

# Multi-precision computation of the complex error function

Pascal Molin

March 23, 2011

## Abstract

We give a very simple algorithm to compute the error and complementary error functions of complex argument to any given accuracy.

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Motivation, and a smooth integral formula . . . . .	2
<b>2</b>	<b>Computation</b>	<b>3</b>
2.1	Trapezoidal formula . . . . .	3
2.2	Quadrature error . . . . .	4
2.3	Truncation error . . . . .	5
2.4	Proof of theorem 2.3 . . . . .	5
<b>3</b>	<b>Practical algorithm</b>	<b>5</b>
3.1	Extension for small real part . . . . .	7
<b>4</b>	<b>Timings</b>	<b>7</b>

## 1 Introduction

With standard normalisation, the error function is defined as

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

and the complementary error function as

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt.$$

Thanks to the relation  $\operatorname{erf}(x) = 1 - \operatorname{erfc}(x)$  and the symmetry  $\operatorname{erfc}(-x) = 2 - \operatorname{erfc}(x)$ , it is sufficient to be able to compute  $\operatorname{erfc}$  on the right half complex plane.

Many formula exist for this calculation, obtained either via integration or with asymptotic developments near zero or infinity. In this note, we study an integration-based formula which is simple, efficient and easy to apply rigorously.

However this formula is not new (we found in particular [MR71]), it seems that it has not been introduced in numerical systems so far. J. Weideman [Wei94] justifies this by the difficulties to find a correct stepsize, and some numerical instabilities. The rigorous treatment we give here aims at removing these prejudices.

## 1.1 Motivation, and a smooth integral formula

In [Mol10], we showed that the trapezoidal scheme yields a very fast method for the computation of integrals of the form

$$\int_{\mathbb{R}} F(u)e^{-Q(u^2)} du \quad (1)$$

where  $F(x) \in \mathbb{R}(x)$  is a rational fraction with no pole on  $\mathbb{R}$  and  $Q(x) \in \mathbb{R}_+[x]$  is a polynomial.

If we focus on the simplest case

$$f(\alpha) = \int_{\mathbb{R}} \frac{e^{-u^2}}{u - \alpha} du, \alpha \notin \mathbb{R}, \quad (2)$$

we have the following formula, which enables us to compute the complementary error function to any accuracy.

### Proposition 1.1

For all  $x$  such that  $\operatorname{Re}(x) > 0$ ,

$$\operatorname{erfc}(x) = \frac{-ie^{-x^2}}{\pi} f(ix). \quad (3)$$

**Proof:**  $f$  is analytic outside  $\mathbb{R}$ , derivating there under the integral and integrating by parts yields the differential equation

$$f'(\alpha) + 2\alpha f(\alpha) + 2\sqrt{\pi} = 0.$$

Writing  $f(\alpha) = \lambda(\alpha)e^{-\alpha^2}$ ,  $\lambda$  satisfies  $\lambda'(\alpha) = -2\sqrt{\pi}e^{\alpha^2}$ , so that for  $\operatorname{Im}(\alpha) > 0$  we have  $\lambda(\alpha) = K - 2\sqrt{\pi} \int_0^\alpha e^{u^2} du$  with  $K \in \mathbb{C}$ . The limit  $\lim_{x \rightarrow \infty} f(ix)e^{-x^2} = 0$  determines  $K$  to be  $2\sqrt{\pi} \int_0^{i\infty} e^{u^2} du$ . As a consequence,  $\lambda = 2\sqrt{\pi} \int_{ix}^{i\infty} e^{-u^2} du = i\pi \operatorname{erfc} x$ , hence the result.  $\square$

More precisely, the general theory ensures that  $f(\alpha)$  can be evaluated to a precision of  $n$  binary digits with a number of points  $n \sim \frac{n \log 2}{\pi}$ .

The results of [Mol10] also assert that this formula is optimal amongst integration-based formulas, which means that any other integration formula will need a larger number of points (for theoretic reason due to the uncertainty principle).

