Publ. RIMS, Kyoto Univ.
19 (1983), 1081–1094

A Method for Evaluation of the Error Function of Real and Complex Variable with High Relative Accuracy

By

Masatake MORI*

Abstract

A method for evaluation of the error function of real and complex variable by means of a trapezoidal sum with a correction term is proposed. This method gives a result with high relative accuracy with small number of operations. A precise upper bound of the relative error of the approximation for real variable is given.

§1. A Representation of the Error Function in Terms of a Trapezoidal Sum with a Correction Term

There are a variety of good methods to evaluate the error function

$$\operatorname{erf} t = \frac{2}{\sqrt{\pi}} \int_0^t \exp(-s^2) ds \tag{1.1a}$$

or

$$\operatorname{erfc} t = \frac{2}{\sqrt{\pi}} \int_t^\infty \exp(-s^2) ds \tag{1.1b}$$

for real values of t. Several minimax polynomial or rational approximations can be found in [1, 4]. These approximations, however, are of limited accuracy in the sense that, if we want to improve the accuracy of one of such kind of approximating formulas, we must recalculate anew the coefficients of the formula. In the present paper we propose an analytic formula for evaluating the error function which gives approximate values with very small relative error for real values of t. By means of this method we can evaluate the error function to any

Received December 21, 1982.
* Institute of Information Sciences and Electronics, University of Tsukuba, Sakura-mura, Niihari-gun, Ibaraki-ken, 305, Japan.

number of significant digits provided that the floating point arithmetic system has sufficient digits. This formula can be applied also to the evaluation of the complex error function and gives a value with very small relative error. Salzer [2] has given two efficient methods similar to ours to evaluate complex erfc t.

An expression of the error function

$$\operatorname{erfc} t = \frac{2t}{\pi} \exp(-t^2) \int_0^\infty \frac{e^{-x^2}}{x^2+t^2} dx \tag{1.2}$$

is known in connection with the Mill's ratio ([1], p. 297, also see Appendix A). Hereafter we use this expression. Let $f(t)$ be the integral in the right hand side of (1.2):

$$f(t) = \int_0^\infty \frac{e^{-x^2}}{x^2+t^2} dx = \frac{1}{2} \int_{-\infty}^\infty \frac{e^{-x^2}}{x^2+t^2} dx . \tag{1.3}$$

Then, since this is an integral over $(-\infty, \infty)$ of an analytic function of x which decays very quickly for large value of $|x|$, $f(t)$ can be approximated with high accuracy by a trapezoidal sum $f_h(t)$ with an equal mesh size of h:

$$f_h(t) = h\left[\frac{1}{2t^2} + \sum_{n=1}^\infty \frac{\exp(-n^2h^2)}{(nh)^2+t^2}\right]. \tag{1.4}$$

The error induced when approximating $f(t)$ by $f_h(t)$ is given by [3]

$$\Delta f_h(t) = f(t) - f_h(t) = \frac{1}{4\pi i} \int_C \Phi_h(z) \frac{e^{-z^2}}{z^2+t^2} dz , \tag{1.5}$$

where

$$\Phi_h(z) = \begin{cases} \dfrac{-2\pi i}{1-\exp\left(-\dfrac{2\pi i}{h} z\right)}; & \operatorname{Im} z > 0 \\[2ex] \dfrac{+2\pi i}{1-\exp\left(+\dfrac{2\pi i}{h} z\right)}; & \operatorname{Im} z < 0 . \end{cases} \tag{1.6}$$

The contour C consists of two curves as shown in Fig. 1.

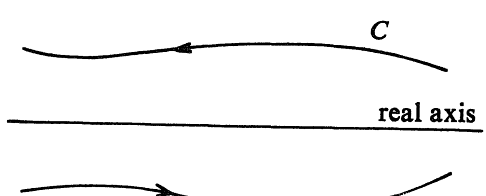

Fig. 1. The path C

We assume for the moment that the variable t is real and positive. If the location of the poles $\pm it$ of the integrand

$$w(z) = \Phi_h(z)\frac{e^{-z^2}}{z^2+t^2} \tag{1.7}$$

of the error integral (1.5) is sufficiently distant from the real axis, then we see that $w(z)$ has two saddle points $\pm s$ which can be approximated by those of

$$\Phi_h(z)\exp(-z^2) \doteqdot \pm 2\pi i \exp\left(\pm\frac{2\pi i}{h}z - z^2\right). \tag{1.8}$$

This gives $s=(\pi/h)i$, so that we have a factor $\exp(-\pi^2/h^2)$ in the error $\Delta f_h(t)$ which becomes very small for small values of h. Therefore, if the distance between the poles $\pm it$ from the real axis is much larger than that of the saddle points $\pm(\pi/h)i$ from the real axis, that is, if

$$t \gg \frac{\pi}{h}, \tag{1.9}$$

then we have an approximate value of $f(t)$ with very high accuracy by the trapezoidal sum (1.4).

