 \leq x \leq 1,$$

where

$$(5.2) \qquad a_r = \int_0^1 f(x) \cos \pi r x \, dx.$$

Using integration by parts, we find

$$(5.3) \qquad a_r = \sum_{j=1}^{p-1} \frac{(-1)^r f^{(2j-1)}(1) - f^{(2j-1)}(0)}{(\pi r)^{2j}} (-1)^{i+1} + a_{r,2p},$$

where

$$(5.4) \qquad a_{r,2p} = \frac{(-1)^p}{(\pi r)^{2p}} \left[-f^{(2p-1)}(1)(-1)^r + f^{(2p-1)}(0) + \int_0^1 f^{(2p)}(x) \cos \pi r x \, dx \right]$$

$$\simeq O(r^{-2p}) \quad \text{as } r \to \infty.$$

Substituting this into (5.1) gives

$$(5.5) \qquad f(x) = If + h_{2p-2}(x) + g_{2p}(x)$$

where

$$(5.6) \qquad h_{2p-2}(x) = \sum_{k=1}^{p-1} 2^{2k} \frac{B_{2k}((1+x)/2)}{(2k)!} f^{(2k-1)}(1) - 2^{2k} \frac{B_{2k}(x/2)}{(2k)!} f^{(2k-1)}(0)$$

and

$$(5.7) \qquad g_{2p}(x) = 2 \sum_{r=1}^{\infty} a_{r,2p} \cos \pi r x.$$

In view of the order relation in (5.4), the half range cosine Fourier coefficients of $g_{2p}(x)$ are of order (r^{-2p}) and, clearly, $h_{2p-2}(x)$ is a polynomial of degree $2p - 2$. The odd order derivatives of $h_{2p-2}(x)$ at the endpoints satisfy the relations

$$(5.8) \qquad h_{2p-2}^{(2j-1)}(1) = f_{2p-2}^{(2j-1)}(1), \qquad h_{2p-2}^{(2j-1)}(0) = f_{2p-2}^{(2j-1)}(0), \qquad j = 1, 2, \cdots, p - 1.$$

In the special case in which $f(x)$ is symmetric about $x = \frac{1}{2}$, the cosine half-range series (5.1) coincides with the full range Fourier series (1.3). In this case, the functions $h_{2p-2}(x)$ given by (5.6) coincide with those given by (1.21), both having the form

$$(5.9) \qquad h_{2p-1}(x) = h_{2p-2}(x) = 2 \sum_{k=1}^{p-1} f^{(2k-1)}(1) B_{2k}(x)/2k!.$$

Unless $f(x)$ is symmetric, the two functions $h_{2p-2}(x)$ given by (5.6) and by (1.21) are of different structure from one another.

The Lidstone series can be derived from

(5.10) $$f(x) = 2 \sum_{r=1}^{\infty} b_r \sin \pi r x, \qquad 0 < x < 1,$$

where

(5.11) $$b_r = \int_0^1 f(x) \sin \pi r x \, dx.$$

Using integration by parts,

(5.12) $$b_r = \sum_{j=0}^{p-1} \frac{(-1)^r f^{(2j)}(1) - f^{(2j)}(0)}{(\pi r)^{2j+1}} (-1)^{j+1} + b_{r,2p+1}$$

with

(5.13) $$b_{r,2p+1} = \frac{(-1)^p}{(\pi r)^{2p+1}} \left[-f^{(2p)}(1)(-1)^r + f^{(2p)}(0) + \int_0^1 f^{(2p+1)}(x) \cos \pi r x \, dx \right]$$

$$\simeq O(r^{-(2p+1)}) \quad \text{as } r \to \infty.$$

Substituting this into (5.10) gives

(5.14) $$f(x) = h_{2p-1}(x) + g_{2p+1}(x)$$

with

(5.15) $$h_{2p-1}(x) = \sum_{j=0}^{p-1} 2^{2j+1} \frac{B_{2j+1}((x+1)/2)}{(2j+1)!} f^{(2j)}(1) - \frac{2^{2j+1} B_{2j+1}(x/2)}{(2j+1)!} f^{(2j)}(0)$$

and

(5.16) $$g_{2p+1}(x) = 2 \sum_{r=1}^{\infty} b_{r,2p+1} \sin \pi r x.$$

In view of the order relation in (5.13), the half range sine Fourier coefficients of $g_{2p+1}(x)$ are of order $(r^{-(2p+1)})$. Clearly, $h_{2p-1}(x)$ is a polynomial of degree $2p - 1$. It is conventionally expressed in the form

(5.17) $$h_{2p-1}(x) = \sum_{j=0}^{p-1} \Lambda_j(x) f^{(2j)}(1) + \Lambda_j(1-x) f^{(2j)}(0)$$

where $\Lambda_j(x)$ is the Lidstone polynomial of degree $2j + 1$ defined by

(5.18) $$\Lambda_j(x) = 2^{2j+1} \frac{B_{2j+1}((x+1)/2)}{(2j+1)!}.$$

The even order derivatives of $h_{2p-1}(x)$ satisfy the relations

(5.19) $$h_{2p-1}^{(2j)}(1) = f_{2p-1}^{(2j)}(1), \qquad h_{2p-1}^{(2j)}(0) = f_{2p-1}^{(2j)}(0), \qquad j = 0, 1, \cdots, p - 1.$$

If $f(x)$ is antisymmetric about $x = \frac{1}{2}$, then the function $h_{2p-1}(x)$ given by (5.17) coincides with the function $h_{2p-1}(x)$ given by (1.21), both having the form

(5.20) $$h_{2p-1}(x) = h_{2p}(x) = 2 \sum_{j=0}^{p-1} f^{(2j)}(1) B_{2j+1}(x)/(2j+1)!.$$

Otherwise these series are different from one another. Yet another series of the same

general nature is the series in Euler polynomials

$$h_p(x) = \sum_{j=0}^{p} \tfrac{1}{2}(f^{(j)}(1) + f^{(j)}(0))E_j(x)/j!$$

with the related property

$$h_p^{(j)}(1) + h_p^{(j)}(0) = f^{(j)}(1) + f^{(j)}(0), \qquad j = 0, 1, \cdots, p.$$

These series differ fundamentally from each other and differ fundamentally from the two-point truncated Taylor expansion which interpolates $f(x)$ and its derivatives at $x = 0$ and $x = 1$.

There is a wide literature about the Lidstone series. See for example Lidstone [5], Widder [6]. However, only in the case in which $f(x)$ happens to be antisymmetric are these results relevant to the Lanczos representation.

PART II. AN APPROXIMATE REPRESENTATION FOR $f(x)$

1. **An Approximate Representation** $F(x)$ **for** $f(x)$. In Part I, the Lanczos representation of a function $f(x)$ was introduced and some of its theoretical properties were described. Briefly, when $f(x) \in AR[0, 1]$, this has the form

$$(1.1) \qquad f(x) = h_{p-1}(x) + g_p(x).$$

Here $h_{p-1}(x)$ is a polynomial of degree $p - 1$ given by

$$(1.2) \qquad h_{p-1}(x) = \sum_{q=1}^{p-1} (f^{(q-1)}(1) - f^{(q-1)}(0))B_q(x)/q!,$$

while $g_p(x)$ has Fourier coefficients which satisfy the asymptotic order relations

$$C^{(r)}g_p \equiv \int_0^1 g_p(x) \cos 2\pi r x \, dx \simeq O(r^{-p}) \quad \text{as } r \to \infty,$$

$$(1.3)$$

$$S^{(r)}g_p \equiv \int_0^1 g_p(x) \sin 2\pi r x \, dx \simeq O(r^{-p}) \quad \text{as } r \to \infty.$$

The work presented here is based on the results given in Sections 1, 2 and 3 of Part I. Since the Fourier coefficients of $g_p(x)$ diminish in magnitude rapidly, we consider a numerical representation $F(x)$ for $f(x)$, based on the exact expansion

$$f(x) = \sum_{q=1}^{p-1} (f^{(q-1)}(1) - f^{(q-1)}(0))B_q(x)/q! + Ig_p$$

$$(1.4)$$

$$+ 2\sum_{r=1}^{\infty} C^{(r)}g_p \cos 2\pi r x + 2\sum_{r=1}^{\infty} S^{(r)}g_p \sin 2\pi r x.$$

Expressions for the Bernoulli polynomials $B_q(x)$ are readily available (Abramowitz and Stegun [1, p. 803]). Thus, given a value of p we require numerical values of

$$(1.5) \qquad \varphi^{(q-1)}(0), \qquad q = 1, 2, \cdots, p - 1,$$

where

$$(1.6) \qquad \varphi(x) = f(x + 1) - f(x)$$

and numerical values of

$$(1.7) \qquad\qquad Ig_p;\ C^{(r)}g_p,\ S^{(r)}g_p, \qquad r = 1, 2, \cdots, s,$$

where

$$(1.8) \qquad\qquad g_p(x) = f(x) - h_{p-1}(x),$$

the value of s depending on the numerical accuracy required. The advantage of such a representation over the Fourier series of $f(x)$ is that, because of order relations (1.3), the value of s is likely to be reasonably small, and so fewer Fourier coefficients may be needed to attain a given accuracy. The disadvantage is that the derivatives (1.5) are required.

Before proceeding, we discard the possibility that, by choosing p sufficiently large, we may attain sufficient accuracy in the approximation without having to evaluate any of the Fourier coefficients of $g_p(x)$ other than Ig_p. This would be the case if

$$(1.9) \qquad\qquad \lim_{p \to \infty}(If + h_p(x)) = f(x).$$

However, we have shown in Section 2 of Part I, that while (1.9) may be established for functions $f(x)$ which are entire functions of exponential type $\tau < 2\pi$, in general, the limit in the left-hand side of (1.9) does not exist and

$$(1.10) \qquad\qquad \lim_{p \to \infty} |h_p(x)| = \infty.$$

Furthermore, even if the limit exists, it may not be equal to $f(x)$. Consequently, a general method based on (1.4) has to employ a moderate value of p and has to include provision for calculating at least some of the Fourier coefficients of $g_p(x)$. One effect of using a larger value of p is to increase the magnitude of the early Fourier coefficients of $g_p(x)$. It is only useful to do this if enough of the later Fourier coefficients are to be calculated that advantage can be taken of the more rapid ultimate convergence rate.

Since the Fourier coefficients of $g_p(x)$ converge reasonably rapidly, the use of the trapezoidal rule for their evaluation is reasonably efficient. In principle, we might employ any offset trapezoidal rule of the form

$$(1.11) \qquad R^{[m,\alpha]}f = R_v^{[m,\alpha]}(f(y)) = \frac{1}{m}\sum_{j=0}^{m-1} \bar{f}((j + t_\alpha)/m), \qquad t_\alpha = (1 + \alpha)/2,$$

and make use of trapezoidal rule sum approximations to the Fourier coefficients (1.3), denoted by

$$(1.12) \qquad\begin{aligned} a_{r,p}^{[m,\alpha]} &= R_v^{[m,\alpha]}(g_p(y)\cos 2\pi ry), \qquad r = 0, 1, \cdots, m-1, \\ b_{r,p}^{[m,\alpha]} &= R_v^{[m,\alpha]}(g_p(y)\sin 2\pi ry), \qquad r = 1, 2, \cdots, m-1. \end{aligned}$$

It turns out that, from a practical standpoint, the use of the endpoint trapezoidal rule

$$(1.13) \qquad R^{[m,1]}f = \frac{1}{m}\left\{\frac{1}{2}f(0) + \sum_{j=1}^{m-1} f(j/m) + \frac{1}{2}f(1)\right\}$$

is likely to be most convenient. This coincides with (1.11) with $\alpha = 1$ or $\alpha = -1$ since $\bar{f}(x) = f(x)$ when $0 < x < 1$ and

(1.14) $$\bar{f}(1) = \bar{f}(0) = \tfrac{1}{2}\{f(1) + f(0)\}.$$

The general theory with other choices of α is marginally more complicated. In Sections 2–4, we treat the following approximation to $f(x)$.

