JOURNAL OF COMPUTATIONAL PHYSICS **56**, 349–355 (1984)

Note

On the Efficient Evaluation of Modified Bessel Functions of Zeroth and First Orders for Arguments of the Form $x \exp(i\pi/4)$

1. INTRODUCTION

Mason and Sykes [1] hereinafter MS, developed a three-dimensional model of wind flow over low hills based on the two-dimensional theory of Jackson and Hunt [2], hereinafter JH. This theory gives approximate analytic solutions to the governing equations of motion for steady flow. In the MS model $2^I \times 2^J$ grid-points represent the terrain in real space, but computations are performed in a Fourier-transformed mode. The number of complex Fourier coefficients at each vertical level selected for solution evaluation is $K = (2^{I-1} + 1)\, 2^J$; since I and J are typically 7 or 8, $K \sim 8{,}300$–$33{,}000$. For each coefficient, the modified Bessel function of the second kind, zeroth order, K_0, must be evaluated with four different arguments, two of which must be recomputed at each level, in order to determine the two components of horizontal wind velocity perturbation. In addition, K_1 must be evaluated twice for each Fourier coefficient for the purpose of computing the shear stress perturbation at the surface.

Simple arithmetic indicates that K_0 or K_1 must be evaluated $2K(N+2)$ times, where N is the number of vertical levels. Typically, $N \sim 1-5$. Hence in the MS model, Bessel-function computations are usually done 5×10^4–5×10^5 times per run. Furthermore, the arguments are complex and the evaluation is done by use of converging series (Abramowitz and Stegun [3], hereinafter AS). Clearly, one would like to do these computations as efficiently as possible while still retaining sufficient accuracy (typically four significant figures for the kind of physical problems to which the model is applied).

The requirement for efficiency becomes even more acute in the model of Walmsley *et al.* [4], hereinafter WST, who extended the MS model to allow pressure perturbations to be functions of height as well as wavenumber in the momentum equations. In WST, terms appear which are additional to those involved in MS. In fact, there are three integrals each of which involves either I_0 or K_0 in the integral. Each integral is evaluated numerically using M-point Gaussian quadrature. One of the integrals is computed over a semi-infinite domain, so evaluation is in two parts (one part involving a change of variable) and hence $2M$ Gaussian points are needed. Functions I_0, I_1, K_0, K_1, or combinations thereof, multiply the integrals. The result is an increase in the number of evaluations of Bessel functions to $4K[(M+1)$

0021-9991/84 $3.00

581/56/2-12

$(N+1)+1]$, approximately 15–20 times that of the MS model for a typical value of $M=10$ and, as before, $N \sim 1$–5. In a recent application of the MS and WST models (Walmsley and Howard [5]), computation time on a CYBER 176 for $I=J=8$, $N=1$, $M=10$ increased from 18 sec for MS to 380 sec for WST.

In Section 2, the calculation of the modified Bessel functions by ascending series and asymptotic expansions is described briefly. Then, in Section 3, a method of table-look-up and interpolation is outlined. This method is sufficiently accurate for most applications and substantially reduces the computation time for the WST model. In the above example, the time was reduced to 147 sec; for $N>1$ the relative improvement is even more spectacular, the tabulated values having already been computed.

With regard to earlier related work, Scarton [6] describes the evaluation of $J_n(z)$ and $I_n(z)$. He does not consider $K_n(z)$ and is mainly concerned with extending the argument, z, into the entire complex domain rather than efficient evaluation with specified accuracy for a particular class of arguments. Sookne [7] also evaluates $J_n(z)$ and $I_n(z)$ in double precision to an accuracy which is machine-dependent. $K_n(z)$ is not considered. Temme [8] gives a computer program for evaluating $K_\nu(z)$ and $K_{\nu+1}(z)$, where ν is real, with a specified accuracy. His method is more general and, provided more than approximately 4,500 evaluations are required, slower than the present method. A comparison will be made in Section 4.