## 2 Computation

We adopt the following notation for the numerical approximation to  $p$  binary digits of absolute precision:

### Definition 2.1

For  $a, b \in \mathbb{C}$ ,

$$a =_p b \Leftrightarrow |a - b| \leq 2^{-p}. \quad (4)$$

Moreover, the consideration of a residue term leads to the following:

### Definition 2.2

Let  $x$  be a complex number with  $\operatorname{Re}(x) \geq 0$ , and  $h > 0$ , and define  $\delta(x, h)$  to be

$$\delta(x, h) = 1 \text{ if } \operatorname{Re}(x) < \frac{\pi}{h}; \quad (5)$$

$$= 0 \text{ otherwise.} \quad (6)$$

The numerical formula is then

### Theorem 2.3

Let  $x$  be a complex number such that  $\operatorname{Re}(x) \geq 1$ , and suppose  $p \geq 2$ . Then for all  $h > 0$  and  $n \in \mathbb{N}$  such that  $h \leq \frac{\pi}{\sqrt{\operatorname{arcsinh}(2^p \sqrt{\pi}) + 2}}$  and  $nh \geq \sqrt{p \log 2}$  we have

$$\operatorname{erfc}(x) =_p \frac{e^{-x^2}}{\pi} \left( \frac{h}{x} + 2hx \sum_{k=1}^n \frac{e^{-(kh)^2}}{x^2 + (kh)^2} \right) - \frac{2\delta(x+1, h)}{e^{\frac{2\pi x}{h}} - 1}. \quad (7)$$

We give a complete proof in the next paragraphs.

### 2.1 Trapezoidal formula

For  $\alpha \in \mathbb{H}$ , we define  $g(x) = \frac{e^{-x^2}}{x-\alpha}$ . The trapezoidal formula for  $f = \int_{\mathbb{R}} g$  reads

#### Lemma 2.4

Let  $\alpha \in \mathbb{H}$ . For all  $h > 0$  and  $N \geq 1$ , we have

$$f(\alpha) = h \sum_{k=-n}^n \frac{e^{-(kh)^2}}{kh - \alpha} + E_t(n, h) - E_q(h) \quad (8)$$

where  $E_q(h) = \sum_{k \neq 0} \hat{g}(\frac{k}{h})$  and  $E_t(N, h) = \sum_{|k| > N} g(kh)$ .

**Proof:** This is the Poisson formula written for  $g$  on the lattice  $h\mathbb{Z}$ , isolating the first Fourier term  $\hat{g}(0) = f(\alpha)$ .  $\square$

## 2.2 Quadrature error

### Lemma 2.5

Let  $\alpha \in \mathbb{H}$  and  $\tau > 0$  such that  $\tau \neq \text{Im}(\alpha)$ , then we have for all  $X > 0$

$$\left| \hat{g}(-X) + \delta 2i\pi e^{-\alpha^2 - 2i\pi X \alpha} \right| \leq \frac{\sqrt{\pi}}{|\tau - \text{Im}(\alpha)|} e^{\tau^2 - 2\pi X \tau} \quad (9)$$

$$|\hat{g}(X)| \leq \frac{\sqrt{\pi}}{|\tau + \text{Im}(\alpha)|} e^{\tau^2 - 2\pi X \tau} \quad (10)$$

where  $\delta = 0$  if  $\tau < \text{Im}(\alpha)$  and  $\delta = 1$  if  $\tau > \text{Im}(\alpha)$ .

**Proof:** We have  $\hat{g}(\pm X) = \int_{\mathbb{R}} \frac{e^{-u^2 \mp 2\pi u X}}{u - \alpha} du$ , and we shift the contour to  $\mathbb{R} \pm i\tau$  thanks to the residue theorem. The first line corresponds to  $\mathbb{R} \rightarrow \mathbb{R} + i\tau$ , which goes through the pole at  $\alpha$  if  $\tau > \text{Im}(\alpha)$ , and the second line to the shift  $\mathbb{R} \rightarrow \mathbb{R} + i\tau$ . We then bound with the  $L^1$  norm, with  $\int_{\mathbb{R}} \left| e^{-(u \pm i\tau)^2 \mp 2i\pi(u \pm i\tau)X} \right| du = e^{\tau^2 - 2\pi X \tau} \int_{\mathbb{R}} e^{-u^2} du = \sqrt{\pi} e^{\tau^2 - 2\pi X \tau}$ .  $\square$

### Lemma 2.6

Assume  $X_0 \geq \frac{1}{\pi} \sqrt{\text{arcsinh}(2^p \sqrt{\pi})}$  such that  $\pi X_0 \notin ]\text{Im}(\alpha) - 1, \text{Im}(\alpha) + 1[$ , then

$$\forall X \geq X_0, \left| \sum_{k \neq 0} \hat{g}(kX) + \delta(x, 1/X_0) \frac{2i\pi e^{-\alpha^2}}{e^{-2i\pi \alpha X} - 1} \right| \leq 2^{-p}. \quad (11)$$