Approximation 1.15.

(1.15)
$$F(x) = \sum_{q=1}^{p-1} \varphi^{(q-1)}(0) \frac{B_q(x)}{q!} + 2 \sum_{r=0}^{m/2}{}'' \ a_{r,p}^{[m,1]} \cos 2\pi r x$$

$$+ 2 \sum_{r=0}^{m/2}{}'' \ b_{r,p}^{[m,1]} \sin 2\pi r x.$$

The double prime on the summation symbol has the meaning that the ultimate terms (as stated), i.e., those with $r = 0$ and $r = m/2$, are assigned a weight factor of $\tfrac{1}{2}$. When m is odd, the summation concludes with the term $r = (m - 1)/2$ with the normal weight factor of unity. Incidentally, it follows from the definition of $b_r^{[m,1]}$ that

(1.16) $$b_{0,p}^{[m,1]} \equiv b_{m/2,p}^{[m,1]} = 0$$

so the ultimate terms in the second sum over r do not in fact occur. The reason for truncating this series at $r = m/2$, and the reason for the ultimate weight factor $\tfrac{1}{2}$ is given in Section 2.

The approximate representation (1.15) requires for its construction the parameters p and m, the derivatives $\varphi^{(q-1)}(0)$, $q = 1, 2, \cdots, p - 1$, and the function values $f(j/m)$, $j = 0, 1, \cdots, m$. In Section 2, we deal with the discretisation error $F(x) - f(x)$. In Section 3, we consider the effect of using approximate derivatives in place of $\varphi^{(q-1)}(0)$ in (1.15). In Section 4, we describe possible implementations. One of these is designed to determine appropriate parameters p and m to provide an overall discretisation error less than a given tolerance ϵ_{req}.

2. The Discretisation Error $F(x) - f(x)$. In this section, we discuss the discretisation error $F(x) - f(x)$ of the approximation $F(x)$ defined by (1.15) to $f(x)$. We note that, since the function $h_{p-1}(x)$ is common to both $f(x)$ and $F(x)$, it follows that

(2.1) $$F(x) - f(x) = G_p(x) - g_p(x)$$

where $G_p(x)$ is a well-known approximation (Gentleman and Sande [8]) to $g_p(x)$, namely

(2.2) $$G_p(x) = 2 \sum_{r=0}^{m/2}{}'' \ (a_{r,p}^{[m,1]} \cos 2\pi r x + b_{r,p}^{[m,1]} \sin 2\pi r x).$$

In this section, we shall drop the subscript p when no confusion is likely to arise. Since

(2.3)
$$g(x) = Ig + 2 \sum_{r=1}^{\infty} C^{(r)}g \cos 2\pi r x$$

$$+ 2 \sum_{r=1}^{\infty} S^{(r)}g \sin 2\pi r x,$$

it follows that

$$G_p(x) - g(x) = 2 \sum_{r=0}^{m/2}{}'' (a_r^{[m,1]} - C^{(r)}g) \cos 2\pi rx$$

$$(2.4) \qquad + 2 \sum_{r=1}^{m/2-1} (b_r^{[m,1]} - S^{(r)}g) \sin 2\pi rx - 2 \sum_{r=m/2}^{\infty}{}'' C^{(r)}g \cos 2\pi rx$$

$$- 2 \sum_{r=m/2}^{\infty} S^{(r)}g \sin 2\pi rx.$$

The purpose of this section is to express this discretisation error in terms of the Fourier coefficients of the function $g_p(x)$. This may be done most conveniently by means of the Poisson summation formula (2.5) below. We verify first that

$$R_x^{[m,\alpha]}(e^{2\pi irz}) = e^{2\pi i(r/m)t\alpha}, \qquad r/m = \text{integer},$$

$$= 0, \qquad\qquad \text{otherwise}.$$

Then, applying the trapezoidal rule operator to the Fourier series of $f(x)$ gives

$$(2.5) \qquad R^{[m,1]}f - If = 2 \sum_{s=1}^{\infty} C^{(ms)}f.$$

We replace $f(x)$ here by $g(x) \cos 2\pi rx$. Since

$$(2.6) \qquad 2C_x^{(ms)}(g(x) \cos 2\pi rx) = C^{(ms+r)}g + C^{(ms-r)}g,$$

we find

$$(2.7) \qquad a_r^{[m,1]} - C^{(r)}g = \sum_{s=1}^{\infty} C^{(ms+r)}g + C^{(ms-r)}g.$$

Similarly,

$$(2.8) \qquad b_r^{[m,1]} - S^{(r)}g = \sum_{s=1}^{\infty} S^{(ms+r)}g - S^{(ms-r)}g.$$

The case with $r = m/2$ is somewhat exceptional. These summations then reduce to

$$(2.9) \qquad a_{m/2}^{[m,1]} - 2C^{(m/2)}g = 2 \sum_{s=1}^{\infty} C^{(ms+m/2)}g,$$

$$(2.10) \qquad b_{m/2}^{[m,1]} = 0.$$

The reason for cutting off the series (1.15) at the term $r = m/2$ becomes apparent on inspection of (2.7) or (2.8). With $r > m/2$, a term $C^{(t)}g$ occurs on the right of (2.7) with $t < r$. Since the Fourier coefficients generally decrease in magnitude, this implies that the approximation error is likely to be greater than the quantity being approximated when $r > m/2$. Substituting (2.7) and (2.8) into (2.4) gives the principal result of this section, namely

$$(2.11) \qquad G(x) - g(x) = \sum_{t=0}^{\infty} C^{(t)}g\alpha_t(x) + \sum_{t=0}^{\infty} S^{(t)}g\beta_t(x),$$

with the values of $\alpha_t(x)$ being given by

$$(2.12) \qquad \alpha_t(x) = 0, \qquad\qquad\qquad\qquad\qquad t \leqq m/2,$$

$$\alpha_{ms+r}(x) = 2(\cos 2\pi r x - \cos 2\pi(ms + r)x), \qquad s \geqq 1, |r| \leqq m/2,$$

$$\beta_t(x) = 0, \qquad\qquad\qquad\qquad\qquad t < m/2,$$

$$(2.13) \qquad \beta_{ms+r}(x) = 2(\sin 2\pi r x - \sin 2\pi(ms + r)x), \qquad s \geqq 1, |r| < m/2,$$

$$\beta_{ms+r}(x) = -2 \sin 2\pi(ms + r)x, \qquad\qquad s \geqq 1, |r| = m/2.$$

The main result of this section follows from the circumstance that by inspection

$$(2.14) \qquad\qquad\qquad \alpha_t(x) = \beta_t(x) = 0, \qquad t < m/2,$$

$$(2.15) \qquad\qquad\qquad |\alpha_t(x)| \leqq 4, \qquad |\beta_t(x)| \leqq 4 \quad \text{all } t.$$

Thus

$$(2.16) \qquad\qquad |G_p(x) - g_p(x)| \leqq 4 \sum_{t=m/2}^{\infty} |C^{(t)} g_p| + |S^{(t)} g_p|.$$

In view of order relations (1.3) and identity (2.1), we find

THEOREM 2.17. *For a fixed value of p, the approximation* 1.15 *has a discretisation error which satisfies the order relation*

$$|F(x) - f(x)| \simeq O(m^{-p+1}) \quad \text{as } m \to \infty.$$

Certain other results of a well-known character follow from (2.11).

THEOREM 2.18. *If* $f(x)$ *is a trigonometrical polynomial of degree* $m/2 - 1$, *then*

$$F(x) \equiv \bar{f}(x).$$

Proof. In this case, $h_{p-1}(x) = 0$ and $f(x) = g_p(x)$. Thus,

$$C^{(t)} f = C^{(t)} g = S^{(t)} f = S^{(t)} g = 0, \qquad t \geqq m/2,$$

and the result follows directly from (2.16) and (2.1).

THEOREM 2.19. *If* $f(x)$ *is an algebraic polynomial of degree* $p - 1$, *then*

$$F(x) \equiv f(x).$$

Proof. In this case, $g_p(x) = If$, a constant. Thus, all the Fourier coefficients occurring in (2.16) are zero.

THEOREM 2.20. *The discretisation error is zero at the abscissas* $x_j = j/m$,

$$F(j/m) = \bar{f}(j/m), \qquad j = 0, 1, \cdots, m.$$

This is an algebraic result, independent of any properties of $f(x)$. Setting $x = j/m$ in (1.15) and simplifying the resulting double summation leads to this result. However, it follows quite simply from (2.11) because $rx_j - (ms + r)x_j = -sj$ is an integer and so, from (2.12) and (2.13),

$$\alpha_t(x_j) = \beta_t(x_j) = 0.$$

3. The Effect of Inaccurate Derivatives $\varphi^{(q-1)}(0)$. The approximation described in the previous sections is a powerful one by commonly accepted standards in computational theory. It has one significant drawback. This is that the values of the derivatives of $\varphi(x)$ given by

$$(3.1) \qquad \varphi(x) = f(x + 1) - f(x)$$

at $x = 0$ are required. In the case that $f(x)$ is an analytic function, and function evaluations in the complex plane are allowed, then this does not present a significant difficulty (see Lyness and Sande [9]). However, it can happen that methods for evaluating $f(x)$ when x is complex are not readily available. In this case, one may have to resort to finite-difference approximations, which are notoriously unreliable.

Previously, the prevailing view has been that, as a practical tool, this approximation has little value simply because these derivatives are required. In fact, some ingenuity has been applied to varying the approximation in such a way that a result of a particular order may be obtained requiring one less derivative evaluation. (Lanczos [4], Jones and Hardy [3]). Certainly, if an equally convenient and accurate method were to be constructed which did not require these derivatives, that method would be preferable. In the present absence of such a method, it seems worthwhile to look more closely into the effect of using approximate derivatives, which we denote by $\tilde{\varphi}^{(q-1)}(0)$ in place of exact derivatives $\varphi^{(q-1)}(0)$. We denote by Δ_{q-1} the error in the $(q - 1)$th derivative. Thus

$$(3.2) \qquad \Delta_{q-1} = \tilde{\varphi}^{(q-1)}(0) - \varphi^{(q-1)}(0).$$

Clearly, the procedure by which $F(x)$ is constructed may be applied using these inaccurate derivatives. One constructs in place of $h_{p-1}(x)$ given by (1.2) a different function $\tilde{h}_{p-1}(x)$ given by

$$(3.3) \qquad \tilde{h}_{p-1}(x) = \sum_{q=1}^{p-1} \tilde{\varphi}^{(q-1)}(0) B_q(x)/q!.$$

In place of $g_p(x)$, the function

$$(3.4) \qquad \tilde{g}_p(x) = f(x) - \tilde{h}_{p-1}(x)$$

is used. In place of $a_{r,p}^{[m,1]}$ and $b_{r,p}^{[m,1]}$, the trapezoidal rule sums

$$(3.5) \qquad \tilde{a}_{r,p}^{[m,1]} = R_y^{[m,1]}(\tilde{g}_p(y) \cos 2\pi r y),$$

$$\tilde{b}_{r,p}^{[m,1]} = R_y^{[m,1]}(\tilde{g}_p(y) \sin 2\pi r y)$$

are calculated, and these are inserted into (1.15) to give, in place of $F(x)$, the function

$$(3.6) \qquad \begin{aligned} \tilde{F}(x) = & \sum_{q=1}^{p-1} \tilde{\varphi}^{(q-1)}(0) \frac{B_q(x)}{q!} + 2 \sum_{r=0}^{m/2}{}'' \tilde{a}_{r,p}^{[m,1]} \cos 2\pi r x \\ & + 2 \sum_{r=0}^{m/2}{}'' \tilde{b}_{r,p}^{[m,1]} \sin 2\pi r x. \end{aligned}$$

Many of the results obtained in Sections 1 and 2 apply equally with tildes attached to each quantity. In fact, all results based only on algebraic manipulation are valid. The results which do not apply are those involving order relations. Thus we find, in place of (2.16), that

$$|\tilde{G}_p(x) - \tilde{g}_p(x)| \leqq 4 \sum_{t=m/2}^{\infty} |C^{(t)} \tilde{g}_p| + |S^{(t)} \tilde{g}_p|$$

but, since the Fourier coefficients $S^{(t)} \tilde{g}_p$ are not of order $O(t^{-p})$ but are of order $O(t^{-1})$, the result corresponding to Theorem 2.17 is simply

$$|\bar{F}(x) - f(x)| \sim O(m^0) \quad \text{as } m \to \infty$$

and not $O(m^{-p+1})$. Consequently, once inaccurate derivatives are used, the main feature of the procedure, its rapid ultimate convergence rate, disappears.