On the other hand, if the distance between the poles $\pm it$ of $\exp(-x^2)/(x^2+t^2)$ and the real axis is small, the value of the error integral $\Delta f_h(t)$ becomes large due to the existence of the poles, so that we can not use the right hand side of (1.4) as an approximation to $f(t)$. The integral in the right hand side of (1.5), however, can be expressed as follows if the contour C is modified beyond the

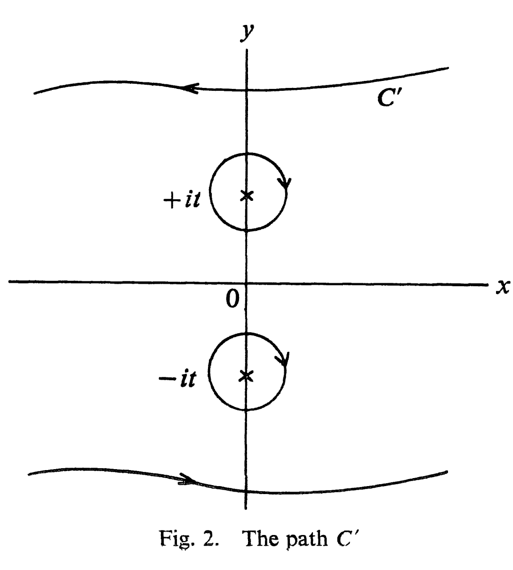

Fig. 2. The path C'

poles $\pm it$ into C' as shown in Fig. 2:

$$\Delta f_h(t) = R(t) + \frac{1}{4\pi i}\int_{C'} \Phi_h(z)\frac{e^{-z^2}}{z^2+t^2}dz. \tag{1.10}$$

$R(t)$ is the term due to the simple poles $\pm it$ of $\exp(-x^2)/(x^2+t^2)$ and can be written explicitly from the residue theorem as

$$R(t) = -2\pi i\left[\frac{1}{4\pi i}\Phi_h(it)\frac{e^{t^2}}{2it} + \frac{1}{4\pi i}\Phi_h(-it)\frac{e^{t^2}}{-2it}\right] \tag{1.11a}$$

$$= -\frac{\dfrac{\pi}{t}e^{t^2}}{\exp\left(\dfrac{2\pi t}{h}\right)-1}, \quad t>0.$$

This equation holds only when $t>0$ because of the definition of (1.6). When $t<0$ we have

$$R(t) = +\frac{\dfrac{\pi}{t}e^{t^2}}{\exp\left(-\dfrac{2\pi t}{h}\right)-1}, \quad t<0. \tag{1.11b}$$

Note that it is easy to evaluate $R(t)$. The contribution from the poles $\pm it$ to the integral along the contour C' is small as is the case stated before, and the value of the integral along C' can be estimated approximately by integrating (1.8) which becomes very small if h is small. Therefore, when h is so small that

$$t \ll \frac{\pi}{h}, \tag{1.12}$$

then, if we add $R(t)$ of (1.11) as a correction term to the trapezoidal sum (1.4), that is, if we compute

$$f_h(t) + R(t), \tag{1.13}$$

we obtain an approximate value of $f(t)$ with a very high accuracy.

When the distance t from the poles $\pm it$ to the real axis is nearly equal to π/h, the behavior of $w(z)$ of (1.7) in the neighborhood of the poles $\pm it$ is a little complicated. In this case, however, the first and the second terms of the right hand side of (1.10), i.e. the magnitude of $R(t)$ and that of the predominating factor $\exp(-\pi^2/h^2)$ of (1.8) are of the same order, so that $R(t)$ may or may not be added to the trapezoidal sum.