For all  $X \geq \frac{1}{\pi} \sqrt{\text{arcsinh}(2^p \sqrt{\pi})} + \frac{2}{\pi}$ ,

$$\left| \sum_{k \neq 0} \hat{g}(kX) + \delta(\text{Im}(\alpha) + 1, 1/X) \frac{2i\pi e^{-\alpha^2}}{e^{-2i\pi \alpha X} - 1} \right| \leq 2^{-p}.$$

**Proof:** Fix  $\tau = \pi X_0$  in the preceding lemma; the hypothesis allows to remove the terms  $|\tau \pm \text{Im}(\alpha)|$ . Summing on  $kX, k \neq 0$  we obtain

$$\left| \sum_{k \neq 0} \hat{g}(kX) + \delta \frac{2i\pi e^{-\alpha^2}}{e^{-2i\pi \alpha X} - 1} \right| \leq 2\sqrt{\pi} \sum_{k > 0} e^{-\pi^2 (kX_0)^2} \leq \frac{\sqrt{\pi}}{\sinh(\pi^2 X_0^2)} \quad (12)$$

hence the first assertion of the lemma. Now take  $X$  as in the second part, and set  $X_2 = X - \frac{2}{\pi}, X_1 = X - \frac{1}{\pi}$ . If  $\text{Im}(\alpha) \leq \pi X_1$ , then we use the first part with  $X_0 = X$

and  $\delta(\text{Im}(\alpha), 1/X) = \delta(\text{Im}(\alpha), 1/X_1) = 1$ . If  $\text{Im}(\alpha) > \pi X_1$  we use the first part with  $X \geq X_0 = X_2$  and  $\delta(\text{Im}(\alpha), 1/X_0) = \delta(\text{Im}(\alpha), 1/X_1) = 0$ . Finally, we use  $\delta(x, 1/X_1) = \delta(x+1, 1/X)$ .  $\square$

### 2.3 Truncation error

#### Lemma 2.7

Assume  $\text{Im}(\alpha) > 1$ , and let  $p \geq 2$ . Then for all  $n$  such that  $nh \geq \sqrt{p \log 2}$ , we have

$$\left| h \sum_{|k|>n} \frac{e^{-(kh)^2}}{kh + \alpha} \right| \leq 2^{-p} \quad (13)$$

**Proof:** Since  $\text{Im}(\alpha) \geq 1$  we have  $|(kh)^2 + \alpha| \geq 1$ , and the left-and side is bounded above by  $2h \int_n^\infty e^{-(th)^2} dt = 2 \text{erfc}(nh)$ . The upper bound  $\text{erfc}(x) \leq \frac{2e^{-x^2}}{\sqrt{\pi}(x + \sqrt{x^2 + \frac{4}{\pi^2}})}$  gives the result (we have  $\sqrt{\pi}(x + \sqrt{x^2 + \frac{4}{\pi^2}}) > 4$  for  $x \geq \sqrt{2 \log 2}$ ).  $\square$

### 2.4 Proof of theorem 2.3

We combine the preceding lemmas for  $\alpha = ix$ , writing

$$-ih \sum_{k=-n}^n \frac{e^{-(kh)^2}}{kh + ix} = \frac{h}{x} + 2xh \sum_{k=1}^n \frac{e^{-(kh)^2}}{x^2 + (kh)^2}$$

The value  $\frac{1}{h} = X = \frac{\sqrt{\text{arcsinh}(2^p \sqrt{\pi})}}{\pi} + \frac{2}{\pi}$  ensures the hypothesis of lemma 2.6 is satisfied, so that

$$\left| E_q(h) + \frac{2i\pi\delta(x, h)}{e^{\frac{2\pi x}{h}} - 1} \right| \leq 2^{-p}.$$

The lemma 2.7 bounds  $|E_t(n, h)|$ , and this proves the computation of  $-\frac{i}{\pi}f(ix)$  to absolute precision  $\frac{2}{\pi}2^{-p} < 2^{-p}$  under the hypothesis of theorem 2.3.

The fact that we obtain the right relative precision on  $\text{erfc}(x)$  follows from the bounds

$$\frac{1}{2|x|+1} \leq \left| \frac{\text{erfc}(x)}{e^{-x^2}} \right| \leq \frac{1}{2|x|-1}$$

valid for any  $x \in \mathbb{C}$  with  $|x| \geq 1$ .

## 3 Practical algorithm

The formula of theorem 2.3 is simple. We describe here how to evaluate it efficiently for multiprecision values.

