JOURNAL OF COMPUTATIONAL PHYSICS 12, 462–470 (1973)

Computation of Fourier Integrals of Exponentials of Truncated Fourier Series*

JOHN D. THOMAS,[†] H. RALPH LEWIS, AND K. J. MELENDEZ

Los Alamos Scientific Laboratory, University of California, Los Alamos, New Mexico 87544

Received July 24, 1972

An efficient and accurate method is described for computing a class of definite integrals that have arisen in plasma physics and that arise in the theory of frequency modulated radio transmission.

1. INTRODUCTION

In this paper we discuss an efficient method for computing a class of definite integrals that can be very time-consuming to evaluate by standard techniques such as the trapezoidal rule or Gauss quadrature. The general form of the integrals in question is

$$I(s; a_k, b_k) = \frac{1}{2\pi} \int_0^{2\pi} d\varphi \exp \left\{ i \left[s\varphi + \sum_{k=1}^{M} (a_k \cos k\varphi + b_k \sin k\varphi) \right] \right\}, \tag{1}$$

where s and k are integers, and the a_k and b_k are real constants. As the magnitudes of the parameters s, a_k, and b_k increase, the integrand becomes increasingly oscillatory.

Our interest in these integrals, which are Fourier integrals of exponentials of truncated Fourier series, arose from the fact that they play an important role in the application of a particular numerical method to a nonlinear problem in plasma physics [1]. In that application it was important to be able to evaluate the integrals rapidly for a rather large range of values of the parameters (for example, $|s| < 10$, $|a_k| < 50$, and $|b_k| < 50$). The integrals are also encountered in radio engineering, as can be seen by noting that Eq. (1) is the general form of the Fourier coefficients of a frequency modulated radio signal for which the modulation itself can be represented as a Fourier series with a finite number of terms.

To evaluate the integral defined by Eq. (1) when the parameters a_k and b_k are

* Work performed under the auspices of the U.S. Atomic Energy Commission.

† Permanent address: New Mexico State University, University Park, New Mexico 88070.

small, we first transform it into a contour integral around the unit circle in the complex plane and apply the residue theorem. An infinite series representation of the the residue can be found that is absolutely convergent for all values of the parameters, and the series can be computed recursively. For fixed s the series is rapidly convergent and well-behaved for sufficiently small values of the a_k and b_k. However, for sufficiently large values of the a_k and b_k, the real and imaginary parts of the sequence of partial sums oscillate between very large positive and negative values before beginning to converge. Although, even in this case, the convergence is very rapid once it begins, the magnitude of the initial oscillations leads to serious inaccuracies in the final result with only modestly large values of the parameters a_k and b_k. However, the integral can be computed rapidly and accurately for a considerably larger range of the parameters a_k and b_k by modifying the method by the introduction of a type of scaling.

In Section 2 we derive our basic method of computing the integral that is based on the series evaluation of a residue, and we discuss the scaling modification that extends the applicability of the method. An *a posteriori* error bound is given for the basic method. In Section 3 a numerical example is presented.

2. Derivation of the Residue Method

A. *The Basic Method*

To convert the integral into a line integral around the unit circle in the complex plane, we make the substitution $z = e^{i\varphi}$ in Eq. (1) and obtain

$$I(s; a_k, b_k) = \frac{1}{2\pi i} \oint_{|z|=1} dz\, z^{s-1} \exp\left[\sum_{k=1}^{M} (\alpha_k z^k - \bar{\alpha}_k z^{-k})\right], \tag{2}$$

where

$$\alpha_k = \tfrac{1}{2}(b_k + i a_k) \tag{2a}$$

and a bar denotes complex conjugation. By the residue theorem, $I(s; a_k, b_k)$ is simply the coefficient of z^{-s} in the Laurent expansion of the function

$$f(z) = \exp\left[\sum_{k=1}^{M} (\alpha_k z^k - \bar{\alpha}_k z^{-k})\right]. \tag{3}$$