2. Series Calculation of Modified Bessel Functions

Formulae given by AS were used to calculate the modified Bessel functions. The ascending series form for $I_n(z)$, where $z=x\exp(i\phi)$, x is real and $n \geqslant 0$ is an integer, is derived from AS Eq. (9.6.10):

$$I_n(z)=(z/2)^n \sum_{k=0}^{\infty} A_k, \qquad n \geqslant 0, \tag{1}$$

where

$$\begin{aligned} A_k &= 1/n!, & k&=0, \\ &\quad A_{k-1}(z^2/4)/[k(n+k)], & k&>0. \end{aligned}$$

For $K_n(z)$, Eqs. (9.6.11) and (9.6.13) may both be expressed as

$$\begin{aligned} K_n(z) = {} & (1/2)(z/2)^{-n} R_n \\ & + (-1)^{n+1} \ln(z/2)\, I_n(z) \\ & + (1/2)(-z/2)^n S_n, \qquad n \geqslant 0, \end{aligned} \tag{2}$$

where

$$R_n = 0, \qquad n = 0$$
$$\sum_{k=0}^{n-1} B_k, \qquad n > 0,$$
$$B_k = (n-1)!, \qquad k = 0$$
$$B_{k-1}(-z^2/4)/[k(n-k)], \qquad k > 0,$$
$$S_n = \sum_{k=0}^{\infty} [\Psi(k+1) + \Psi(n+k+1)] A_k.$$

Here $\Psi(m)$, the Psi or Digamma function is given by AS Eq. (6.3.2):

$$\Psi(m) = -\gamma, \qquad m = 1$$
$$\Psi(m-1) + 1/(m-1), \qquad m > 1, \tag{3}$$

where $\gamma = 0.577216...$ is Euler's constant. The recursive forms for A_k, B_k, C_k and $\Psi(m)$ facilitate numerical computation. The infinite series in (1) and (2), which converge for all $z \neq 0$, are considered to have converged when the ratio, R_k, of the kth term to the sum of the preceding terms is such that $2^{-p} \leqslant R_k < \varepsilon$, where ε is specified and p is the computer precision in bits (excluding the bits used for the exponent).

For the case of large values of $|z|$, both (1) and (2) converge slowly and are subject to computational round-off errors. This problem is avoided by the use of asymptotic expansions derived from As Eqs. (9.7.1) and (9.7.2):

$$I_n(z) = e^z (2\pi z)^{-1/2} \sum_{k=0}^{\infty} (-1)^k C_k, \qquad n \geqslant 0, |\phi| < \pi/2, \tag{4}$$

where

$$C_k = 1, \qquad k = 0$$
$$= C_{k-1}[\mu - (2k-1)^2]/8kz, \qquad k > 0,$$
$$\mu = 4n^2,$$
$$K_n(z) = e^{-z}(\pi/2z)^{1/2} \sum_{k=0}^{\infty} C_k, \qquad n \geqslant 0, |\phi| < 3\pi/2. \tag{5}$$

For a fixed value of $|z|$ which is large but finite the series in both (4) and (5) diverge, as the ratio of successive terms exceeds unity. This situation often occurs with asymptotic expansions (see Whittaker and Watson [9, p. 150]). By truncating the series at the point where the magnitude of the terms is a minimum, however, an

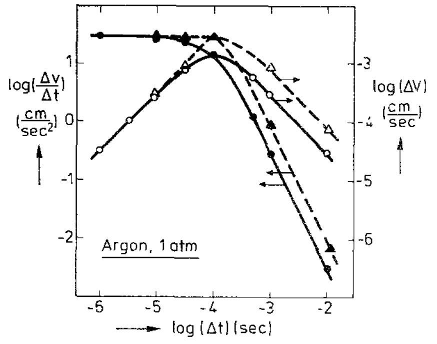

FIG. 1. Gas acceleration (●, ▲) and incremental change of gas velocity (○, △) as functions of time step length. Convection problem. Box side length $L = 5$ cm; $\Delta T = 30$ deg; number of space intervals: 17×17. 3-point-backward time differencing scheme (drawn lines) and trapezoidal time differencing scheme (dashed lines).