First, we write  $\lambda = \frac{x}{h}$  to put formula (19) into the form

$$\operatorname{erfc}(x) =_p \frac{e^{-x^2}}{\pi} \left( \frac{1}{\lambda} + 2\lambda \sum_{k=1}^n \frac{(e^{-h^2})^{-k^2}}{\lambda^2 + k^2} \right) - \delta \frac{2}{e^{2\pi\lambda} - 1} \quad (14)$$

We then write  $U_k = e^{-(kh)^2}$ ,  $V_k = e^{-(2k+1)h^2}$  and  $D_k = \lambda^2 + k^2$ , so that the main sum becomes

$$\sum_{k=1}^n \frac{U_k}{D_k},$$

subject to the recursions

$$U_{k+1} = U_k V_k \quad (15)$$

$$V_{k+1} = e^{-2h^2} V_k \quad (16)$$

$$D_{k+1} = D_k + 2k + 1. \quad (17)$$

Finally, thanks to the loose condition we have on  $h$ , we can improve the computation if we constraint the factor  $u = e^{-2h^2}$  to be exactly a small precision float, as soon as  $h = \sqrt{-\log \sqrt{u}}$  satisfies theorem 2.3.

This way, the computation of each term of the sum is reduced to the following operations

- a multiprecision multiplication  $U_k V_k$ ;
- a multiplication between a small number  $u$  and multiprecision  $V_k$ ;
- a multiprecision division  $U_k / D_k$ ;
- a small addition

and we have

**Theorem 3.1 (complexity)**

*The complex error function can be evaluated to  $p$  binary digits in complexity  $\frac{p \log 2(1+\lambda)}{\pi} M(p) + o(pM(p))$ , where  $M(p)$  denotes the complexity of a multiplication of size  $p$  and  $\lambda$  is the number of multiplications needed to perform a division.*

**Remark :** Of course this complexity can be lowered for negative real part, since there the result approaches 2 and the precision required to compute the difference can be decreased.

### 3.1 Extension for small real part

The theorem 2.3 can be extended to  $\operatorname{Re}(x) < 1$  if we shift the path of integration. Indeed, the function  $f$  defined for  $\operatorname{Im}(\alpha) > 0$  in (2) extends analytically to  $\operatorname{Im}(\alpha) > -d$  for any  $d > 0$  with

$$f(\alpha) = \int_{\mathbb{R}-id} \frac{e^{-u^2}}{u - \alpha} du. \quad (18)$$

We then have

#### Theorem 3.2

Let  $x$  be a complex number such that  $\operatorname{Re}(x) \geq 0$ . Then for all  $h > 0$  and  $n \in \mathbb{N}$  such that  $h < \frac{\pi}{\sqrt{\operatorname{arcsinh}(2^p \sqrt{\pi})+2}}$  and  $nh \geq \sqrt{(p+1) \log 2}$ , we have

$$\operatorname{erfc}(x) = {}_p \frac{e^{-x^2+1}}{\pi} \left( \frac{1}{\lambda} + 2 \sum_{k=1}^n \frac{(\lambda \cos(2kh) + k \sin(2kh)) e^{-(kh)^2}}{\lambda^2 + k^2} \right) - \frac{2\delta(x+1, h)}{e^{2\pi\lambda} - 1} \quad (19)$$

where  $\lambda = \frac{x+1}{h}$ .

**Proof:** Using (18), the main Riemann sum for  $f(\alpha)$  takes the form

$$\begin{aligned} S(\alpha) &= h \sum_{k=-n}^n \frac{e^{-(kh)^2} e^{2idkh} e^{d^2}}{kh - \tilde{\alpha}}, \quad \tilde{\alpha} = \alpha + id \\ &= -\frac{he^{d^2}}{\tilde{\alpha}} + he^{d^2} \sum_{k=1}^n e^{-(kh)^2} \left( \frac{2ikh \sin(2dkh) + 2\tilde{\alpha} \cos(2dkh)}{(kh)^2 - \tilde{\alpha}^2} \right). \end{aligned}$$

If  $\operatorname{Re}(x+d) \geq 1$ , the truncation error is bounded by  $2 \int_{nh}^{\infty} e^{-t^2+d^2} \leq e^{d^2} E_t(n, h)$ , so that for  $d = 1$  is enough to take  $nh > \sqrt{(p+3) \log 2}$ .