It is convenient to define functions $P(z)$ and $Q(z)$ by

$$P(z) = \sum_{k=1}^{M} \alpha_k z^k \quad \text{and} \quad Q(z) = -\sum_{k=1}^{M} \bar{\alpha}_k z^k, \tag{4}$$

so that $f(z)$ can be written as

$$f(z) = e^{P(z)} e^{Q(1/z)}. \tag{5}$$

We obtain the Laurent series for $f(z)$ by multiplying the power series for $\exp P(z)$ by that for $\exp Q(1/z)$. Thus, if we set

$$e^{P(z)} = \sum_{n=0}^{\infty} c_n z^n \qquad \text{and} \quad e^{Q(1/z)} = \sum_{n=0}^{\infty} d_n z^{-n}, \tag{6}$$

and adopt the convention

$$c_k = d_k = 0 \qquad \text{for} \quad k < 0, \tag{6a}$$

then

$$f(z) = \sum_{n=-\infty}^{\infty} e_n z^n, \tag{7}$$

where

$$\begin{aligned} e_n &= \sum_{i=0}^{\infty} c_{i+n} d_i \\ &= \sum_{i=0}^{\infty} c_i d_{i-n} \,. \end{aligned} \tag{8}$$

The coefficients c_n and d_n can be determined from recursion relations. To derive the recursion relation satisfied by the c_n, we first define

$$g(z) = e^{P(z)}, \tag{9}$$

and then differentiate $g(z)$ to obtain

$$g' = gP'. \tag{9a}$$

By applying the Leibnitz rule for the differentiation of a product to Eq. (9a) we immediately obtain

$$\begin{aligned} \frac{1}{n!} g^{(n)} &= \sum_{k=1}^{n} \frac{1}{n!} \binom{n-1}{k-1} g^{(n-k)} P^{(k)} \\ &= \frac{1}{n} \sum_{k=1}^{n} k \frac{g^{(n-k)}}{(n-k)!} \frac{P^{(k)}}{k!} \,. \end{aligned} \tag{10}$$

Finally, by combining Eq. (10) with the formula $c_n = (1/n!)\, g^{(n)}(0)$, we arrive at the recursion relation

$$c_n = \frac{1}{n} \sum_{k=1}^{M} k \alpha_k c_{n-k} \,. \tag{11}$$

The recursion relation for the d_n can be derived by a completely analogous treatment of exp $Q(z)$. The result is

$$d_n = -\frac{1}{n}\sum_{k=1}^{M} k\bar{\alpha}_k d_{n-k}\,. \tag{12}$$

These M-term recursion relations can be used to determine the c_n and d_n starting from Eq. (6a) and the values of c_0 and d_0 given by

$$c_0 = e^{P(0)} = 1 \qquad \text{and} \qquad d_0 = e^{Q(0)} = 1. \tag{13}$$

The convergence properties of the sequence formed by the numbers c_n can be discussed conveniently in terms of numbers t_n defined by

$$c_n = \frac{(\alpha_M)^{n/M}}{\Gamma(1 + (n/M))}\, t_n\,. \tag{14}$$

The M-term recursion relation, or difference equation, for the t_n corresponding to Eq. (11) is

$$t_n = \sum_{k=1}^{M-1} f_k(n)\, t_{n-k} + t_{n-M}\,, \tag{15}$$

where

$$f_k(n) = \frac{k\alpha_k(\alpha_M)^{-k/M}\,\Gamma(n/M)}{M\Gamma(1 + (n-k)/M)}\,, \tag{16}$$

and the initial data are

$$\text{and} \qquad \left.\begin{aligned} t_0 &= 1, \\ t_n &= 0 \qquad \text{for} \quad n < 0. \end{aligned}\right\} \tag{17}$$

For large n we can use the asymptotic representation of the gamma function corresponding to Stirling's formula to obtain

$$\begin{aligned} f_k(n) &= \frac{k\alpha_k(\alpha_M)^{-k/M}}{M}\,\frac{1}{(x+\delta)^{\delta}}\,\frac{e^{\delta}}{(1 + (\delta/x))^{x}}\left(1 + \frac{\delta}{x}\right)^{1/2} + O\left(\frac{1}{n}\right) \\ &= \frac{k\alpha_k(\alpha_M)^{-k/M}}{M}\left(1 + \frac{n-k}{M}\right)^{-(M-k)/M} + O\left(\frac{1}{n}\right), \end{aligned} \tag{18}$$

