School of Mathematical
and Physical Sciences

University of Reading

Department of Mathematics and Statistics

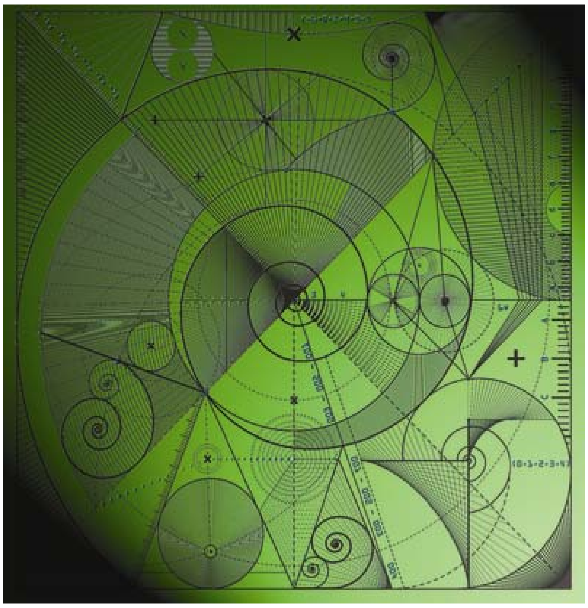
Preprint MPS-2012-20

17 September 2012

Computing Fresnel Integrals via Modified Trapezium Rules

by

Mohammad Alazah, Simon Chandler-Wilde and Scott La Porte



Noname manuscript No.
(will be inserted by the editor)

Computing Fresnel Integrals via Modified Trapezium Rules

Mohammad Alazah ·
Simon N. Chandler-Wilde ·
Scott La Porte

Dedicated to David Hunter on the occasion of his 80th birthday

September 16, 2012

Abstract In this paper we propose methods for computing Fresnel integrals based on truncated trapezium rule approximations to integrals on the real line, these trapezium rules modified to take into account poles of the integrand near the real axis. Our starting point is a method for computation of the error function of complex argument due to Matta and Reichel (*J. Math. Phys.* **34** (1956), 298–307) and Hunter and Regan (*Math. Comp.* **26** (1972), 539–541). We construct approximations which we prove are exponentially convergent as a function of N, the number of quadrature points, obtaining explicit error bounds which show that accuracies of 10^{-15} uniformly on the real line are achieved with $N = 12$, this confirmed by computations. The approximations we obtain are attractive, additionally, in that they maintain small relative errors for small and large argument, are analytic on the real axis (echoing the analyticity of the Fresnel integrals), and are straightforward to implement. In a last section we explore the implications of our results for the computation of the error function of real and complex argument.

Mathematics Subject Classification (2000) 65D30 · 33B32

Mohammad Alazah · Simon N. Chandler-Wilde
Department of Mathematics and Statistics, University of Reading, Whiteknights, PO Box 220, Reading RG6 6AX, UK
E-mail: m.a.m.alazah@pgr.reading.ac.uk
E-mail: S.N.Chandler-Wilde@reading.ac.uk

Scott La Porte
Department of Mathematical Sciences, John Crank Building, Brunel University, Uxbridge UB8 3PH, UK
E-mail: scottis@ntlworld.com

1 Introduction

Let $C(x)$, $S(x)$, and $F(x)$ be the Fresnel integrals defined by

$$C(x) := \int_0^x \cos\left(\tfrac{1}{2}\pi t^2\right)\,\mathrm{d}t, \quad S(x) := \int_0^x \sin\left(\tfrac{1}{2}\pi t^2\right)\,\mathrm{d}t, \tag{1}$$

and

$$F(x) := \frac{\mathrm{e}^{-\mathrm{i}\pi/4}}{\sqrt{\pi}} \int_x^\infty \mathrm{e}^{\mathrm{i}t^2}\,\mathrm{d}t. \tag{2}$$

Our definitions in (1) are those of [2] and [1, §7.2(iii)], and F, C and S are related through

$$\sqrt{2}\,\mathrm{e}^{\mathrm{i}\pi/4}F(x) = \tfrac{1}{2} - C\left(\sqrt{2/\pi}\,x\right) + \mathrm{i}\left(\tfrac{1}{2} - S\left(\sqrt{2/\pi}\,x\right)\right). \tag{3}$$

Another definition in common use [1, §7.2(iii)] is

$$\mathcal{F}(x) := \int_x^\infty \mathrm{e}^{\frac{1}{2}\pi\mathrm{i}t^2}\,\mathrm{d}t = \sqrt{2}\,\mathrm{e}^{\mathrm{i}\pi/4}F\left(\sqrt{\pi/2}\,x\right). \tag{4}$$