The quadrature error, once corrected by the residue

$$\frac{2\pi\delta(x+d, h)e^{x^2}}{e^{\frac{2\pi(x+d)}{h}} - 1},$$

is bounded by  $\frac{\sqrt{\pi}}{\tau \pm \operatorname{Re}(x+d)} e^{(\tau \pm d)^2 - 2\pi X \tau}$ . For  $X > 0$  the denominator is greater than  $d$  and this can be bound by  $\frac{\sqrt{\pi}}{d} e^{-(\pi X - d)^2}$  so that the value  $X_1$  of lemma 2.6 is enough; for  $X < 0$  we apply the same bounds as in lemma 2.6. This gives the result, fixing the value  $d = 1$ .  $\square$

## 4 Timings

We did a PARI/gp implantation which proves to be quite efficient. In table 1 we compared its running time on a few inputs to MPFR [FHL<sup>+</sup>07] and Maple 14.

digits	value	Maple 14	MPFR 3.0.0	integration
100	3	0.01s	0.012s	0.010s
100	200	0.02s	0.003s	0.008s
100	10 000	0.01s	0.003s	0.008s
100	$\pi + i$	0.03s	*	0.008s
100	$\pi + 1\,000i$	0.01s	*	0.008s
1 000	3	0.00s	0.021s	0.021s
1 000	200	0.03s	0.005s	0.021s
1 000	10 000	0.03s	0.005s	0.028s
1 000	$\pi + i$	0.06s	*	0.035s
1 000	$\pi + 1\,000i$	0.05s	*	0.036s
10 000	3	0.08s	0.246s	3.923s
10 000	200	16.78s	47.840s	3.917s
10 000	10 000	3.48s	0.301s	3.899s
10 000	$\pi + i$	14.26s	*	8.072s
10 000	$\pi + 1\,000i$	16.34s	*	8.059s

Table 1: timings (on a Intel Core2 Quad, 2.40GHz)

It turns out that the integration formula is not very interesting in the asymptotic range (near zero or infinity), where the asymptotic developpements of erf and erfc are very accurate. However, it could be considered in the transition range, when the modulus of the argument comes around the square root of the precision.

We also remark that the integration formula gives a running time which depends only on the required precision, and to a minor extent of the nature (real or complex) of the argument.

## References

- [FHL<sup>+</sup>07] Laurent Fousse, Guillaume Hanrot, Vincent Lefèvre, Patrick Pélissier, and Paul Zimmermann. Mpfr: A multiple-precision binary floating-point library with correct rounding. *ACM Trans. Math. Softw.*, 33, June 2007.
- [Mol10] Pascal Molin. Intégration numérique par la méthode double-exponentielle. <http://hal.archives-ouvertes.fr/hal-00491561/fr/>, jun 2010. 46 pages.
- [MR71] F. Matta and A. Reichel. Uniform computation of the error function and other related functions. *Mathematics of Computation*, 25(114):pp. 339–344, 1971.
- [Wei94] J. A. C. Weideman. Computation of the complex error function. *SIAM Journal on Numerical Analysis*, 31(5):pp. 1497–1518, 1994.



---

**Algorithm 1:** Computation of the complementary error function

---

**Input:**  $x$  such that  $\operatorname{Re}(x) \geq 1$

**Input:** binary precision  $p \geq 1$

**Output:**  $z$  such that  $|(\operatorname{erfc} x - z)/z| < 2^{-p}$

**begin** choose parameters

$$h_0 = \pi / (2 + \sqrt{\operatorname{arcsinh}(2^p \sqrt{\pi})});$$

$$u_0 = e^{-2h_0^2};$$

$$u = \lfloor 2^v u_0 \rfloor;$$

$$n = \lceil p \log 2 / \pi \rceil;$$

**end**

**begin** main multiprecision computation

$$h = \sqrt{-\log(u/2^v)/2};$$

$$\lambda = x/h;$$

$$U = 1; /* U_k = e^{-(kh)^2} */$$

$$V = \sqrt{u/2^v}; /* V_0 = e^{-h^2}, V_k = e^{-(2k+1)h^2} */$$

$$D = \lambda^2; /* D_k = \lambda^2 + k^2 */$$

$$z = 0;$$

**for**  $k=1$  to  $n$  **do**

$$U *= V;$$

$$V *= u/2^v;$$

$$D += 2k + 1;$$

$$z += U/D;$$

**end**

$$z *= 2\lambda;$$

$$z += 1/\lambda; /* term for k = 0 */$$

$$z *= e^{-x^2} / \pi;$$

**end**

**begin** residue term

**if**  $\delta(x, h) = 1$  **then**

$$z -= 2/(e^{2\pi\lambda} - 1);$$

**end**

**end**

**return**  $z$

---