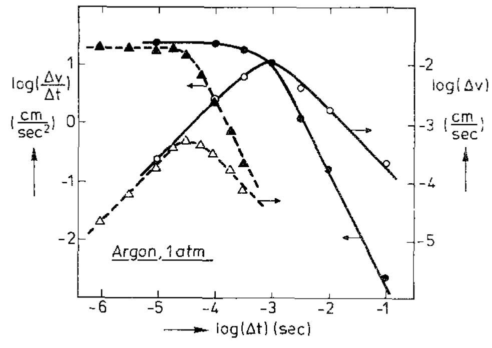

FIG. 2. As Fig. 1. Convection problem. Space intervals: 17×17. $L = 50$ cm, $\Delta T = 200$ deg (drawn lines); $L = 2.5$ cm, $\Delta T = 30$ deg (dashed lines).

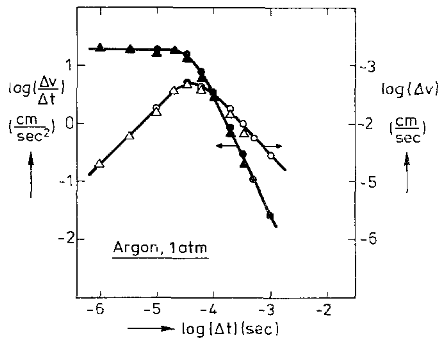

FIG. 3. As Fig. 1. Convection problem $L = 2.5$ cm, $\Delta T = 30$ deg. Space intervals: 9×9 (▲, △) and 17×17 (●, ○).

TABLE II

Normalizing Factors Used in Tabulating Modified Bessel Functions for Arguments of the Form $z = x \exp(i\pi/4) = x(1+i)/\sqrt{2}$

Range of Argument $x = \lvert z \rvert$	I_0	I_1	K_0	K_1
$0.01 - 1$	1	$z/2$	$-yI_0$	$I_0/z + (y-1)I_1$
$1 - 40$	$e^z\left(\frac{1+a}{b}\right)$	$e^z\left(\frac{1-3a}{b}\right)$	$e^{-z}\left(\frac{1-a}{c}\right)$	$e^{-z}\left(\frac{1+3a}{c}\right)$
$y = \ln(z/2) + \gamma;\ a = 1/8z;\ b = (2\pi z)^{1/2};\ c = (2z/\pi)^{1/2}$				

The normalizing factors for $I_n(z)$ are derived by truncation of the converging series after one term in (1) and after two terms in (4). To obtain the normalizing factors for $K_n(z)$ with $|z| \leqslant 1$, (2) is first expressed in the Neumann series form (AS Eq. 9.6.53 and 9.6.54). Truncation after the $k = 0$ terms of those formulations gives the required factors. For $|z| \geqslant 1$, normalizing factors for $K_n(z)$ result from truncation after two terms in (5).

The normalized functional values of $I_n(z)$ and $K_n(z)$ having been tabulated for $n = 0$ and 1, interpolated values can be derived for $0.01 \leqslant |z| \leqslant 40$ as follows. If x_1 and x_N are the lower and upper limits of the range of one of the three sub-tables and N is the number of points in that sub-table, and if the value of x at the interpolated point is such that $x_m \leqslant x < x_{m+1}$, then m is determined from

$$m = \langle (N-1)(x - x_1)/(x_N - x_1) \rangle + 1, \tag{7}$$

TABLE III

Method of Determination of Normalized Value $L(x)$, of Modified Bessel Functions I_0, I_1, K_0 and K_1 for Arguments of the Form $z = x(1+i)/\sqrt{2}$