Fresnel integrals arise in applications throughout science and engineering, especially in problems of wave diffraction and scattering (*e.g.*, [4, §8.2], [5]), so that methods for the efficient and accurate computation of these functions are of wide application. The purpose of this paper is to present new approximations for the Fresnel integrals, based on N-point trapezium rule approximations to integral representations for $F(x)$, these trapezium rules modified to take into account the presence of poles in the integrand near the path of integration. The resulting approximation to $F(x)$ is

$$F_N(x) := \frac{1}{2} + \frac{\mathrm{i}}{2}\tan\left(A_N x \mathrm{e}^{\mathrm{i}\pi/4}\right) + \frac{x}{A_N}\,\mathrm{e}^{\mathrm{i}(x^2+\pi/4)} \sum_{k=1}^N \frac{\mathrm{e}^{-t_k^2}}{x^2 + \mathrm{i}t_k^2} \tag{5}$$

$$= \frac{1}{\exp\left(2A_N x \mathrm{e}^{-\mathrm{i}\pi/4}\right) + 1} + \frac{x}{A_N}\,\mathrm{e}^{\mathrm{i}(x^2+\pi/4)} \sum_{k=1}^N \frac{\mathrm{e}^{-t_k^2}}{x^2 + \mathrm{i}t_k^2}, \tag{6}$$

where

$$t_k := \frac{(k-1/2)\,\pi}{\sqrt{(N+1/2)\,\pi}}, \quad A_N := t_{N+1} = \sqrt{(N+1/2)\pi}. \tag{7}$$

The corresponding approximations to $C(x)$ and $S(x)$ (obtained by substituting this approximation in (3) and separating real and imaginary parts) are

$$C_N(x) := \frac{1}{2}\,\frac{\sinh\left(\sqrt{\pi}\,A_N\,x\right) + \sin\left(\sqrt{\pi}\,A_N\,x\right)}{\cos(\sqrt{\pi}\,A_N\,x) + \cosh(\sqrt{\pi}\,A_N\,x)}$$
$$+ \frac{\sqrt{\pi}\,x}{A_N}\left(a_N\left(\frac{\pi}{2}x^2\right)\sin\left(\frac{\pi}{2}x^2\right) - b_N\left(\frac{\pi}{2}x^2\right)\cos\left(\frac{\pi}{2}x^2\right)\right) \tag{8}$$

and

$$S_N(x) := \frac{1}{2}\,\frac{\sinh(\sqrt{\pi}\,A_N\,x) - \sin(\sqrt{\pi}\,A_N\,x)}{\cos(\sqrt{\pi}\,A_N\,x) + \cosh(\sqrt{\pi}\,A_N\,x)} - \frac{\sqrt{\pi}\,x}{A_N}\left(a_N\left(\frac{\pi}{2}x^2\right)\cos\left(\frac{\pi}{2}x^2\right) + b_N\left(\frac{\pi}{2}x^2\right)\sin\left(\frac{\pi}{2}x^2\right)\right), \quad (9)$$

where

$$a_N(s) := s\sum_{k=1}^{N}\frac{\mathrm{e}^{-t_k^2}}{s^2+t_k^4}, \quad b_N(s) := \sum_{k=1}^{N}\frac{t_k^2\,\mathrm{e}^{-t_k^2}}{s^2+t_k^4}. \quad (10)$$

These approximations, designed for computation of $F(x)$, $C(x)$ and $S(x)$ for all $x \in \mathbb{R}$, are attractive in several respects. Firstly, they are provably exponentially convergent as N increases: a main result of the paper is to show in §2 that, for $N \in \mathbb{N}$, where

$$E_N(x) := F(x) - F_N(x), \quad (11)$$

we have

$$|E_N(x)| \leq c_N\,\frac{\mathrm{e}^{-\pi N}}{\sqrt{N+1/2}} < \frac{\mathrm{e}^{-\pi N}}{\sqrt{N+1/2}}, \quad x \in \mathbb{R}. \quad (12)$$

Here c_N is a decreasing sequence of positive constants, $c_1 > c_2 > \ldots$, given explicitly by

$$c_N = \frac{20\sqrt{2}\mathrm{e}^{-\pi/2}}{9\pi\left(1-\mathrm{e}^{-2A_N^2}\right)}\left(1 + 2\sqrt{\pi}\,\mathrm{e}^{-\beta A_N^2}\right) + \frac{(2\pi+1)\mathrm{e}^{-\pi/2}}{2\sqrt{2}\,\pi^{3/2}A_N}, \quad (13)$$

where $\beta \approx 0.0536$ is given by (42), so that

$$c_1 \approx 0.825 \quad \text{and} \quad \lim_{N\to\infty} c_N = \frac{20\sqrt{2}\mathrm{e}^{-\pi/2}}{9\pi} \approx 0.208. \quad (14)$$