We summarize here the formula for computing erfc t with real positive t:

$$
(1.14)\qquad \operatorname{erfc} t=\begin{cases} \dfrac{2th}{\pi}e^{-t^2}\left[\dfrac{1}{2t^2}+\displaystyle\sum_{n=1}^{\infty}\dfrac{\exp(-n^2h^2)}{(nh)^2+t^2}\right] \\ \qquad\qquad +\dfrac{2t}{\pi}e^{-t^2}E(t)\,;\ \dfrac{\pi}{h}\leqq t \\ \dfrac{2th}{\pi}e^{-t^2}\left[\dfrac{1}{2t^2}+\displaystyle\sum_{n=1}^{\infty}\dfrac{\exp(-n^2h^2)}{(nh)^2+t^2}\right] \\ \qquad -\dfrac{2}{\exp\left(\dfrac{2\pi t}{h}\right)-1}+\dfrac{2t}{\pi}e^{-t^2}E(t)\,;\ 0<t<\dfrac{\pi}{h}. \end{cases}
$$

$E(t)$ is the error integral of (1.5) or the second term of the right hand side of (1.10) defined by

$$
(1.15)\qquad E(t)=\frac{1}{4\pi i}\int_{\tilde{C}}\Phi_h(z)\frac{e^{-z^2}}{z^2+t^2}dz,
$$

and the path $\tilde{C}$ is taken in such a way that it is located closer to the real axis than the poles $\pm it$ when the correction term $R(t)$ is not added and that it is located farther from the real axis than the poles when $R(t)$ is added.

When t is small, the denominator of (1.11) should be evaluated by means of the Taylor expansion

$$
(1.16)\qquad \exp\left(\frac{2\pi t}{h}\right)-1=\frac{2\pi t}{h}+\frac{1}{2!}\left(\frac{2\pi t}{h}\right)^2+\cdots
$$

in order to avoid the loss of significant digits. In the actual computation of $f_h(t)$, $b_n=\exp(-n^2h^2)$ can be computed as follows:

$$
(1.17)\qquad \begin{cases} a_0=\exp(h^2), \quad b_0=1, \quad c=1/a_0{}^2 \\ a_{n+1}=a_n\times c, \quad b_{n+1}=b_n\times a_{n+1}, \quad n=0,1,2\ldots \end{cases}
$$

§2. Error Analysis

By choosing the path $\tilde{C}$ of the error integral $E(t)$ appropriately we can obtain an exact upper bound of the error. We assume here again that t is real and

$$
(2.1)\qquad t>0.
$$

For the path $\tilde{C}$ we choose two lines which are parallel to the real axis the distance of which from the real axis is $t+\alpha$ $(-t<\alpha,\ \alpha\neq 0)$ as shown in Fig. 3. Then, since along the path in the upper half plane

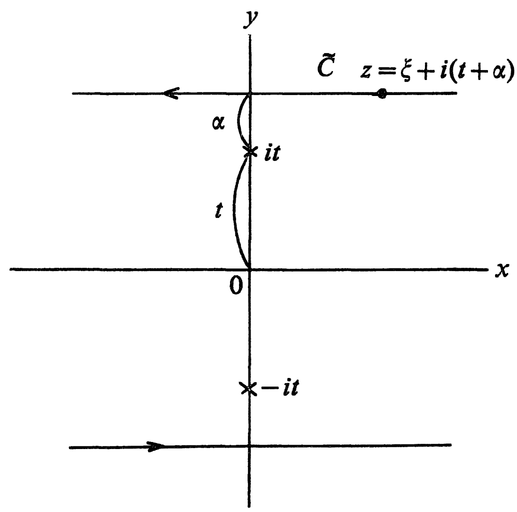

Fig. 3. The path $\tilde{C}$

$$z=\xi+i(t+\alpha), \quad -\infty<\xi<\infty \tag{2.2}$$

we have

$$\begin{aligned}|z^2+t^2|&=|z-it|\cdot|z+it|\\&=|\xi+i\alpha|\cdot|\xi+i(2t+\alpha)|\geqq|\alpha|\cdot(2t+\alpha),\end{aligned} \tag{2.3}$$

so that for the absolute value of the error integral (1.5) or for that of the second term in the right hand side of (1.10) along this path we have

$$\begin{aligned}&\left|\frac{1}{4\pi i}\int_{-\infty}^{\infty}\frac{-2\pi i}{1-\exp\left(\frac{2\pi i}{h}\xi\right)\exp\left(\frac{2\pi}{h}(t+\alpha)\right)}\right.\\&\qquad\qquad\left.\times\frac{\exp\left(-\xi^2+(t+\alpha)^2-2i(t+\alpha)\xi\right)}{(\xi+i(t+\alpha))^2+t^2}d\xi\right|\\&\leqq\frac{\exp(t+\alpha)^2}{2\left\{\exp\left(\frac{2\pi}{h}(t+\alpha)\right)-1\right\}\{|\alpha|(2t+\alpha)\}}\int_{-\infty}^{\infty}e^{-\xi^2}d\xi\\&=\frac{\sqrt{\pi}}{2|\alpha|(2t+\alpha)}\times\frac{\exp\left(t+\alpha-\frac{\pi}{h}\right)^2}{1-\exp\left(-\frac{2\pi}{h}(t+\alpha)\right)}\exp\left(-\frac{\pi^2}{h^2}\right),\end{aligned} \tag{2.4}$$