where $x = n/M$ and $\delta = (M - k)/M$. Thus, $f_k(n)$ tends toward zero for large n, and the difference equation for t_n tends toward a linear difference equation with constant coefficients. The general solution of Eq. (15) when the $f_k(n)$ are constant is an arbitrary linear combination of the nth powers of the roots of the polynomial

$$y^M - \sum_{k=1}^{M-1} f_k\, y^{M-k} - 1$$

as long as all of the roots are distinct. For large n this polynomial approaches $y^M - 1$, all of whose roots are distinct and of unit modulus. Therefore, for every positive real number ϵ, if the sequence $| t_n |$ is unbounded, the terms must grow more slowly than $(1 + \epsilon)^n$, and if the sequence tends toward zero, it must do so more slowly than $(1 - \epsilon)^n$. That is, either the sequence $| t_n |$ converges to a nonzero complex number, or it diverges very slowly, or tends toward zero very slowly. Therefore, because of the gamma function in the denominator of Eq. (14), the $| c_n |$ must eventually converge quite rapidly to zero. Exactly the same argument can be made for the d_n by replacing α_k by $-\bar{\alpha}_k$. Since the sequences $| c_n |$ and $| d_n |$ converge to zero rapidly, it is clear that the series for e_n given by Eq. (8) converges absolutely and rapidly.

There is a useful *a posteriori* bound on the error associated with approximating e_n by truncating the sums in Eq. (8). Let l and N be integers such that

$$l \geqslant 0 \text{ and } N \geqslant A = \sum_{k=1}^{M} k \mid \alpha_k \mid ,$$

and approximate e_l and e_{-l} by

$$\tilde{e}_l = \sum_{i=0}^{N} c_{i+l} d_i \qquad \text{and} \quad \tilde{e}_{-l} = \sum_{i=0}^{N} c_i d_{i+l} ,$$

respectively. In order to establish bounds on $| e_l - \tilde{e}_l |$ and $| e_{-l} - \tilde{e}_{-l} |$, we first prove the following two lemmas.

LEMMA 1: *For any $\epsilon > 0$ and any integer $s \geqslant 0$, if*

$$\max\{| c_{N-k} |, | d_{N-k} |; k = 1,\ldots, M\} \leqslant \epsilon,$$

then

$$| c_{N+s} | \leqslant \frac{\epsilon A}{N + s} \qquad \text{and} \quad | d_{N+s} | \leqslant \frac{\epsilon A}{N + s} .$$

The proof is by induction. Suppose that the lemma is true for $0 \leqslant s \leqslant s' - 1$.

Then,

$$
\begin{aligned}
| c_{N+s'} | &= \left| \frac{1}{N+s'} \sum_{k=1}^{M} k\alpha_k c_{N+s'-k} \right| \\
&\leqslant \frac{1}{N+s'} \sum_{k=1}^{M} k \, | \alpha_k | \, | c_{N+s'-k} | \\
&\leqslant \frac{\epsilon A}{N+s'},
\end{aligned}
$$

since $| c_{N+s'-k} | \leqslant \epsilon$ when $1 \leqslant k - s' \leqslant M$ and $| c_{N+s'-k} | \leqslant \epsilon A/(N + s' - k) \leqslant \epsilon$ when $0 \leqslant s' - k \leqslant s' - 1$. Similarly,

$$
| d_{N+s'} | \leqslant \frac{\epsilon A}{N+s'}.
$$

Thus, the lemma is true for $s = s'$ if it is true for $0 \leqslant s \leqslant s' - 1$. But the lemma is true for $s = 0$. Therefore, it is true for all $s \geqslant 0$.