Classification of Argument	Range of Argument $x = \lvert z \rvert$	Normalized Value $L(x)$
Very small	$x < 0.01$	1
Small	$0.01 \leqslant x \leqslant 1$	Eq. (8)
Intermediate	$1 < x \leqslant 8$	Eq. (8)
Large	$8 < x \leqslant 40$	Eq. (8)
Very large	$40 < x$	1

where $\langle\ \rangle$ denotes integer truncation. Letting $L_m \equiv L(x_m)$ represent the tabulated normalized values of $I_n(z)$ or $K_n(z)$, then the linearly interpolated value is

$$L(x) = L_m + (x - x_m)(L_{m+1} - L_m)/\Delta x. \tag{8}$$

Table III shows the range of x for which (8) is applied.

4. Comparison of Accuracy and Efficiency with Temme's Method

Since the tabulated values are computed only once, it makes sense to perform the calculations with high accuracy, thereby reducing to a minimum the errors arising in the interpolated values. Due to the availability of a CDC CYBER 176 computer with precision $p = 48$ (i.e., approximately 14 decimal digits), it was decided to set $\varepsilon = 10^{-10}$. Interpolated values of $|I_n(z)|$ and $|K_n(z)|$ had relative errors less than 10^{-4} at all values of $z = x(1+i)/\sqrt{2}$. Cyber 176 CPU execution time was 0.75 sec to set up the tables and 0.47 sec for 4,290 evaluations of each of K_0 and K_1 in the range $|z| = 10^{-3}$ to 10^2.

Temme's [8] program was recoded in FORTRAN (with one typographic error corrected) and run with parameter $eps = 10^{-5}$ which gave relative errors for $K_0(z)$ and $K_1(z)$ less than 10^{-4} at all values of $z = x(1+i)/\sqrt{2}$. *CPU* execution time was 1.89 sec for 4,290 evaluations of each of K_0 and K_1 at the same arguments as for the interpolation method. Thus, for comparable accuracy, a simple calculation shows that the interpolation method is faster provided the total number of evaluations exceeds approximately 4,500 which is clearly the case in the applications described in Section 1. It should be noted, furthermore, that the tabular values may be computed and stored, thus eliminating the 0.75 sec start-up time. The interpolation method would then be faster, regardless of the number of evaluations.

Acknowledgments

I am grateful to Jim Salmon for suggesting the method of table-look-up, to Peter Taylor for reviewing the manuscript and to Evonna Mathis for typing it. Two referees also made useful suggestions, most of which have been incorporated.

References

1. P. J. Mason and R. I. Sykes, *Quart. J. Roy. Meteorol. Soc.* **105** (1979), 383–395.
2. P. S. Jackson and J. C. R. Hunt, *Quart. J. Roy. Meteorol. Soc.* **101** (1975), 929–955.
3. M. Abramowitz and I. Stegun, "Handbook of Mathematical Functions," Dover, New York, 1965.
4. J. L. Walmsley, J. R. Salmon, and P. A. Taylor, *Boundary-Layer Meteorol.* **23** (1982), 17–46.
5. J. L. Walmsley and A. D. Howard, submitted to *J. Geophys. Res.*
6. H. A. Scarton, *J. Comput. Phys.* **8** (1971), 295–299.
7. D. J. Sookne, *J. Res. Nat. Bur. Standards B* **77** (1973), 111–114.

8. N. M. TEMME, *J. Comput. Phys.* **19** (1975), 324–337.
9. E. T. WHITTAKER AND G. N. WATSON, "A Course of Modern Analysis," 4th ed., Cambridge Univ. Press, London, 1927.

RECEIVED: April 26, 1983; REVISED: January 20, 1984

JOHN L. WALMSLEY

Boundary Layer Research Division
Atmospheric Environment Service
4905 Dufferin Street
Downsview, Ontario M3H 5T4, Canada