where h is taken so small that $\exp(2\pi(t+\alpha)/h)>1$ holds. If we take the path in the lower half plane symmetric with respect to that in the upper half plane, we obtain an estimate for the error integral $E(t)$ as follows:

(2.5) $$|E(t)| \leqq \frac{\sqrt{\pi}}{|\alpha|(2t+\alpha)} \times \frac{\exp\left(t+\alpha-\frac{\pi}{h}\right)^2}{1-\exp\left(-\frac{2\pi}{h}(t+\alpha)\right)} \exp\left(-\frac{\pi^2}{h^2}\right).$$

Note that this inequality holds for any α satisfying $\alpha > -t$ $(\alpha \neq 0)$. A choice of such a good α that makes the right hand side of (2.5) small gives a good estimate for the error.

First we assume $t \neq \pi/h$ and take $\alpha = (\pi/h) - t$. Then we have

(2.6) $$|E(t)| \leqq \frac{\sqrt{\pi}}{\left|t^2-\left(\frac{\pi}{h}\right)^2\right|} \times \frac{1}{1-\exp\left(-\frac{2\pi^2}{h^2}\right)} \exp\left(-\frac{\pi^2}{h^2}\right).$$

The factor $\exp(-\pi^2/h^2)$ in the right hand side of (2.6) becomes very small when h is small as already mentioned.

This inequality does not hold when $t=\pi/h$. This corresponds to making the path $\tilde{C}$ pass through the poles $\pm it$. In this case we choose α such that $\alpha = (\pi/h) - t + (1/\sqrt{2})$ so that the path $\tilde{C}$ does not pass through the poles. Then we have

(2.7) $$|E(t)| \leqq \frac{\sqrt{\pi}}{\left|t^2-\left(\frac{\pi}{h}+\frac{1}{\sqrt{2}}\right)^2\right|} \times \frac{\exp\left(\frac{1}{2}\right)}{1-\exp\left(-\frac{2\pi}{h}\left(\frac{\pi}{h}+\frac{1}{\sqrt{2}}\right)\right)} \exp\left(-\frac{\pi^2}{h^2}\right).$$

The inequality (2.6) holds as long as $t \neq \pi/h$, while the inequality (2.7) holds as long as $t \neq (\pi/h)+1/\sqrt{2}$. Therefore the upper bound of the error $|E(t)|$ for the entire range of $0<t$ is given by the smaller one of the right hand side of (2.6) or that of (2.7). The value of t which makes the right hand side of (2.6) equal to that of (2.7) is given by

(2.8) $$t=\beta=\left[\frac{1}{1+\lambda}\left\{\left(\frac{h}{\pi}+\frac{1}{\sqrt{2}}\right)^2+\lambda\left(\frac{\pi}{h}\right)^2\right\}\right]^2$$

where

(2.9) $$\lambda = \frac{1-\exp\left(-\frac{2\pi^2}{h^2}\right)}{1-\exp\left(-\frac{2\pi}{h}\left(\frac{\pi}{h}+\frac{1}{\sqrt{2}}\right)\right)} \exp\left(\frac{1}{2}\right).$$

When $t<\beta$ the right hand side of (2.7) is smaller, while when $\beta<t$ the right hand side of (2.6) is smaller, so that the error induced when erfc t is computed by means of (1.14) can be estimated by (2.7) if $t<\beta$ and by (2.6) if $t>\beta$.