LEMMA 2: *For any $\epsilon > 0$, and any integers $r \geqslant 0$ and $s \geqslant 0$, if*

$$
\max\{| c_{N+s-k} |, | d_{N+r-k} |; k = 1,..., M\} \leqslant (\epsilon/A)^{1/2},
$$

then

$$
\left| \sum_{k=1}^{\infty} c_{N+s+k} d_{N+r+k} \right| \leqslant \epsilon.
$$

The proof is:

$$
\begin{aligned}
\left| \sum_{k=1}^{\infty} c_{N+s+k} d_{N+r+k} \right| &\leqslant \sum_{k=1}^{\infty} | c_{N+s+k} | \, | d_{N+r+k} | \\
&\leqslant \sum_{k=1}^{\infty} \frac{\epsilon A}{(N+s+k)(N+r+k)} \qquad \text{(by Lemma 1)} \\
&\leqslant \epsilon A \sum_{k=1}^{\infty} \frac{1}{(N+k)^2} \leqslant \epsilon A \int_0^{\infty} \frac{dk}{(N+k)^2} \\
&\leqslant \frac{\epsilon A}{N} \leqslant \epsilon.
\end{aligned}
$$

The error bounds are expressed by the following theorem.

THEOREM: (a) *If*

$$\max\{|c_{N+l-k}|, |d_{N-k}|; k = 1,..., M\} \leqslant (\epsilon/A)^{1/2},$$

then

$$|e_l - \tilde{e}_l| \leqslant \epsilon. \tag{19a}$$

(b) *If*

$$\max\{|c_{N-k}|, |d_{N+l-k}|; k = 1,..., M\} \leqslant (\epsilon/A)^{1/2},$$

then

$$|e_{-l} - \tilde{e}_{-l}| \leqslant \epsilon. \tag{19b}$$

The proof of part (a) follows from Lemma 2 with $r = 0$ and $s = l$; the proof of part (b) follows from Lemma 2 with $r = l$ and $s = 0$.

Despite the fact that the series for e_n given by Eq. (8) converges absolutely and rapidly, the results so far do not suffice for the practical computation of $I(s; a_k, b_k)$ except when the a_k and b_k are rather small. Even for a_k and b_k of moderate magnitudes, the terms in the sequences c_n and d_n that come before rapid convergence can grow so large as to render the final value of e_n computed from Eq. (8) very inaccurate. As an illustration, consider the case $M = 1$. Then we have

$$c_n = \alpha_1{}^n/n!, \quad d_n = (-1)^n \bar{\alpha}_1{}^n/n! \qquad \text{and} \qquad e_0 = \sum_{n=0}^{\infty} (-1)^n (A^n/n!)^2,$$

where $A = |\alpha_1|$. The sequence $A^n/n!$ increases until n exceeds A. By using Stirling's formula for $n!$, which is quite good for $n \gtrsim 10$, we find that the largest term in the sequence is approximately $e^A/(2\pi A)^{1/2}$. This quantity is 2.78×10^3 for $A = 10$ and 4.33×10^7 for $A = 20$. Thus, although $|e_0|$ is bounded by unity (because the integrand in Eq. (1) is of unit modulus), the series for e_0 will contain terms of order 10^{15} if A is 20, resulting in a loss of accuracy in e_0 of at least 15 digits.

B. *Scaling Modification for Larger Values of a_k and b_k*

To modify the basic method for computing $I(s; a_k, b_k)$ so that larger values of the a_k and b_k can be allowed, we start from the observation that $I(s; \lambda a_k, \lambda b_k)$, where λ is a positive integer, is closely related to $I(s; a_k, b_k)$. In particular, if $I(s; a_k, b_k)$ is the coefficient of z^{-s} in the Laurent expansion of $f(z)$ as defined by Eq. (3), then $I(s; \lambda a_k, \lambda b_k)$ is the coefficient of z^{-s} in the Laurent expansion of $[f(z)]^\lambda$. Suppose that we want to compute $I(s; A_k, B_k)$, where the A_k and B_k are too large for direct application of the basic method described in the previous section. We first define another set of parameters, a_k and b_k, by

$$a_k = A_k/\lambda \qquad \text{and} \qquad b_k = B_k/\lambda, \tag{20}$$

where λ is a positive integer. By choosing λ sufficiently large, we arrange that the c_n and d_n computed from the a_k and b_k via Eqs. (11) and (12) die out rapidly and never exceed some predetermined magnitude that is consistent with the accuracy desired for the integral. We then compute the Laurent series for $f(z)$ by means of Eq. (8), and raise that series to the power λ. In practice we choose λ to be a power of 2 so that the series for $[f(z)]^\lambda$ can be computed from the series for $f(z)$ by repeated squaring. The integral $I(s; A_k, B_k)$ is the coefficient of z^{-s} in the series for $[f(z)]^\lambda$.