On the other hand, one would expect that if the errors in the derivatives were very small, the difference $\bar{F}(x) - F(x)$ would also be small and that the method would be almost as powerful as if the derivatives were exact. In other words, while the order reduces from $O(m^{-p+1})$ to $O(m^0)$, the numerical value of the difference might be small.

Consequently, we investigate the effect of inaccurate derivatives in this section. First, we find an expression for $\bar{F}(x) - F(x)$ (Theorem 3.16 below). Then we determine a bound for $|\bar{F}(x) - F(x)|$ (Theorem 3.21 below). These involve the quantities $\Delta_{q-1}(x)$. Finally, we make some realistic assumptions about the magnitude of Δ_{q-1} and, on the basis of these, discuss the overall effect on the calculation. It turns out that this is not catastrophic. In some cases, there is no noticeable effect at all. In others, there is an increase in the required value of m, but in a calculational context, it is effectively the same as replacing $O(m^{-p+1})$ by $O(m^{-p+2})$.

We now proceed to obtain expressions for $\bar{F}(x) - F(x)$ in terms of Δ_{q-1}. Taking the difference between (1.2) and (3.3), we find

$$(3.7) \qquad \tilde{g}_p(x) - g_p(x) = -\tilde{h}_{p-1}(x) + h_{p-1}(x) = \sum_{q=1}^{p-1} - \Delta_{q-1} B_q(x)/q!.$$

Taking the difference between (1.12) and (3.5) and using (3.7),

$$(3.8) \qquad \begin{aligned} \tilde{a}_{r,p}^{[m,1]} - a_{r,p}^{[m,1]} &= R_y^{[m,1]}((\tilde{g}_p(y) - g_p(y)) \cos 2\pi ry) \\ &= -\sum_{q=1}^{p-1} \Delta_{q-1} R_y^{[m,1]}((B_q(y)/q!) \cos 2\pi ry). \end{aligned}$$

It is convenient to denote trapezoidal rule approximations to the Fourier coefficients of the Bernoulli polynomials by $X_{r,q}^{[m,\alpha]}$ and $Y_{r,q}^{[m,\alpha]}$. These are defined as the real and imaginary parts of

$$(3.9) \qquad Z_{r,q}^{[m,\alpha]} = X_{r,q}^{[m,\alpha]} + i Y_{r,q}^{[m,\alpha]} = R_y^{[m,\alpha]}((B_q(y)/q!) e^{2\pi i ry}).$$

Using (1.15), (3.6), (3.7), (3.8) and the corresponding relation to (3.8) for $\tilde{b}_{r,p}^{[m,1]}$, straightforward algebraic manipulation yields

$$(3.10) \qquad \bar{F}(x) - F(x) = \sum_{q=1}^{p-1} \Delta_{q-1} E_q^{[m,1]}(x)$$

where

$$(3.11) \qquad E_q^{[m,1]}(x) = \frac{B_q(x)}{q!} - 2 \sum_{r=0}^{m/2}{''} X_{r,q}^{[m,1]} \cos 2\pi rx - 2 \sum_{r=1}^{m/2-1} Y_{r,q}^{[m,1]} \sin 2\pi rx.$$

An alternate form for $E_q^{[m,1]}(x)$ may be obtained by rederiving (3.10) directly from (2.11). In view of (2.1), we have

$$(3.12) \qquad F(x) - f(x) = \sum_{t=0}^{\infty} (C^{(t)} g_p \alpha_t(x) + S^{(t)} g_p \beta_t(x))$$

where the functions $\alpha_t(x)$ and $\beta_t(x)$ are given by (2.12) and (2.13). An equation similar to (3.12) is valid with $F(x)$ replaced by $\bar{F}(x)$ and g_p replaced by $\tilde{g}_p$. Taking the difference

between this equation and (3.12) and using (3.6) again, we recover (3.10) above, with $E_q^{[m,1]}(x)$ expressed in a different form, namely

$$(3.13) \qquad E_q^{[m,1]}(x) = -\sum_{t=0}^{\infty} \left(\alpha_t(x) C_\nu^{(t)}\left(\frac{B_q(y)}{q!}\right) + \beta_t(x) S_\nu^{(t)}\left(\frac{B_q(y)}{q!}\right) \right).$$

The Fourier coefficients of the Bernoulli polynomials are very straightforward. Thus, with $q > 0$,

$$(3.14)$$
$$C_\nu^{(t)}\left(\frac{B_q(y)}{q!}\right) = \frac{(-1)^{q/2+1}}{(2\pi t)^q}, \qquad S_\nu^{(t)}\left(\frac{B_q(y)}{q!}\right) = 0, \qquad q \quad \text{even},$$

$$S_\nu^{(t)}\left(\frac{B_q(y)}{q!}\right) = \frac{(-1)^{q/2+1/2}}{(2\pi t)^q}, \qquad C_\nu^{(t)}\left(\frac{B_q(y)}{q!}\right) = 0, \qquad q \quad \text{odd}.$$

Since $\alpha_t(x) = 0$ when $t \le m/2$, we find

$$(3.15)$$
$$E_q^{[m,1]}(x) = \sum_{t=m/2+1}^{\infty} (-1)^{q/2} \alpha_t(x)/(2\pi t)^q, \qquad q \quad \text{even},$$

$$E_q^{[m,1]}(x) = \sum_{t=m/2}^{\infty} (-1)^{q/2-1/2} \beta_t(x)/(2\pi t)^q, \qquad q \quad \text{odd}.$$

The results obtained so far in this section may be summarised as follows:

THEOREM 3.16. *The approximation $\tilde{F}(x)$ obtained using derivatives $\tilde{\varphi}^{(q-1)}(0) = \varphi^{(q-1)}(0) + \Delta_{q-1}$ differs from $F(x)$ given by (1.15) by an amount*

$$\tilde{F}(x) - F(x) = \sum_{q=1}^{p-1} \Delta_{q-1} E_q^{[m,1]}(x)$$

where the coefficients are given explicitly by (3.11) *or* (3.15).

We now turn to deriving a bound on the magnitude of the difference $\tilde{F}(x) - F(x)$. To this end, we require bounds on the magnitude of $E_q^{[m,1]}(x)$. Since $|\alpha_t(x)| \le 4$ and $|\beta_t(x)| < 4$, we find a simple bound in the cases in which $q \ge 2$ in terms of the generalised zeta function

$$(3.17) \qquad \zeta(q, a) = \sum_{r=0}^{\infty} 1/(a + r)^q, \qquad q > 1, a > 0,$$

and its well-known bound

$$(3.18) \qquad \zeta(q, a) < 1/(q - 1)(a - \tfrac{1}{2})^{q-1}, \qquad q > 1, a > \tfrac{1}{2}.$$

Using this in (3.15) yields

$$(3.19)$$
$$|E_q^{[m,1]}(x)| < \frac{4}{(2\pi)^q} \sum_{t=m/2}^{\infty} \frac{1}{t^q} = \frac{4}{(2\pi)^q} \zeta(q, m/2)$$

$$< \frac{2/\pi}{(\pi(m-1))^{q-1}(q-1)}, \qquad q > 1.$$

The case with $q = 1$ has to be treated separately. In Appendix 2, it is shown that

$$(3.20) \qquad\qquad |E_1^{[m,1]}(x)| < \tfrac{2}{3}.$$

THEOREM 3.21. *The quantity expressed in Theorem 3.16 is bounded by*

$$(3.21) \qquad |\tilde{F}(x) - F(x)| < \tfrac{2}{3}|\Delta_0| + \frac{2}{\pi} \sum_{q=2}^{p-1} \frac{|\Delta_{q-1}|}{(\pi(m-1))^{q-1}(q-1)}.$$

A marginally worse bound, which is convenient for the subsequent discussion, is

$$(3.22) \qquad |\tilde{F}(x) - F(x)| < \tfrac{2}{3} \sum_{q=1}^{p-1} |\Delta_{q-1}|/(\pi(m-1))^{q-1}.$$

In the rest of this section, we shall consider in some detail the behaviour of various terms as m becomes large. In this context, we shall assume that the value of p is fixed.

The analysis carried out above is valid for arbitrary values of Δ_{q-1}. Taking the limit in (3.21), we see

$$(3.23) \qquad \lim_{m \to \infty} |\tilde{F}(x) - F(x)| \leq \tfrac{2}{3}\Delta_0.$$

Since

$$(3.24) \qquad |F(x) - f(x)| \simeq O(m^{-p+1}) \quad \text{as} \quad m \to \infty$$

it follows that

$$(3.25) \qquad \lim_{m \to \infty} |\tilde{F}(x) - f(x)| \leq \tfrac{2}{3}\Delta_0.$$

We recall that, while we have referred to the quantities Δ_{q-1} as errors in derivatives, the quantity

$$(3.26) \qquad \Delta_0 = \bar{\varphi}(0) - \varphi(0)$$

is in fact an error in function values. Thus, in a situation in which all function values are exactly calculated and the whole calculation is carried out exactly (apart from the use of approximate derivatives $\bar{\varphi}^{(q-1)}(0)$, $q = 2, 3, \cdots, p-1$), it follows that $\Delta_0 = 0$ and (3.25) indicates that

$$(3.27) \qquad \lim_{m \to \infty} \tilde{F}(x) = f(x).$$

Of course, if approximate derivatives $\bar{\varphi}^{(q-1)}$, $q = 2, 3, \cdots, p-1$, were assigned quite arbitrarily, the convergence might be very slow indeed, the dominant term in the error $\tilde{F}(x) - f(x)$ being $\Delta_1 E_2^{[m,1]}(x) \simeq O(m^{-1})$ as $m \to \infty$.

In order to proceed, we have to make some hypotheses about the magnitude of the quantities Δ_q. In fact these will be that there exists numbers ϵ_f and k such that Δ_q and the required accuracy ϵ_{req} satisfy (3.34) below. We now justify this hypothesis in a computational context.