Next we proceed to the estimation of the relative error. Since the relative errors induced when computing erfc t and $f(t)$ are equal, we consider here the relative error

$$(2.10)\qquad e(t)=\frac{E(t)}{f(t)}.$$

As for the magnitude of $f(t)$ itself it is known that

$$(2.11)\qquad \frac{1}{f(t)}<\frac{1}{\sqrt{\pi}}t(t+\sqrt{t^2+2})$$

([1], p. 298, see also Appendix B). Hence we have for the relative error

$$(2.12)\qquad |e(t)|<\frac{1}{\sqrt{\pi}}t(t+\sqrt{t^2+2})|E(t)|.$$

In order to see the behavior of the right hand side of (2.12), we multiply $t(t+\sqrt{t^2+2})/\sqrt{\pi}$ to the right hand side of (2.7) and define a function $\varepsilon_1(t)$ by

$$(2.13)\qquad \varepsilon_1(t)=\frac{t(t+\sqrt{t^2+2})}{\left(\frac{\pi}{h}+\frac{1}{\sqrt{2}}\right)^2-t^2}\frac{\exp\left(\frac{1}{2}\right)}{1-\exp\left(-\frac{2\pi}{h}\left(\frac{\pi}{h}+\frac{1}{\sqrt{2}}\right)\right)}\exp\left(-\frac{\pi^2}{h^2}\right).$$

Then $\varepsilon_1(t)$ is monotone increasing in $0<t<\beta$ and attains its maximum at $t=\beta$. On the other hand, if we multiply $t(t+\sqrt{t^2+2})/\sqrt{\pi}$ to the right hand side of (2.6) and define $\varepsilon_2(t)$ by

$$(2.14)\qquad \varepsilon_2(t)=\frac{t(t+\sqrt{t^2+2})}{t^2-\left(\frac{\pi}{h}\right)^2}\frac{1}{1-\exp\left(-\frac{2\pi^2}{h^2}\right)}\exp\left(-\frac{\pi^2}{h^2}\right),$$

then $\varepsilon_2(t)$ is monotone decreasing in $\beta<t$ and it attains its maximum at $t=\beta$. We note here that

$$(2.15)\qquad \varepsilon_1(t)\sim\frac{\sqrt{2}\exp\left(\frac{1}{2}\right)}{\left(\frac{\pi}{h}+\frac{1}{\sqrt{2}}\right)^2}\frac{\exp\left(-\frac{\pi^2}{h^2}\right)}{1-\exp\left(-\frac{2\pi}{h}\left(\frac{\pi}{h}+\frac{1}{\sqrt{2}}\right)\right)}t\qquad \text{as } t\to 0$$

and that

$$(2.16)\qquad \varepsilon_2(t)\sim\frac{2\exp\left(-\frac{\pi^2}{h^2}\right)}{1-\exp\left(-\frac{2\pi^2}{h^2}\right)}\qquad \text{as } t\to\infty.$$

Here we define $\varepsilon(t)$ by

$$\varepsilon(t)=\begin{cases}\varepsilon_1(t) & 0<t\leqq\beta,\\ \varepsilon_2(t) & \beta<t.\end{cases} \tag{2.17}$$

Then $\varepsilon(t)$, which is a function of h, is an upper bound of the relative error $|e(t)|$. In Fig. 4 $\varepsilon(t)$ for various values of h are shown. From this figure we can choose an appropriate value of the mesh size h for (1.4) corresponding to the required accuracy.

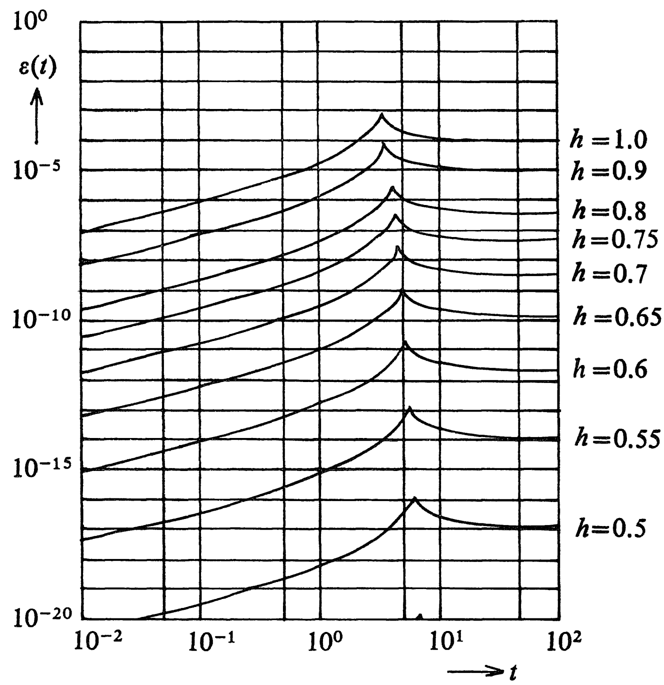

Fig. 4. $\varepsilon(t)$ for various values of the mesh size of h.