This modification is very effective even for values of A_k and B_k that are considerably larger than those for which the basic method is practicable. By choosing λ large enough we are able to calculate the series for $f(z)$ accurately and rapidly with the recursive formulas of the basic method. Once that is done, we are assured that the numbers that occur in computing the series for $[f(z)]^\lambda$ by repeated squaring of the series for $f(z)$ will not give rise to large subtraction errors. The reason is that each coefficient in the series for $[f(z)]^N$, where N is any positive integer, is bounded in magnitude by unity, because each such coefficient is the value of an integral like that defined by Eq. (1) and these integrals are all bounded by unity. This property of the series for $f(z)$ is important for the practical computation of $[f(z)]^\lambda$.

In the computation of the series for $[f(z)]^\lambda$ it is only necessary to keep those terms which, to within the accuracy required for the integral, contribute to $[f(z)]^\lambda$ on the *unit circle*, because the integral that we are evaluating is a line integral around the unit circle. Suppose, for example, that one of the series to be computed as a step in the computation of the series for $[f(z)]^\lambda$ is

$$[f(z)]^N = \sum_{i=-\infty}^{\infty} \hat{e}_i z^i, \tag{21}$$

where N is an integer satisfying $1 \leqslant N \leqslant \lambda$. Because the coefficients e_n in the series for $f(z)$ die out rapidly for increasing $|n|$, the coefficients $\hat{e}_i$ die out rapidly for increasing $|i|$. Therefore, we can approximate $[f(z)]^\lambda$ adequately on the *unit circle* by a truncated series $h(z)$ of the form

$$h(z) = \sum_{i=NL}^{NU} \hat{e}_i z^i, \tag{22}$$

where the limits NL and NU are so chosen that the $\hat{e}_i$ are sufficiently small for $i < NL$ and $i > NU$. Because of the rapid convergence of the $\hat{e}_i$, the $\hat{e}_i$, the difference between $[f(z)]^\lambda$ and $h(z)$ on the unit circle is of the same order of magnitude as the larger of $|\hat{e}_{NL}|$ and $|\hat{e}_{NU}|$. Truncating the series for $f(z)$ and its successive squares in this fashion renders the computation considerably simpler and faster.

3. A Numerical Example

The success of the scaling procedure is illustrated by the following example. We set $M = 3$, $a_k = 10/k$ and $b_k = 25/k$, so that $k\alpha_k = 12.5 + 5i$. With these data $I(s; a_k, b_k)$ was calculated for $-12 \leqslant s \leqslant 12$ from Eqs. (8), (11) and (12) in both single and double precision with a CDC 6600 computer. The number of terms in Eq. (8) needed for convergence varied between 110 and 116. Since the largest terms were of order 10^{12}, we expected errors of order 10^{-2} in the single-precision computation, and of order 10^{-16} in the double-precision computation; these expectations were realized.

We then scaled the input parameters by choosing $\lambda = 8$, and used Eqs. (8), (11), and (12) to compute in single precision all of the e_n which are of magnitude greater than 10^{-14}. This required computing the e_n for $-52 \leqslant n \leqslant 45$. The number of terms in Eq. (8) required for convergence varied between 21 and 30. For each e_n, no term in Eq. (8) was larger in magnitude than 3 nor larger than about $10 \mid e_n \mid$. The series for $f(z)$ was squared and truncated, retaining only those $\hat{e}_i$ defined by Eq. (22) for which $-71 \leqslant i \leqslant 58$; it was squared again and truncated, this time retaining the $\hat{e}_i$ for which $-102 \leqslant i \leqslant 79$; then the expression was squared a final time. The $I(s; a_k, b_k)$, thus calculated, agreed with the results of the double-precision calculation to better than 10^{-13} for all s in the range $-12 \leqslant s \leqslant 12$.

Acknowledgment

This work was performed under the auspices of the U. S. Atomic Energy Commission while J. D. Thomas had an AWU fellowship at LASL.

Reference

1. H. R. Lewis, "Methods in Computational Physics," Vol. 9, pp. 332–335, Academic Press, New York, 1970.