The results $F(x)$ and $\tilde{F}(x)$ depend on the function evaluations $f(j/m)$. However carefully the calculation is carried out, the accuracy of the final result is limited by the accuracy of the initial function values used. Let us suppose that these function values have an 'overall' absolute accuracy ϵ_f. This may be defined as follows. If $\tilde{f}_i$ is the approximation to $f(j/m)$ used in the calculation, then

$$(3.28) \qquad \epsilon_f = \max_{0 \leq i \leq m} |\tilde{f}_i - f(j/m)|.$$

If now a reasonable attempt to calculate $\varphi(0)$ were made, the error in $\varphi(0)$ is given by

$$(3.29) \qquad \Delta_0 = \bar{\varphi}(0) - \varphi(0) = \tilde{f}_m - f(1) - (\tilde{f}_0 - f(0))$$

and so

(3.30)
$$|\Delta_0| \leqq 2\epsilon_f.$$

However they are calculated, the error in the higher order derivatives is likely to be larger. If these are calculated using finite-difference approximations, the error rises rapidly with the order of the derivative involved. It is convenient to assign a geometric bound and assert

(3.31)
$$|\Delta_{q-1}| \leqq 2\epsilon_f k^{q-1}, \qquad q = 1, 2, \cdots, p - 1.$$

Since only a finite number of bounds are involved, and all quantities are finite, a number k (which depends on p) exists, defined by

(3.32)
$$k = \max_{1 \leqq q \leqq p-1} \left| \frac{\tilde{\varphi}^{(q-1)}(0) - \varphi^{(q-1)}(0)}{2\epsilon_f} \right|^{1/(q-1)}.$$

The value of k is not usually known. If one assumes that one loses five binary digits of precision per differentiation, then $k = 32$. When finite differences are used, the rate of increase of Δ_{q-1} depends on the nature of the function and the actual magnitude of the derivatives. Thus k may be greater than or less than 32 and depends on the value of p.

The existence of a limit ϵ_f on the accuracy of the function values has another consequence. If one envisages a calculation involving a moderate value of p, say $p = 10$, then, in view of the various additions and subtractions involved in the calculation, one would be very rash to ask for an overall accuracy less than $2p\epsilon_f$. In fact, one would be much more cautious than this. Thus it is consistent to suppose that the required accuracy satisfies

(3.33)
$$\epsilon_{req} \geqq 2p\epsilon_f.$$

The discussion above provides a justification for the use of the following constraints on Δ_q and on ϵ_{req}. These are:

There exist numbers ϵ_f and k having the property

(3.34)
$$|\Delta_{q-1}| \leqq 2\epsilon_f k^{q-1}, \qquad q = 1, 2, \cdots, p - 1,$$

$$\epsilon_{req} \geqq 2p\epsilon_f.$$

We now return to the discussion of the approximation error. We assume again that an exact calculation takes place, but that the required error and the error in the derivatives satisfy (3.34). Substituting (3.34) into (3.22) gives

(3.35)
$$|\tilde{F}(x) - F(x)| < \frac{4\epsilon_f}{3} \sum_{q=1}^{p-1} \left(\frac{k}{\pi(m-1)} \right)^{q-1}.$$

This may be further weakened, and written in the form:

THEOREM 3.36.   *Under the hypothesis of Theorems* 3.16 *and* 3.21,

$$|\tilde{F}(x) - F(x)| \leqq (4\epsilon_f(p-1)/3)(k/\pi(m-1))^{p-2}, \qquad m \leqq k/\pi + 1,$$

$$\leqq (4\epsilon_f(p-1)/3), \qquad\qquad\qquad m \geqq k/\pi + 1.$$

The significance of this result lies in the fact that the coefficient $4\epsilon_f(p-1)/3$ is small; in fact, in view of (3.34), it is less than $2\epsilon_{req}/3$ and so is smaller than the required accuracy.

For a fixed value of p then, the total discretisation error is bounded by

(3.37) $$|\tilde{F}(x) - f(x)| \leqq |F(x) - f(x)| + |\tilde{F}(x) - F(x)|.$$

As m is increased, the first term on the right reduces, having ultimate order $O(m^{-p+1})$. The second term also reduces, satisfying the bound given in Theorem 3.36. At first, while m is significantly less than k/π, it reduces in a manner consistent with an ultimate order $O(m^{-p+2})$. However, when m reaches the value $k/\pi + 1$ or exceeds it, this second term falls below the value of the required accuracy and its effect on the total discretisation error is unimportant. These remarks may be placed on a mathematical footing as follows. Since $\lim_{m\to\infty} F(x) = f(x)$, we may assert that there exists an integer valued function $m_E(\epsilon)$ having the property

(3.38) $$|F(x) - f(x)| < \epsilon \quad \text{for all } m \geqq m_E(\epsilon).$$

If we employ (3.37), (3.34) and Theorem 3.36, we find immediately

THEOREM 3.39. *For* $\epsilon \geqq \epsilon_{req}$,

$$|\tilde{F}(x) - f(x)| < \epsilon \quad \text{for all } m > m_A(\epsilon),$$

where $m_A(\epsilon) \geqq \max(m_E(\epsilon/3), 1 + k/\pi)$.

Consequently, if we were able to compare numerically the progress of a calculation using exact derivatives and one using approximate derivatives, we would find the following situation. If the exact derivative calculation gives a result whose tolerance is less than $\epsilon_{req}/3$ for values of m greater than $m_E(\epsilon_{req}/3)$, then the approximate derivative calculation gives a result whose tolerance is less than ϵ_{req} for values of m greater than

(3.40) $$m_A = \max(m_E(\epsilon_{req}/3), 1 + k/\pi).$$

This statement is rigorously exact when all calculations are exact, the only conditions being that Δ_{q-1} and ϵ_{req} satisfy (3.34).

The consequence of this section is the important result in this paper. Namely, one may apply the method and obtain a representation, even if the derivatives are not exact, so long as they are computationally reasonable. The sole consequence of using approximate derivatives is that possibly more function evaluations may be needed to attain a particular accuracy. The ultimate accuracy of the result is not impaired.

There is one further point about the use of approximate derivatives which is important in the implementation of the calculation. This is that they should not affect any practical convergence criterion by possibly replacing a smooth approach to a limit by a spasmodic or erratic approach. This point is considered briefly in Appendix 3 where it is shown that in the implementation to be described in Section 4 this is not a hazard.

The foregoing discussion has been directed towards the behaviour of the overall error in the approximation $\tilde{F}(x)$. As such it represents a discussion of the breakdown of Theorem 2.17 when inaccurate derivatives are used and of how this theorem can be replaced.

It is interesting to note that Theorem 2.20 is valid quite independently of the accuracy of the derivatives used. Thus

THEOREM 3.41.

$$\tilde{F}(j/m) = F(j/m) = \tilde{f}(j/m), \qquad j = 0, 1, \cdots, m.$$

The proof of this follows precisely the same lines as the proof of Theorem 2.20. That proof depends on the vanishing of multiplying factors $\alpha_t(j/m)$ and $\beta_t(j/m)$ in (2.11) and is independent of the accuracy of Fourier coefficients $C^{(t)}g$ and $S^{(t)}g$.

In general, one would expect Theorems 2.18 and 2.19, which refer to the trigonometric and algebraic degree of $F(x)$, to be completely invalidated if the derivatives used are not exact. However, a state of affairs of minor theoretical interest comes about when these derivatives are approximated by means of finite-difference approximations of the form

$$(3.42) \qquad \tilde{f}^{(q)}(x_0) = L_d^{(q)}(x_0)f = \sum_i a_i^{(q)} f(x_0 + x_i)$$

which are exact when $f(x)$ is a polynomial of degree d or less. When approximations of this nature are used, the degrees of $\tilde{F}(x)$ may be retained. Theorems analogous to 2.18 and 2.19 include the following:

THEOREM 3.43.   *If $f(x)$ is a trigonometrical polynomial of degree $m/2 - 1$ and if each pair of derivatives $f^{(q-1)}(1)$ and $f^{(q-1)}(0)$ are approximated by the same finite-difference approximations*

$$\tilde{f}^{(q-1)}(x_0) = L_d^{(q-1)}(x_0)f, \qquad x_0 = 1, 0,$$

then

$$(3.44) \qquad \tilde{F}(x) = F(x) = f(x).$$

THEOREM 3.45.   *If $f^{(q-1)}(1)$ and $f^{(q-1)}(0)$ are approximated by finite-difference approximations of algebraic degree d_1 and d_2, respectively, then*

$$\tilde{F}(x) = F(x) = f(x),$$

whenever $f(x)$ is an algebraic polynomial of degree $\min(d_1, d_0, p - 1)$.

For example, if forward differences at $x = 0$ and backward differences at $x = 1$ are employed, the algebraic degree of the approximation may be retained, but the trigonometrical degree is reduced to zero. On the other hand, if central (or forward) differences of the same order are used at both endpoints, the trigonometrical degree of the approximation is retained (and is $m/2 - 1$) and also the algebraic degree may be retained.

4. Organisation of the Calculation.   In this section, the organisation of a calculation based on the preceding formulas is described.

The purpose of the calculation is to determine values of a set of parameters

$$(4.1) \qquad \begin{aligned} p; \lambda_{q-1}, & \qquad q = 1, 2, \cdots, p - 1, \\ m; \mu_r, \nu_r, & \qquad r = 0, 1, \cdots, m/2, \end{aligned}$$

for use in an approximate representation

$$(4.2) \qquad \tilde{F}(x) = \sum_{q=1}^{p-1} \lambda_{q-1} B_q(x)/q! + 2 \sum_{r=0}^{m/2}{}'' \mu_r \cos 2\pi rx + \nu_r \sin 2\pi rx$$

which hopefully has the property

(4.3) $$|\tilde{F}(x) - f(x)| < \epsilon_{req}, \qquad 0 \leqq x \leqq 1,$$

where ϵ_{req} is the required tolerance. For the purposes of this description, we suppose that the user may evaluate $f(x)$ for all values of x within the unit interval $[0, 1]$. However, all that is required is a table of function values on an equally spaced sufficiently fine mesh together with the ability to obtain approximations to the derivatives $f^{(q)}(1)$ and $f^{(q)}(0)$.

It is convenient to describe such a program in terms of six stages. First, these stages are briefly listed. Then some of the computational aspects of each stage are discussed in more detail.

Stage 1. Assign a value of p.

Stage 2. Determine a set of approximations

(4.4) $$\tilde{f}^{(q-1)}(1), \qquad \tilde{f}^{(q-1)}(0), \qquad q = 2, 3, \cdots, p - 1,$$

to the corresponding derivatives $f^{(q-1)}(1)$, $f^{(q-1)}(0)$. Then set

(4.5) $$\lambda_0 = f(1) - f(0),$$
$$\lambda_{q-1} = \tilde{f}^{(q-1)}(1) - \tilde{f}^{(q-1)}(0), \qquad q = 2, 3, \cdots, p - 1,$$

and define the function $\tilde{g}_p(x)$ as

(4.6) $$\tilde{g}_p(x) = f(x) - \sum_{q=1}^{p-1} \lambda_{q-1} B_q(x)/q!.$$

Stage 3. Assign a value of m.

Stage 4. Determine μ_r and ν_r as a set of trapezoidal rule sums as follows:

(4.7) $$\mu_r = \tilde{a}_{r,p}^{[m,1]} = R_x^{[m,1]}(\tilde{g}_p(x) \cos 2\pi r x), \qquad r = 0, 1, \cdots, m/2,$$
$$\nu_r = \tilde{b}_{r,p}^{[m,1]} = R_x^{[m,1]}(\tilde{g}_p(x) \sin 2\pi r x), \qquad r = 0, 1, \cdots, m/2.$$

By definition $\nu_0 = \nu_{m/2} = 0$.

Stage 5. Calculate an error estimate. (This is discussed below.) If this estimate is less than ϵ_{req}, then the calculation is complete.

Stage 6. If this estimate is not less than ϵ_{req}, either reassign p and return to Stage 2 above, or reassign m and return to Stage 4 above.

Almost any serious attempt to calculate a set of parameters is likely to conform to some extent to the description given above. The degree of sophistication of the program will depend on the context in which it is subsequently used and on the human time devoted to its construction by the user.