When we compute the summation (1.4), we truncate it at a certain value of n, say N, as follows:

$$\frac{1}{h}f_h(t) \doteqdot \frac{1}{2t^2}+\sum_{n=1}^{N}\frac{\exp(-n^2h^2)}{(nh)^2+t^2}. \tag{2.18}$$

Let the required relative accuracy be 10^{-m}. Then N can be determined from

$$\sum_{n=N+1}^{\infty}\frac{\exp(-n^2h^2)}{(nh)^2+t^2}<\frac{1}{2t^2}\times 10^{-m}. \tag{2.19}$$

Since

$$(2.20)\qquad \sum_{n=N+1}^{\infty} \frac{\exp(-n^2h^2)}{(nh)^2+t^2} < \frac{1}{(N+1)^2h^2+t^2} \sum_{n=N+1}^{\infty} \exp(-n^2h^2)$$

$$< \frac{1}{t^2}\int_{Nh}^{\infty} e^{-x^2}dx < \frac{1}{t^2}\int_{Nh}^{\infty} xe^{-x^2}dx < \frac{1}{2t^2}e^{-(Nh)^2},$$

we replace the left hand side of (2.19) by $\exp(-(Nh)^2)/(2t^2)$ and obtain

$$\exp(-(Nh)^2) < \exp(-m \log 10).$$

Therefore, if we choose the smallest value of N satisfying

$$(2.21)\qquad N \geqq \frac{1}{h}\sqrt{2.31m}$$

and compute the sum up to the N-th term, we have a result with relative accuracy less than 10^{-m}. Note that the number N of the terms is very small even when we want to obtain a result with high accuracy. For example, if we want to compute erfc 1 with $m=15$, then from Fig. 4 we should take $h=0.55$ and $N=11$ from (2.21).

An appropriate value of the number N of the terms to be summed up when only the mesh size h is given can be determined approximately from the relation $\exp(N^2h^2) \doteqdot \exp(-\pi^2/h^2)$ which gives.

$$(2.22)\qquad N \doteqdot \frac{\pi}{h^2}.$$

We see that from (2.13) and (2.14) the right hand side of (2.12) is bounded by $\varepsilon_1(\beta)=\varepsilon_2(\beta)$, so that we have an error estimate for the relative error $e(t)$ as follows:

$$(2.23)\qquad |e(t)| < \frac{\beta(\beta+\sqrt{\beta^2+2})}{\left\{\beta^2-\left(\frac{\pi}{h}\right)^2\right\}\left\{1-\exp\left(-\frac{2\pi^2}{h^2}\right)\right\}} \exp\left(-\frac{\pi^2}{h^2}\right).$$

Note that the right hand side of (2.23) gives a uniform error bound from above in the sense that it does not depend on t although it is too much conservative especially for small t.

In Table 1 the relative error $|e(t)|$ of several computed values of erfc t by (1.14) and the corresponding upper bound $\varepsilon(t)$ are shown. Also the numbers N of the terms actually summed up are given. From this table we can see that the upper bound $\varepsilon(t)$ of (2.17) is sufficiently precise.

Table 1. Relative error $|e(t)|$ of the computed values and the theoretical upper bound $\varepsilon(t)$. For $t=0.01$, 0.1, 1.0 and 5.0, $e(t)$ is the relative error of erfc t, while for $t=10.0$ and 100.0, $e(t)$ is that of e^t erfc t. The underline indicates that the correction term $R(t)$ is added. N is the number of the terms actually summed up.

h	t	0.01	0.1	1.0	5.0	10.0	100.0	N
0.5	$e(t)$	1.99×10^{-21}	2.20×10^{-20}	4.72×10^{-19}	2.01×10^{-17}	2.31×10^{-17}	1.44×10^{-17}	13
	$\varepsilon(t)$	3.44×10^{-21}	3.67×10^{-20}	6.74×10^{-19}	2.52×10^{-17}	2.38×10^{-17}	1.44×10^{-17}	—
0.6	$e(t)$	4.90×10^{-16}	5.41×10^{-15}	1.17×10^{-13}	8.13×10^{-12}	3.38×10^{-12}	2.49×10^{-12}	9
	$\varepsilon(t)$	8.25×10^{-16}	8.79×10^{-15}	1.63×10^{-13}	1.01×10^{-11}	3.43×10^{-12}	2.49×10^{-12}	—
0.75	$e(t)$	1.44×10^{-11}	1.59×10^{-10}	3.50×10^{-9}	1.18×10^{-7}	5.78×10^{-8}	4.80×10^{-8}	6
	$\varepsilon(t)$	2.35×10^{-11}	2.50×10^{-10}	4.70×10^{-9}	1.64×10^{-7}	5.85×10^{-8}	4.81×10^{-8}	—
1.0	$e(t)$	5.24×10^{-8}	5.79×10^{-7}	1.31×10^{-5}	1.59×10^{-4}	1.14×10^{-4}	1.04×10^{-4}	4
	$\varepsilon(t)$	8.20×10^{-8}	8.74×10^{-7}	1.69×10^{-5}	1.74×10^{-4}	1.15×10^{-4}	1.04×10^{-4}	—