In a context in which function evaluation is not expensive, the user may attempt to assign values of p and m which hopefully produce a sufficiently accurate result without recourse to iteration. As a very rough rule of thumb, the choices $p = 7$ or 8 and $m = 32$ produce an accuracy of order

(4.8) $$\epsilon \simeq 10^{-7} \max_{0 \leqq x \leqq 1} |f(x)|$$

for 'well behaved' functions. If, at Stage 5 (discussed below), one finds that the result is not sufficiently accurate, the iteration could consist of simply replacing m by $2m$ and returning to Stage 4.

We now discuss the various stages separately.

Stage 2. Whatever value of p is assigned, it is usually convenient to calculate

approximations of all derivatives up to a maximum (say $p = 12$) the first time Stage 2 is encountered. Then if subsequently a new choice of p is made, all that is entailed is a change in the value of p in (4.6). The derivatives may be approximated using forward and backward differences at $x = 0$ and 1, or, more accurately, using central differences at both $x = 0$ and 1; again, a method based on Cauchy's integral representation (see Lyness and Sande [9]) may be used. These all involve only function values either within the interval [0, 1] or near the endpoints $x = 0$ and $x = 1$ or in the complex plane. Alternatively, one may use analytic expressions for the derivatives or approximate analytic expressions. The choice simply depends on what is readily available or convenient. There is no need for the later derivatives to be particularly accurate. One might expect successive derivatives to be successively less accurate, the final derivative having perhaps 10% or 20% accuracy.

Stage 4. There are many formulas which can be used in various circumstances to shorten this calculation. (In the following description, only the formulas relevant to the even part of the trigonometrical expansion are given explicitly. Precisely analogous formulas with sin replacing cos exist.)

In view of (4.6), it follows that

$$(4.9) \qquad R_x^{[m,1]}(\tilde{g}_p(x) \cos 2\pi rx) = R_x^{[m,1]}(f(x) \cos 2\pi rx) - \sum_{q=1}^{p-1} \lambda_{q-1} X_{r,q}^{[m,1]}$$

where

$$(4.10) \qquad X_{r,q}^{[m,1]} = R_x^{[m,1]}((B_q(x)/q!) \cos 2\pi rx).$$

These coefficients were introduced in Section 3 and simple analytic forms are given in Appendix 2. A precomputed table of these coefficients may be useful, particularly if the calculation is to be repeated with different integrand functions $f(x)$.

If Stage 4 is reentered with a different value of p but with the same value of m, again formula (4.9) should be used. Since $X_{r,q}^{[m,1]} = 0$ when q is even (and $Y_{r,q}^{[m,1]} = 0$ when q is odd), it follows that the quantities in (4.7) satisfy

$$(4.11) \qquad \begin{array}{l} \tilde{a}_{r,p-1}^{[m,1]} = \tilde{a}_{r,p}^{[m,1]}, \\[6pt] \tilde{b}_{r,p}^{[m,1]} = \tilde{b}_{r,p+1}^{[m,1]}, \end{array} \qquad p \ \text{ even.}$$

Thus, only half the approximations (4.7) need be altered when p is altered to $p + 1$, m remaining fixed.

If Stage 4 is reentered with a new value of m which is double the previous value, then, in addition to the previous function values $f(j/2m)$, $j = 0, 2, 4, \cdots, 2m$, one requires m new function values $f(j/2m)$, $j = 1, 3, 5, \cdots, 2m - 1$. Convenient formulas for this updating procedure may be based on the identity

$$(4.12) \qquad R^{[2m,1]}\psi = \tfrac{1}{2}[R^{[m,1]}\psi + R^{[m,0]}\psi].$$

This may be used in several ways. For example, one may set

$$R_x^{[2m,1]}(\tilde{g}_p(x) \cos 2\pi rx) = \tfrac{1}{2}R^{[m,1]}(f(x) \cos 2\pi rx) + \tfrac{1}{2}R^{[m,0]}(f(x) \cos 2\pi rx)$$

$$(4.13)$$

$$- \sum_{q=1}^{p-1} \lambda_{q-1} X_{r,q}^{[2m,1]}$$

in which case only formulas for $X_{r,q}^{[2m,1]}$ are required. Alternatively, one may use the

formulas for $X_{r,q}^{[m,0]}$, given in (A.2.17), to calculate

$$(4.14) \qquad R_x^{[m,0]}(\tilde{g}_p(x) \cos 2\pi rx) = R^{[m,0]}(f(x) \cos 2\pi rx) - \sum_{q=1}^{p-1} \lambda_{q,1} X_{r,q}^{[m,0]}$$

and then apply (4.12) directly with $\psi(x) = \tilde{g}_p(x) \cos 2\pi rx$.

In whatever manner Stage 4 is coded, the fast Fourier transform technique (see for example Gentleman and Sande [9]) may be used for the evaluation of the sets of finite Fourier transforms of the form $R_x^{[m,\alpha]}(\varphi(x) \cos 2\pi rx)$ as and when they occur. However, in view of the small value of m involved, the saving is not very significant.

Stage 5 (Error Estimate).

(i) *The Observed Error Estimate.* The simplest, most reliable but least economic method is to evaluate the error

$$(4.15) \qquad\qquad\qquad E(x) = f(x) - \tilde{F}(x)$$

at a few carefully selected values of x. It follows from Theorem 2.20 that

$$(4.16) \qquad\qquad E(j/m) = 0, \qquad j = 0, 1, \cdots, m,$$

and an evaluation at one or two of these points provides an estimate of the overall roundoff error level (and a check against coding errors). The largest values of $E(x)$ occur in general near the ends of the interval and in regions where $f(x)$ has high derivatives such as peaks or rapid oscillations. Consequently, one may evaluate $E((2j-1)/2m)$ with j taking a few selected values, including $j = 1$ and $j = m$, and take the value with largest magnitude as an error estimate. This estimate involves further function evaluation, but if the error is larger than ϵ_{req}, these function values are required subsequently when m is replaced by $2m$.

(ii) *A Theoretical Error Estimate.* An alternative error estimate which does not involve further function evaluation may be based on the currently available set of approximations $a_{r,p}^{[m,1]}$ and $b_{r,p}^{[m,1]}$ to $C^{(r)}g_p$ and $S^{(r)}g_p$. This is an asymptotic estimate, based on the error bound

$$(4.17) \qquad |F(x) - f(x)| \leqq 4 \sum_{t=m/2+1}^{\infty} |C^{(t)}g_p| + 4 \sum_{t=m}^{\infty} |S^{(t)}g_p|$$

(see (2.16)) and on the asymptotic behaviour of the Fourier coefficients

$$(4.18) \qquad\qquad C^{(r)}g_p \simeq O(r^{p_1}), \qquad S^{(r)}g_p \simeq O(r^{p_2})$$

where

$$(4.19) \qquad \begin{array}{lll} p_1 = p, & p_2 = p + 1, & p \text{ even,} \\ p_1 = p + 1, & p_2 = p, & p \text{ odd.} \end{array}$$

One used the values of $a_{r,p}^{[m,1]}$ and $b_{r,p}^{[m,1]}$ to estimate bounds K_1 and K_2 which hopefully satisfy

$$(4.20) \qquad\qquad |C^{(r)}g_p| < K_1 r^{p_1}, \qquad |S^{(r)}g_p| < K_2/r^{p_2}.$$

For example, one may set

$$K_1 = \max\left(|a^{[m,1]}_{m/4,p}| \left(\frac{m}{4}\right)^{p_1}, \; |a^{[m,1]}_{3m/8,p}| \left(\frac{3m}{8}\right)^{p_1}, \; \tfrac{1}{2}|a^{[m,1]}_{m/2,p}| \left(\frac{m}{2}\right)^{p_1}\right),$$

(4.21)

$$K_2 = \max\left(|b^{[m,1]}_{m/4,p}| \left(\frac{m}{4}\right)^{p_2}, \; |b^{[m,1]}_{3m/8,p}| \left(\frac{3m}{8}\right)^{p_2}\right),$$

but there are many other more sophisticated ways of doing this. If K_1 and K_2 do in fact satisfy (4.20), it follows from (4.17) that

(4.22) $$|F(x) - f(x)| \leqq 4K_1\zeta(p_1, 1 + m/2) + 4K_2\zeta(p_2, m/2).$$

Using standard bounds on the generalized zeta function $\zeta(s, a)$, we find

(4.23) $$|F(x) - f(x)| \leqq \frac{4K_1}{(p_1 - 1)((m + 1)/2)^{p_1-1}} + \frac{4K_2}{(p_2 - 1)((m - 1)/2)^{p_2-1}}.$$

This is a rigorous bound only if K_1 and K_2 satisfy (4.20) for $r \geqq 1 + m/2$ and $r \geqq m$, respectively.

In the numerical example given below, the theoretical error estimate is

$$E_2 = \max\left(\frac{4K_1}{(p_1 - 1)((m + 1)/2)^{p_1-1}}, \; 2|a^{[m,1]}_{m/2,p}|\right)$$

(4.24)

$$+ \max\left(\frac{4K_2}{(p_2 - 1)((m - 1)/2)^{p_2-1}}, \; 4|b^{[m,1]}_{m/2-1,p}|\right).$$

This bound is justifiable theoretically if the derivatives are calculated from an analytic formula and if the value of m is large enough (in an asymptotic region) for (4.20) to form a valid approximation when $r \geqq m/4$. The numerical example given below illustrates the sort of values of m and p for which this estimate is valid.

In cases in which the derivatives are calculated numerically, this theoretical estimate cannot be justified. Nevertheless, it has been found to work quite well in several examples. A proper procedure would be based on inferring the value of the sum of the Fourier coefficients in (4.17) from the approximations already available without relying on order relations (4.18). It is shown in Appendix 3 that the use of approximate derivatives does not introduce erratic behaviour into the sequence of approximations, and so a 'practical convergence criterion' based on such a sequence can be constructed.

Numerical Example. Some of these points are illustrated in the following numerical example.

(4.25) $$f(x) = 1/(x - \lambda)^2 + \mu^2, \qquad \lambda = 0.3; \beta = 0.2.$$

This function has a peak at $x = \lambda = 0.3$ where $f(\lambda) = 25$. Its values at the endpoints are $f(0) = 7.7$ and $f(1) = 1.9$. It is positive definite and its integral (or mean value) over the interval $[0, 1]$ is $If = 11.4$.

The numbers given in the table are some of the results obtained using a double precision Fortran code on an IBM 195 computer. The machine accuracy parameter is 0.2×10^{-15}. The derivatives were calculated from analytical formulas. In the table are given only the observed error estimate and the theoretical error estimate for $1 \leqq p \leqq 12$ and $m = 16, 32, 64$.

TABLE 1

p	m = 16		m = 32		m = 64	
	Observed	Theoretical	Observed	Theoretical	Observed	Theoretical
1	.29E + 1	—	.29E + 1	—	.29E + 1	—
2	.25E 0	.92E 0	.12E 0	.41E 0	.58E − 1	.21E 0
3	.15E − 1	.22E 0	.36E − 2	.54E − 2	.83E − 3	.15E − 2
4	.16E − 2	.52E − 1	.12E − 3	.63E − 3	.15E − 4	.38E − 4
5	.18E − 2	.34E − 1	.48E − 5	.32E − 3	.27E − 6	.39E − 6
6	.18E − 2	.11E − 1	.71E − 6	.13E − 3	.21E − 7	.42E − 7
7	.18E − 2	.10E − 1	.32E − 6	.56E − 4	.56E − 8	.69E − 8
8	.18E − 2	.85E − 2	.93E − 7	.23E − 4	.30E − 9	.15E − 8
9	.18E − 2	.85E − 2	.10E − 6	.10E − 4	.46E − 10	.11E − 8
10	.18E − 2	.85E − 2	.10E − 6	.46E − 5	.30E − 11	.31E − 9
11	.18E − 2	.85E − 2	.10E − 6	.23E − 5	.50E − 11	.22E − 9
12	.18E − 2	.85E − 2	.10E − 6	.10E − 5	.10E − 10	.65E − 10

Table of error estimates corresponding to numerical example (4.25). The observed error estimate is

$$E_1 = \max_{1 \le i \le m} |\tilde{F}((2j - 1)/2m) - f((2j - 1)/2m)|$$

The theoretical error estimate is (4.24) above.

Stage 6. The overall efficiency of the program depends significantly on the care with which the error estimate (Stage 5) and the subsequent decision process in Stage 6 is programmed. There are several points to bear in mind.

(a) For a fixed m, as p is increased from 1, the error at first decreases, then levels off and then (unless $f(x)$ is an entire function of order 1) starts to increase in a rather spasmodic manner. The value of p at which this leveling off occurs is different for different values of m, generally increasing monotonically with m.

(b) If m is to be increased at all, at least m further function values will be required. Thus, the only economy effected by choosing a new value of m less than $2m$ is in the data manipulation section involving the fast Fourier transform in Stage 4. In special cases, such as where there are storage space constraints or in which the value of m is unduly large, there may be an advantage in choosing the new value of m to be less than double the current value. But, in general, the most convenient procedure is to double the current value.

(c) In subsequent applications of the representation, one may wish to use as short a trigonometrical series as possible. With such applications in mind, one may still construct a series with an unduly large value of m; once a representation, together with an error estimate is available, the tail of the trigonometric series may be truncated before use in applications. This procedure is more reliable than trying to gauge an appropriate value of m in the course of the calculation.

For example, suppose in the example illustrated one had sought a representation with $\epsilon_{req} = 0.5 \times 10^{-7}$ and one had ultimately decided on $p = 9$; $m = 64$. The

theoretical error estimate is $.011 \times 10^{-7}$ and this is likely to be an overestimate. Thus, we have an amount of error 0.39×10^{-7} in hand and we may omit from the final representation any set of terms whose total contribution is known to be less than this amount. Examination of the list of μ_r and ν_r (not given here) shows that

$$2 \sum_{r=17}^{32} |\mu_r| \simeq 0.19 \times 10^{-7}, \qquad 2 \sum_{r=17}^{32} |\nu_r| \simeq 0.02 \times 10^{-7}.$$

Thus, if the trigonometrical series is truncated by omitting terms $2\mu_r \cos 2\pi rx$ and $2\nu_r \sin 2\pi rx$ for which $r \geq 17$, the error incurred as a result of the truncation is less than 0.21×10^{-7}. An overall error estimate is $(0.21 + 0.011) \times 10^{-7} \simeq 0.221 \times 10^{-7}$. This is a bound so long as a theoretical estimate is indeed a bound.

Inaccurate Derivatives. To assess the effect of inaccurate derivatives, the author has constructed a straightforward numerical differentiation routine. This requires a step size H and employs function values at x_0, $x_0 \pm (2j + 1)H/2$, $j = 1, 2, \cdots, 7$. The routine is based on extrapolation. It constructs a generalised Romberg table, discards extreme elements and averages the values of the retained elements. It returns a set of approximation $\tilde{f}^{(s)}(x_0)$, $s = 1, 2, \cdots, 12$, together with a crude error estimate for each approximation.

The calculation was carried out 12 more times with approximate derivatives obtained from the numerical differentiation routine with $H = 1/16, 1/32, 1/64, \cdots$, $1/32768$, respectively. The first value of H gave wild results and the second gave usable results, an accuracy of 10^{-5} being attainable with $p \simeq 6$ and $m = 64$. The next four, i.e., those with $H = 1/64, 1/128, 1/256$ and $1/512$, gave results which virtually reproduced Table 1 (to within either 10% or 10^{-11}). Taking successively smaller values of H had the effect of destroying a successively larger portion of the lower part of the table, leaving the upper portion virtually identical with Table 1. For example, the results with $H = 1/4096$ correspond almost precisely with those in Table 1 up to and including $p = 7$. Thereafter, the error estimates for each value of m increase rapidly giving numbers of order 10^{-2} at $p = 11$ and 10 at $p = 12$.

General Remarks. In the example just given, the function $f(x)$ has a slowly converging Fourier series. In fact, approximately 10^6 terms of this series are required before the magnitudes of the terms fall below 10^{-7}. It is simply in order to avoid this sort of calculation that the method given here is constructed. However, for functions which have a rapidly converging Fourier series, the method given here is unnecessary. Thus, if $f(x)$ is periodic with period 1 (and of course analytic), then the Fourier coefficients decay exponentially. In this case,

$$f^{(q)}(1) - f^{(q)}(0) = 0$$

and so $h_p(x) = 0$. A straightforward finite Fourier transform method is indicated in this case.

It should be noticed that the calculation can be split into parts of even and odd parity. We may define

$$f_E(x) = \tfrac{1}{2}(f(x) + f(1 - x)), \qquad f_O(x) = \tfrac{1}{2}(f(x) - f(1 - x))$$

and the representation of $f_E(x)$ includes only terms involving $B_q(x)$ with q even and the cosine term part of the Fourier series. Naturally, for a function which is known to be even (or odd), only the corresponding half of the calculation described above

need be carried out. (Only half the number of function evaluations is then necessary since $f(x) = \pm f(1 - x)$.) However, care is necessary in a case in which the even part of $f(x)$ has significantly different characteristics from the odd part. For example, if the even part is very much smaller than the odd part, then a table of error estimates corresponding to Table 1 would show correspondingly little change when p is increased from an odd integer to an even integer, the significant change occurring only when p is increased to an odd integer. Any assessment on which an automatic updating procedure (in Stage 6) is based should take this possibility into account. The same sort of phenomenon occurs when one part is periodic with period 1 but the other part is not.

5. Concluding Remarks. The idea of expanding a given function in a series of Bernoulli polynomials is classical and it is well known that such a series is usually not convergent. The idea of truncating this series and expressing the remainder term as a Fourier series (which is rapidly convergent) is described by Lanczos in [4]. While he does give some examples, he limits attention to cases where only very low-order derivatives are required. In the notation of this paper, he uses $p = 1$ or 2.

The main result of this paper is that one may use large values of p and even if the derivatives are inaccurate, the technique can be used safely to obtain results of high accuracy.

One of the advantages of this representation is that it is relatively simple to construct when the allowed error ϵ_{req} is preset, and the user may have reasonable confidence that

$$|\tilde{F}(x) - f(x)| < \epsilon_{req}, \qquad 0 \leq x \leq 1.$$

The disadvantage is that it involves both algebraic and trigonometric polynomials. This restricts to some extent the possible applications.

The author is grateful to the referee whose comments were extremely helpful in preparing the final form of this paper.

Appendix 1. The Discretisation Error for the Offset Trapezoidal Rule. In this appendix, we outline the theory which corresponds to that of Section 2 in the case in which a general offset trapezoidal rule

$$(A1.1) \qquad R_x^{[m,\alpha]}(f(x)) = \frac{1}{m} \sum_{j=0}^{m-1} \tilde{f}((j + t_\alpha)/m), \qquad t_\alpha = (1 + \alpha)/2,$$

is used in place of the endpoint rule (with $\alpha = 1$). Thus

$$(A1.2) \qquad g(x) = Ig + 2 \sum_{r=1}^{\infty} (C^{(r)}g \cos 2\pi rx + S^{(r)}g \sin 2\pi rx)$$

is approximated by

$$
\begin{aligned}
G(x) = {}& a_0^{[m,\alpha]} + 2 \sum_{r=1}^{m/2-1} a_r^{[m,\alpha]} \cos 2\pi rx + \lambda a_{m/2}^{[m,\alpha]} \cos 2\pi mx \\
& + 2 \sum_{r=1}^{m/2-1} b_r^{[m,\alpha]} \sin 2\pi rx + \mu b_{m/2}^{[m,\alpha]} \sin 2\pi mx
\end{aligned}
$$

(A1.3)

where

(A1.4) $a_r^{[m,\alpha]} = R_v^{[m,\alpha]}(g(y) \cos 2\pi ry)$, $b_r^{[m,\alpha]} = R_v^{[m,\alpha]}(g(y) \sin 2\pi ry)$.

The coefficients λ and μ are usually chosen to be

(A1.5) $\lambda = \mu = 1$.

However, we leave this assignment open pro tem. Our first aim is to express the discretisation error $G(x) - g(x)$ in terms of the Fourier coefficients of $g(x)$. To this end, we employ a finite form of the Poisson summation formula, namely

(A1.6) $R_z^{[m,\alpha]}(f(x)) = If + 2\sum_{s=1}^{\infty}(C^{(ms)}f \cos 2\pi st_\alpha + S^{(ms)}f \sin 2\pi st_\alpha)$.

This may be applied directly to the function $g(x)$ to obtain the error incurred in approximating Ig by $a_0^{[m,\alpha]}$. For the corresponding approximations for the Fourier coefficients, we apply (A1.6) to the functions $g(x) \cos 2\pi rx$ and $g(x) \sin 2\pi rx$. Using the identity

(A1.7) $2C_z^{(ms)}(g(x) \cos 2\pi rx) = C^{(ms+r)}g + C^{(ms-r)}g$

together with three other similar identities, we find without difficulty that

(A1.8a)
$$a_r^{[m,\alpha]} - C^{(r)}g = \sum_{s=1}^{\infty}(C^{(ms+r)}g + C^{(ms-r)}g) \cos 2\pi st_\alpha$$
$$+ \sum_{s=1}^{\infty}(S^{(ms+r)}g + S^{(ms-r)}g) \sin 2\pi st_\alpha$$

and

(A1.8b)
$$b_r^{[m,\alpha]} - S^{(r)}g = \sum_{s=1}^{\infty}(-S^{(ms+r)}g + S^{(ms-r)}g) \sin 2\pi st_\alpha$$
$$+ \sum_{s=1}^{\infty}(S^{(ms+r)}g - S^{(ms-r)}g) \cos 2\pi st_\alpha.$$

Substituting these expressions into the expression for $G(x) - g(x)$, obtained by taking the difference between (A1.2) and (A1.3), gives

(A1.9) $G(x) - g(x) = \sum_{t=1}^{\infty} C^{(t)}g\alpha_t(x) + S^{(t)}g\beta_t(x)$

with the functions $\alpha_t(x)$ and $\beta_t(x)$ given by

$$\alpha_t(x) = \beta_t(x) = 0, \quad t < m/2,$$
$$\alpha_t(x) = 2\{\cos 2\pi(st_\alpha + rx) - \cos 2\pi(ms + r)x\},$$
$$\beta_t(x) = 2\{\sin 2\pi(st_\alpha + rx) - \sin 2\pi(ms + r)x\},$$

(A1.10)
$$t = ms + r, \quad |r| < m/2,$$
$$\alpha_{ms+m/2}(x) = 2v \cos \pi(2s+1)t_\alpha - 2\cos(\pi m(2s+1)x),$$
$$\beta_{ms+m/2}(x) = 2v \sin \pi(2s+1)t_\alpha - 2\sin(\pi m(2s+1)x),$$

(A1.11) $v = v(\lambda, \mu, t_\alpha, mx) = \lambda \cos \pi t_\alpha \cos \pi mx + \mu \sin \pi t_\alpha \sin \pi mx$.

If we denote by M_v the maximum value attainable by v, namely,

(A1.12)
$$|v| \leqq M_v = \left|\frac{\lambda + \mu}{2}\right| + \left|\frac{\lambda - \mu}{2}\right|$$

and set

(A1.13)
$$M_1 = \max(M_v, 1),$$

then

(A1.14)
$$|\alpha_t(x)| \leqq 4M_1, \qquad |\beta_t(x)| \leqq 4M_1.$$

Thus

(A1.15)
$$|G(x) - g(x)| \leqq 4M_1 \sum_{t=m/2}^{\infty} |C^{(t)}g| + |S^{(t)}g|.$$

In the normal case, $\lambda = \mu = 1$ and $M_1 = 1$ and so inequality (A1.15) is identical with the corresponding inequality (2.16) of Section 2. However, with any assignment of λ and μ, the analogue of Theorem 2.17 follows from (A1.15).