§3. Computation of erfc t for Complex Values of t

In the discussion above we did not use the fact that t is real except in (2.11). Therefore, in principle, we can apply the method proposed in the preceding sections to the evaluation of erfc t for complex values of t, that is, we can obtain a result with high relative accuracy by just substituting a complex value for t in (1.14). We must note that (1.11a) holds only when Re $t>0$ because of the definition of (1.6). If Re $t<0$ we must use (1.11b).

The correction term $R(t)$ should be added only when its absolute value is larger than that of $E(t)$. As is evident from the error analysis in the previous section, the order of magnitude of the absolute value of the error integral $E(t)$ is approximately equal to $\exp(-\pi^2/h^2)$ as long as the poles $\pm it$ are not extremely close to the saddle points. On the other hand, the order of magnitude of the absolute value of $R(t)$ is approximately equal to

$$(3.1) \qquad \left|\exp\left(t^2-\frac{2\pi t}{h}\right)\right| = \exp\left(\tau^2-\sigma^2-\frac{2\pi\tau}{h}\right), \quad t=\tau+i\sigma .$$

Therefore the domain of t for which $R(t)$ should be added is determined approximately by the relation $\exp(\tau^2-\sigma^2-2\pi\tau/h)>\exp(-\pi^2/h^2)$ when $\tau=\text{Re}\,t>0$, that it

$$(3.2) \qquad 0<\tau<\pm\sigma+\frac{\pi}{h}$$

In other words, if we include the case for Re $t<0$, $R(t)$ should be added if $t=\tau+i\sigma$ falls within the square shown in Fig. 5, and it should not be added if t does not fall within the square.

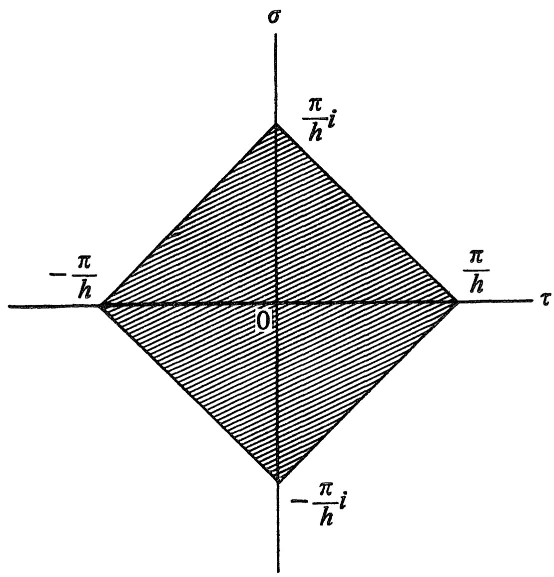

Fig. 5. The domain in which the correction term $R(t)$ should be added.

Inequalities with respect to the magnitude of erfc t itself like (2.11) are not known for complex value of t, and hence we can not obtain an exact upper bound of the relative error when t is complex. However, it is expected that a result with very high accuracy is obtained also when t is complex if we make h sufficiently small and add $R(t)$ if necessary. In Table 2 the absolute value of the relative error $e(t)$ of several computated values for complex t by (1.14) are shown along with the number N of the terms actually summed up.

Table 2. Absolute value of the relative error $e(t)$. For $t=0.01+1.0i$, $0.1+1.0i$, $1.0+1.0i$ and $5.0+1.0i$, $e(t)$ is the relative error of erfc t, while for $t=10.0+1.0i$ and $100.0+1.0i$, $e(t)$ is that of e^t erfc t. The underline indicates that the correction term $R(t)$ is added. N is the number of the terms actually summed up.