The series manipulation described above which leads to Eq. (A1.9) requires justification as some of the series involved may be only conditionally convergent. In fact, it follows that, so long as the Fourier series (A1.2) is convergent for values of x required in the rule sum evaluations, namely $x = (j + t_\alpha)/m$, then the series (A1.6) and (A1.8) are also convergent, since these are finite sums of (A1.2) for various values of j with finite weight coefficients. The step to (A1.9) requires reordering, but this is only a local reordering and may be justified without difficulty using standard techniques.

We conclude this appendix by outlining the situation with respect to the choice of λ and μ. First, we note that the final terms in (A1.3), those with $r = m$ are in essence simply alternating sums over the function values. Thus

(A1.16)
$$\lambda a_{m/2}^{[m,\alpha]} \cos \pi m x + \mu b_{m/2}^{[m,\alpha]} \sin \pi m x$$
$$= \frac{1}{m} \sum_{j=0}^{m-1} (-1)^j \bar{g}\left(\frac{t_\alpha + j}{m}\right) v(\lambda, \mu, t_\alpha, mx).$$

The condition for the analogue of Theorem 2.20 to be valid, i.e., for

(A1.17)
$$\bar{G}\left(\frac{t_\alpha + j}{m}\right) = \bar{g}\left(\frac{t_\alpha + j}{m}\right), \qquad j = 0, 1, 2, \cdots, m - 1,$$

is

(A1.18)
$$\lambda \cos^2 \pi t_\alpha + \mu \sin^2 \pi t_\alpha = 1.$$

While $G(x)$ is an approximation to $g(x)$ of trigonometric degree $m/2 - 1$, no choice of λ and μ will make it of degree $m/2$. If we require the function

(A1.19)
$$\cos(\pi m x - \theta), \qquad 0 \leqq \theta \leqq \pi,$$

to be integrated exactly, λ and μ satisfy

(A1.20) $\lambda \cos \pi t_\alpha = \cos \theta / \cos(\pi t_\alpha - \theta), \qquad \mu \sin \pi t_\alpha = \sin \theta / \cos(\pi t_\alpha - \theta).$

If λ and μ satisfy these conditions, it follows that they also satisfy condition (A1.18)

above and it also follows that

(A1.21) $$v(\lambda, \mu, t_\alpha, mx) = \cos(\pi mx - \theta)/\cos(\pi t_\alpha - \theta).$$

Thus, if the calculation is carried out with some value of α, and subsequently the result is to be adjusted to fit the function $\cos(\pi mx - \theta)$, only the final term (A1.16) need be adjusted. This can be done for all values of θ except the one which makes $\cos(\pi t_\alpha - \theta) = 0$.

Finally, an obvious choice for λ and μ is

$$\lambda = 1; \quad \mu = 1; \quad t_\alpha = \theta/\pi$$

giving

$$v = \cos \pi(mx - t_\alpha) \quad \text{and} \quad M_v = 1.$$

Appendix 2. Properties of $X_{r,q}^{[m,\alpha]}$ and $Y_{r,q}^{[m,\alpha]}$, the Discrete Fourier Coefficients of the Bernoulli Functions. In Section 3, quantities involving trapezoidal rule sums over functions which are products of Bernoulli polynomials and sine or cosines occur. They may be defined by

(A2.1) $$Z_{r,q}^{[m,\alpha]} = X_{r,q}^{[m,\alpha]} + i Y_{r,q}^{[m,\alpha]} = R_v^{[m,\alpha]}\left(\frac{B_q(y)}{q!} e^{2\pi i r \nu}\right).$$

In the actual calculation, numerical values are needed for $\alpha = 1$ and 0, $q = 1, 2, \cdots$, $p - 1$, $r = 0, 1, 2, \cdots, \bar{m}/2$ and $m = 4, 8, 16, \cdots, \bar{m}$, or some other sequence of mesh ratios. Typical values of p and $\bar{m}$ are 10 and 32, respectively. In practice, a straightforward way of calculating these is simply to carry out the summation in each case. However, they do have simple analytic forms and simple expansions some of which are derived here.

Two special cases are:

(i) $q = 0$.

$$X_{r,0}^{[m,\alpha]} = \cos 2\pi(r/m)t_\alpha, \qquad r/m = \text{integer or zero},$$

(A2.2) $$Y_{r,0}^{[m,\alpha]} = \sin 2\pi(r/m)t_\alpha, \qquad r/m = \text{integer or zero},$$

$$X_{r,0}^{[m,\alpha]} = Y_{r,0}^{[m,\alpha]} = 0, \qquad \text{other values of } r.$$

(ii) $r = 0$.

(A2.3) $$X_{0,q}^{[m,\alpha]} = B_q(t_\alpha)/m^q q!, \qquad Y_{0,q}^{[m,\alpha]} = 0.$$

Making use of the generating function for the Bernoulli polynomials

(A2.4) $$\sum_{q=0}^{\infty} \frac{B_q(x)}{q!} t^q = \frac{te^{xt}}{e^t - 1}, \qquad |t| < 2\pi,$$

(see Abramowitz and Stegun [1]), we construct a generating function for $Z_{r,q}^{[m,\alpha]}$ as follows: In view of (A2.1), we have

(A2.5) $$\sum_{q=0}^{\infty} Z_{r,q}^{[m,\alpha]} t^q = R_v^{[m,\alpha]}\left(\frac{te^{\nu t}}{e^t - 1} e^{2\pi i r \nu}\right).$$

The trapezoidal rule sum on the right is simply a geometric progression which may

be summed. If we set

(A2.6) $t = 2my, \qquad \rho = \pi r/m,$

we find

(A2.7) $\displaystyle\sum_{q=0}^{\infty} y^q (2m)^q Z_{r,q}^{[m,\alpha]} = -\frac{2ye^{(2\nu+2i\rho)t_\alpha}}{1 - e^{(2\nu+2i\rho)}}, \qquad |\alpha| < 1, \quad y \neq 0.$

This relation is also valid for $y = 0$, $\rho/2\pi \neq$ integer. The right-hand side depends on r and m only in the combination $\rho = \pi r/m$. Consequently, the function $(2m)^q Z_{r,q}^{[m,\alpha]}$ is invariant under transformations which leave r/m invariant. For example,

(A2.8) $Z_{2r,q}^{[2m,\alpha]} = Z_{r,q}^{[m,\alpha]}/(2m)^q.$

The generating function (A2.7) shows that, for a fixed ρ, the function $Z_{r,q}^{[m,\alpha]}$ is a polynomial of degree $q - 1$ in t_α. This polynomial is, apart from scaling factors, the C-polynomial discussed in Chakravarti [7]. The precise relationship is

$$Z_{r,q}^{[m,\alpha]} = (q - 1)!\, m^{2-q} e^{-2\rho t_\alpha} C_{q-1}(t_\alpha, 2i\rho).$$

In the rest of this section, we confine ourselves to the cases in which $\alpha = 0$ or 1 corresponding to the trapezoidal midpoint and endpoint rules. For $\alpha = 1$, we take the mean of (A2.7) as α approaches -1 and $+1$, respectively, to obtain

(A2.9) $\displaystyle\sum_{q=0}^{\infty} y^q (2m)^q Z_{r,q}^{[m,1]} = y \coth(y + i\rho).$

For $\alpha = 0$, we find that (A2.7) reduces to

(A2.10) $\displaystyle\sum_{q=0}^{\infty} y^q (2m)^q Z_{r,q}^{[m,0]} = y \operatorname{cosech}(y + i\rho).$

Differentiating (A2.9) q times and setting $y = 0$ gives

$$q!\,(2m)^q Z_{r,q}^{[m,1]} = \left. \frac{d^q}{dy^q}\left(y \coth(y + i\rho)\right) \right|_{\nu=0}$$

$$= q \left. \frac{d^{q-1}}{dy^{q-1}} \coth(y + i\rho) \right|_{\nu=0}.$$

Thus

(A2.11)
$$Z_{r,q}^{[m,1]} = \frac{1}{(2m)^q (q-1)!} \left. \frac{d^{q-1}}{dz^{q-1}} \coth z \right|_{z=i\rho}$$

$$= \frac{1}{(2m)^q (q-1)!\, i^q} \left. \frac{d^{q-1}}{dx^{q-1}} \cot x \right|_{z=\rho}.$$

Since $X_{r,q}^{[m,1]}$ and $Y_{r,q}^{[m,1]}$ are real and, except for the factor i^q, the right-hand side of (A2.11) is real, we find at once that

(A2.12)
$$X_{r,q}^{[m,1]} = 0, \qquad q \text{ odd},$$

$$Y_{r,q}^{[m,1]} = 0, \qquad q \text{ even},$$

while one may construct a list of analytic forms for the nonzero quantities by dif-

ferentiation. Thus

$$Y_{r,1}^{[m,1]} = -(2m)^{-1} \cot \rho,$$

$$X_{r,2}^{[m,1]} = +(2m)^{-2}(\cot^2 \rho + 1),$$

(A2.13)

$$Y_{r,3}^{[m,1]} = +(2m)^{-3}(\cot^3 \rho + \cot \rho),$$

$$X_{r,4}^{[m,1]} = -(2m)^{-4}(\cot^4 \rho + \tfrac{4}{3} \cot^2 \rho + \tfrac{1}{3}),$$

and so on. Here, $\rho = \pi r/m$ and $0 < \rho < 2\pi$. A series representation follows from the series for $\cot x$, namely

$$\text{(A2.14)} \qquad \cot x = \frac{1}{x} + \sum_{k=1}^{\infty} 2^k (-1)^k \frac{B_{2k}}{(2k)!} x^{2k-1}, \qquad 0 < |x| < \pi.$$

Using this in (A2.11) gives

$$(-1)^n (2m)^{2n} X_{r,2n}^{[m,1]}$$

$$= \frac{-1}{\rho^{2n}} + \frac{1}{(2n-1)!} \sum_{k=n}^{\infty} 2^k (-1)^k \frac{B_{2k}}{(2k)!} \frac{(2k-1)!}{(2k-2n)!} \rho^{2k-2n-1},$$

(A2.15)

$$(-1)^n (2m)^{2n-1} Y_{r,2n-1}^{[m,1]}$$

$$= \frac{1}{\rho^{2n-1}} + \frac{1}{(2n-2)!} \sum_{k=n}^{\infty} 2^k (-1)^k \frac{B_{2k}}{(2k)!} \frac{(2k-1)!}{(2k-2n+1)!} \rho^{2k-2n+1}.$$

Similar results for $X_{r,q}^{[m,0]}$ and $Y_{r,q}^{[m,0]}$ may be obtained following a precisely analogous procedure, starting with (A2.10) in place of (A2.9). This leads to

$$\text{(A2.16)} \qquad Z_{r,q}^{[m,0]} = \frac{1}{(2m)^q (q-1)! \, i^{q-2}} \frac{d^{q-1}}{dx^{q-1}} \operatorname{cosec} x \bigg|_{x=\rho}$$

and

$$Y_{r,1}^{[m,0]} = -(2m)^{-1} \operatorname{cosec} \rho,$$

$$X_{r,2}^{[m,0]} = +(2m)^{-2} \operatorname{cosec} \rho \cot \rho,$$

(A2.17)

$$Y_{r,3}^{[m,0]} = +(2m)^{-3} \operatorname{cosec} \rho(\cot^2 \rho + \tfrac{1}{2}),$$

$$X_{r,4}^{[m,0]} = -(2m)^{-4} \operatorname{cosec} \rho(\cot^3 \rho + \tfrac{5}{6} \cot \rho).$$

These values may also be obtained from the corresponding values with $\alpha = 1$ since, as each is a trapezoidal rule sum,

$$\text{(A2.18)} \qquad \tfrac{1}{2}(Z_{r,q}^{[m,0]} + Z_{r,q}^{[m,1]}) = Z_{r,q}^{[2m,1]}.$$

We require in Appendix 3 a bound on the difference between these discrete Fourier transforms of $B_q(x)/q!$ and the corresponding Fourier coefficients. This is given in Theorem A2.25 below. As a preliminary, we consider the function

$$\text{(A2.19)} \qquad f(x) = 1/x - \cot x, \qquad |x| < \pi.$$

In view of (A2.14), this function has a Taylor expansion about $x = 0$ all of whose coefficients are nonnegative. It follows that

LEMMA A2.20. $f(x)$ and all its derivatives are monotonic increasing positive valued functions of x in the interval $0 < x < \pi$.

Next, we require bounds on the values of the derivatives of $f(x)$ at $x = \pi/2$. Since $\cot x = -\tan(x - \pi/2)$, the derivatives of $\cot x$ at $x = \pi/2$ coincide with the negatives of the derivatives of $\tan x$ at $x = 0$. Using this, we find without difficulty that

(A2.21) $\qquad f^{(2q)}(\pi/2) = (2q)!/(\pi/2)^{2q+1}$,

(A2.22) $\qquad f^{(2q-1)}(\pi/2) = -(2q - 1)!/(\pi/2)^{2q} + 2^{2q}(2^{2q} - 1)\,|B_{2q}|/2q$.

This second expression may be rewritten and bounded as follows

$$
f^{(2q-1)}(\pi/2) = \frac{(2q - 1)!}{(\pi/2)^{2q}}\left\{1 + \frac{2}{3^{2q}} + \frac{2}{5^{2q}} + \cdots\right\}
$$

(A2.23)

$$
\leqq \left(\frac{\pi^2 - 4}{4}\right)(2q - 1)!/(\pi/2)^{2q}, \qquad q \geqq 1.
$$

It follows from (A2.11) and (3.14) that

(A2.24) $\qquad Z_{r,q}^{[m,1]} - C_x^{(r)}\left(\dfrac{B_q(x)}{q!}\right) - iS_x^{(r)}\left(\dfrac{B_q(x)}{q!}\right) = \dfrac{-i^{-q}f^{(q-1)}(\rho)}{(2m)^q(q - 1)!}$

where, as before, $\rho = \pi r/m$.

Taking the real and imaginary parts of (A2.24), confining ourselves to values of ρ in the interval $[0, \pi/2]$, using Lemma A2.20 and the extreme values (A2.21) and (A2.23), we find the following theorem:

THEOREM A2.25. *For $q \geqq 1$ and $r \leqq m/2$,*

$$
0 < (-1)^{(q+2)/2}(X_{r,q}^{[m,1]} - C_x^{(r)}(B_q(x)/q!)) \leqq K_q/(\pi m)^q,
$$

$$
0 < (-1)^{(q+1)/2}(Y_{r,q}^{[m,1]} - S_x^{(r)}(B_q(x)/q!)) \leqq L_q/(\pi m)^q,
$$

where

$$
K_2 = (\pi^2 - 4)/4, \qquad K_{2j} > K_{2j+2} > 1,
$$

$$
K_q = 0, \qquad L_q = 1, \qquad q\ odd,
$$

$$
L_q = 0, \qquad\qquad\qquad q\ even.
$$

The bounded quantities are monotonic increasing functions of r (or identically zero) attaining the upper bound when $r = m/2$.

We conclude this section with rather a trivial theorem, which is convenient to use in Section 3.

THEOREM A2.26.

$$
E_1^{[m,1]}(x) = \bar{B}_1(x) - 2\sum_{r=1}^{m/2-1} Y_{r,1}^{[m,1]}\sin 2\pi r x
$$

satisfies inequality (3.20), namely $|E_1^{[m,1]}(x)| < \tfrac{2}{3}$.

We establish this by comparison with a function

(A2.27) $\qquad \psi(x) = \bar{B}_1(x) - 2\displaystyle\sum_{r=1}^{m/2-1} S^{(r)}(B_1(x))\sin 2\pi r x$

which represents the difference between the saw tooth function $\bar{B}_1(x)$ and the first $m/2 - 1$ terms of its Fourier expansion. It is well known that

(A2.28) $$|\psi(x)| \leqq \tfrac{1}{2}.$$

Now

(A2.29)
$$|E_1^{[m,1]}(x) - \psi(x)| = \left| 2 \sum_{r=1}^{m/2-1} (Y_{r,1}^{[m,1]} - S^{(r)}(B_1(x))) \sin 2\pi rx \right|$$
$$< \frac{1}{m} \sum_{r=1}^{m/2-1} f(\pi r/m),$$

where the inequality is based on replacing $\sin 2\pi rx$ by 1 and using (A2.24) with $q = 1$. Since $f(0) = 0$ and $f(\pi/2) = 2/\pi$ and $f'(x)$ is nonnegative in the interval $[0, \pi/2]$, it follows that

(A2.30) $$f(\pi r/m) < 4r/\pi m, \qquad 0 < r < m/2.$$

Thus

(A2.31) $$|E_1^{[m,1]}(x) - \psi(x)| < \frac{1}{m} \sum_{r=1}^{m/2-1} (4r/\pi m) = \frac{1}{2\pi} \left(1 - \frac{2}{m} \right).$$

Theorem A2.26 follows from inequalities (A2.28) and (A2.31).

Appendix 3. Effect of Approximate Derivatives on Practical Convergence Criterion. In this section, we discuss briefly the effect of approximate derivatives on the actual quantities which are used in the practical convergence criterion. The purpose of such a discussion is to verify that spasmodic behaviour is not introduced into the Fourier coefficients or their trapezoidal rule approximations in a way which might hamper the operation of the criterion.

It follows quite simply from (3.7) and other formulas in Section 3 that

(A3.1)
$$C^{(r)} \tilde{g}_p - C^{(r)} g_p = \sum_{q=1}^{p-1} \Delta_{q-1} C_x^{(r)} \left(\frac{B_q(x)}{q!} \right)$$
$$= -\frac{\Delta_1}{(2\pi r)^2} + \frac{\Delta_3}{(2\pi r)^4} + \cdots + \frac{(-1)^l \Delta_{2l-1}}{(2\pi r)^{2l}}$$

where

(A3.2) $$l = [(p-1)/2].$$

Similarly

(A3.3) $$\tilde{a}_{r,p}^{[m,1]} - a_{r,p}^{[m,1]} = \sum_{q=1}^{p-1} \Delta_{q-1} X_{r,q}^{[m,1]}.$$

Thus

(A3.4)
$$(\tilde{a}_{r,p}^{[m,1]} - C^{(r)} \tilde{g}_p) - (a_{r,p}^{[m,1]} - C^{(r)} g_p)$$
$$= \sum_{q=1}^{p-1} \Delta_{q-1} \left(X_{r,q}^{[m,1]} - C_x^{(r)} \left(\frac{B_q(x)}{q!} \right) \right).$$

The first point to note is that the functions on the right-hand side of these equations are colloquially 'smooth' functions of the integer variable r. By themselves, they are not likely to mislead a carefully constructed practical convergence criterion. Thus, if,

in practice, it is found that the sequence $\tilde{a}_{r,p}{}^{[m,1]}$, $r = 1, 2, \cdots$, is behaving erratically, this behaviour is likely to be a consequence of the fact that the sequence $a_{r,p}{}^{[m,1]}$, $r = 1, 2, \cdots$, behaves erratically and it is most unlikely to be induced by the use of approximate derivatives.

Next, we look at the magnitude of the right-hand side of (A3.1). For small values of r, this may be quite large. However, for large values of r, this difference approaches zero. Specifically, using bounds (3.34) on the values of Δ_{q-1},

(A3.5)
$$|C^{(r)}\tilde{g}_p - C^{(r)}g_p| < \frac{\epsilon_f p}{k}\left(\frac{k}{2\pi r}\right)^{2l}, \qquad r \lessgtr k/2\pi,$$

$$< \frac{\epsilon_f p}{2\pi r}\left(\frac{k}{2\pi r}\right), \qquad r \gtrless k/2\pi.$$

(While this difference approaches zero, the actual value of the coefficients also approaches zero. $C^{(r)}\tilde{g}_p$ approximates $C^{(r)}g_p$ in an absolute sense and not a relative sense.) Since the function $\tilde{g}_p(x)$ may be quite different from $g_p(x)$, one expects their Fourier coefficients to be different. This result indicates that the numerical difference is largely accounted for by the early Fourier coefficients.

Finally, we treat (A3.4) in the same way. Using the bound stated in Theorem A2.25, we find

(A3.6)
$$|(\tilde{a}_{r,p}{}^{[m,1]} - C^{(r)}\tilde{g}_p) - (a_{r,p}{}^{[m,1]} - C^{(r)}g_p)|$$

$$\leqq 2.2\frac{\epsilon_f p}{k}\left(\frac{k}{\pi m}\right)^{2l}, \qquad m \lessgtr k/\pi,$$

$$\leqq 2.2\frac{\epsilon_f p}{\pi m}\left(\frac{k}{\pi m}\right), \qquad m \gtrless k/\pi.$$

This inequality compares the error in the trapezoidal rule approximation in the two cases. These bounds are independent of r. For all values of m, the second bracket on the left-hand side is of order $O(m^{-2l-2})$. Thus for small values of m, the first bracket on the right behaves as if there were an additional term of order $O(m^{-2l})$. When m exceeds k/π, this additional term is below the required accuracy $\epsilon_{req} < 2p\epsilon_f$ by a comfortable margin. Thus, even though for small values of r, $C^{(r)}\tilde{g}_p$ and $C^{(r)}g_p$ are quite different, the different trapezoidal rule approximations approach their respective limits at much the same rate.

Completely analogous formulas relating to the sine Fourier coefficients exist, and analogous conclusions hold.

Argonne National Laboratory
9700 South Cass Avenue
Argonne, Illinois 60439

1. M. ABRAMOWITZ & I. STEGUN (Editors), *Handbook of Mathematical Functions, With Formulas, Graphs and Mathematical Tables*, Nat. Bur. Standards Appl. Math. Series, 55, Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964, Chap. 7. MR **29** #4914.
2. R. P. BOAS, *Entire Functions*, Academic Press, New York, 1954. MR **16**, 914.
3. W. B. JONES & G. HARDY, "Accelerating convergence of trigonometric approximations," *Math. Comp.*, v. 24, 1970, pp. 547–560. MR **43** #2823.

4. C. LANCZOS, *Discourse on Fourier Series*, Hafner, New York, 1966. MR **33** #7772.

5. G. J. LIDSTONE, "Notes on the extension of Aitken's theorem (for polynomial interpolation) to the Everett types," *Proc. Edinburgh Math. Soc.* (2), v. 2, 1929, pp. 16–19.

6. D. V. WIDDER, "Completely convex functions and Lidstone series," *Trans. Amer. Math. Soc.*, v. 51, 1942, pp. 387–398. MR **3**, 293.

7. P. C. CHAKRAVARTI, *Integrals and Sums*, Univ. of London, The Athlone Press, 1970.

8. W. M. GENTLEMAN & G. SANDE, *Fast Fourier Transforms—For Fun and Profit*, Proc. AF1PS 1966 FJCC, v. 29, Spartan Books, New York, pp. 563–578.

9. J. N. LYNESS & G. SANDE, "ENTCAF and ENTCRE: Evaluation of normalised Taylor coefficients of an analytic function, algorithm 413," *Comm. ACM*, v. 14, 1971, pp. 669–675.