h \ t	$0.01+1.0i$	$0.1+1.0i$	$1.0+1.0i$	$5.0+1.0i$	$10.0+1.0i$	$100.0+1.0i$	N
0.5	2.74×10^{-19}	2.96×10^{-19}	7.55×10^{-19}	1.85×10^{-17}	2.26×10^{-17}	1.44×10^{-17}	13
0.6	6.67×10^{-14}	7.21×10^{-14}	1.85×10^{-13}	6.42×10^{-12}	3.34×10^{-12}	2.49×10^{-12}	9
0.75	1.93×10^{-9}	2.09×10^{-9}	5.44×10^{-9}	1.02×10^{-7}	5.75×10^{-8}	4.80×10^{-8}	6
1.0	6.84×10^{-6}	7.39×10^{-6}	1.97×10^{-5}	1.50×10^{-4}	1.14×10^{-4}	1.04×10^{-4}	4

There may be more efficient methods if t varies in a rather restricted range. The merit of the present method is that, not only for real t but also for wide range of complex t, approximate values of erfc t can be obtained with high relative accuracy with very small number of operations.

Appendix A[1]

From the identity

$$\text{(A.1)} \qquad \frac{1}{x^2+t^2}=\int_0^\infty e^{-y(x^2+t^2)}dy\,,$$

we have

$$\text{(A.2)} \qquad \begin{aligned} f(t) &=\int_0^\infty \frac{e^{-x^2}}{x^2+t^2}dx=\int_0^\infty e^{-x^2}dx\int_0^\infty e^{-y(x^2+t^2)}dy \\ &=\int_0^\infty e^{-yt^2}dy\int_0^\infty e^{-(1+y)x^2}dx \\ &=\frac{\sqrt{\pi}}{2}\int_0^\infty \frac{e^{-yt^2}}{\sqrt{1+y}}dy. \end{aligned}$$

A change of variable $y=s^2/t^2-1$ results in

$$\text{(A.3)} \qquad \begin{aligned} f(t) &=\frac{\sqrt{\pi}}{2}\int_t^\infty \frac{t}{s}e^{-(s^2-t^2)}\frac{2s}{t^2}ds \\ &=\frac{\sqrt{\pi}}{t}\,e^{t^2}\int_t^\infty e^{-s^2}ds. \end{aligned}$$

Therefore we have

$$\text{(A.4)} \qquad \operatorname{erfc} t=\frac{2}{\sqrt{\pi}}\int_t^\infty e^{-s^2}ds=\frac{2t}{\pi}\,e^{-t^2}\int_0^\infty \frac{e^{-x^2}}{x^2+t^2}dx.$$

Appendix B[2]

The inequality (2.11) is equivalent to

$$\text{(B.1)} \qquad g(t)=e^{t^2}\int_t^\infty e^{-x^2}dx-\frac{\sqrt{t^2+2}-t}{2}>0,\quad t>0\,,$$

so that we prove (B.1). It is evident from (B.1) that

(1) By Masaaki Sugihara. Private communication.

(2) By Kazuo Murota. Private communication.

(B.2) $$g(0)=\frac{\sqrt{\pi}}{2}-\frac{1}{\sqrt{2}}>0 \quad \text{and} \quad \lim_{t\to\infty} g(t)=0.$$

On the other hand, since

(B.3) $$\begin{aligned} g'(t) &= 2te^{t^2}\int_t^\infty e^{-x^2}dx-1-\frac{1}{2}\left(\frac{t}{\sqrt{t^2+2}}-1\right) \\ &\leqq 2te^{t^2}\int_t^\infty e^{-x^2}dx-1-\left(2t\frac{\sqrt{t^2+2}-t}{2}-1\right)=2tg(t), \end{aligned}$$

we have

(B.4) $$g'(t)\leqq 2tg(t), \quad t>0.$$

Suppose that $g(t)<0$ at $t=t_1>0$. Then we have from (B.4) that $g'(t)<0$ at $t=t_1$. On the other hand $g(t)\to 0$ as $t\to\infty$ and this is a contradiction. Therefore we have

(B.5) $$g(t)\geqq 0.$$

We can remove the equality sign from (B.5) in a similar way, and we conclude $g(t)>0$.

References

[1] Abramowitz, M. and Stegun, A., ed., *Handbook of Mathematical Functions,* NBS Appl. Math. Ser. No. **55** (1964).

[2] Salzer, H. E., Formulas for calculating the error function of a complex variable, *MTAC*, **5** (1951), 67–70.

[3] Takahasi, H. and Mori, M., Error estimation in the numerical integration of analytic functions, *Rep. Comput. Centre Univ. Tokyo,* **3** (1970), 41–108.

[4] Toda, H., The normal distribution (in Japanese), Yamauchi, Z. ed., *Statistical Tables,* Japanese Standards Association (1972), 1–16.