The bound (12) implies exponential convergence also of the approximations for $C(x)$ and $S(x)$. Indeed, since (3) holds with F, C, and S replaced by F_N, C_N, and S_N, respectively, it holds for $x \in \mathbb{R}$ (see §3) that

$$|C(x) - C_N(x)| \leq \sqrt{2}\,|E_N(\sqrt{\pi/2}\,x)| \quad \text{and} \quad |S(x) - S_N(x)| \leq \sqrt{2}\,|E_N(\sqrt{\pi/2}\,x)|. \quad (15)$$

We will present in §4 numerical computations which demonstrate this exponential convergence and show that, with only $N = 12$ quadrature points, the absolute error $|E_N(x)| \leq 10^{-15}$ and the same bound holds for the relative error, *i.e.*, $|E_N(x)| \leq 10^{-15}|F(x)|$, both these bounds holding for all $x \in \mathbb{R}$. Thus these approximations are efficient; they achieve close to double precision machine accuracy with very little computation.

A second attractive feature of these approximations is their smoothness. The Fresnel integrals $F(z)$, $C(z)$, and $S(z)$, considered as functions of $z \in \mathbb{C}$ in the complex plane, are all entire functions. The approximations $F_N(z)$, $C_N(z)$ and $S_N(z)$ are meromorphic rather than entire functions, but are analytic in a strip which surrounds the real axis, in particular are infinitely differentiable on the real line. But more than this, the bound (12) extends into the complex plane, holding in the first and third quadrants, and, in modified form (see (57)), in the strip $|\mathrm{Im}(z)| < A_N/(2\sqrt{2})$ around the real axis. This implies exponentially convergent error estimates, presented in §2 and §3, for the difference between the coefficients in the Maclaurin series of F, C, and S and those in the corresponding series for F_N, C_N and S_N. In turn (see §3), this implies that the approximations are particularly accurate and retain small relative error for $|x|$ small, and the computations in §4 demonstrate this.

The third attractive feature of these approximations is that they inherit certain symmetries of the Fresnel integrals. In particular, our normalisation of $F(x)$ is such that

$$F(-x) = 1 - F(x), \tag{16}$$

so that, in particular, $F(0) = 1/2$. It is clear from (5) that the same holds for $F_N(x)$, *i.e.*,

$$F_N(-x) = 1 - F_N(x). \tag{17}$$

Similarly, where an overline denotes a complex conjugate,

$$\overline{F(z)} = F(\mathrm{i}\bar{z}) \text{ and } \overline{F_N(z)} = F_N(\mathrm{i}\bar{z}). \tag{18}$$

Both these symmetries can be deduced as a consequence of the structure of C and S and their approximations: by inspection of (8) and (9) we see that

$$C_N(x) = x f_C(x^4), \quad S_N(x) = x^3 f_S(x^4), \tag{19}$$

where f_C and f_S are analytic in a neighbourhood of the real line and are real-valued for real arguments. This is the same structure as C and S (see (67) below). In particular, (19) implies that C_N and S_N, like C and S, are odd functions.

The final attractive feature is that these approximations are straightforward to code. Table 1 shows the simple Matlab code used to evaluate $F_N(x)$ for all the computations in this paper. Of course this code is easily converted to other computer languages.

Let us make some comments regarding the antecedents of our method. The methodology we employ has a long history, though we introduce in this paper significant improvements in implementation and in error analysis. The derivation of our approximation $F_N(x)$ makes use of the relationship between the Fresnel integral and the error function, that

$$F(x) = \tfrac{1}{2}\mathrm{erfc}(\mathrm{e}^{-\mathrm{i}\pi/4}x) = \tfrac{1}{2}\,\mathrm{e}^{\mathrm{i}x^2}\, w\left(\mathrm{e}^{\mathrm{i}\pi/4}x\right) \tag{20}$$

where erfc is the complementary error function, defined by

$$\mathrm{erfc}(z) := \frac{2}{\sqrt{\pi}} \int_z^\infty \mathrm{e}^{-t^2}\mathrm{d}t,$$

and

$$w(z) := \mathrm{e}^{-z^2}\mathrm{erfc}(-\mathrm{i}z).$$

```
function f = fresnel(x,N)
% Evaluates the approximation F_N(x) to the Fresnel integral F(x).
% x is a real scalar or matrix,
% N is the positive integer controlling accuracy (suggest N=12),
% f is the corresponding scalar or matrix of values of F_N(x).
f = zeros(size(x));
select = x>=0;
if any(select), f(select) = F(x(select),N); end
if any(~select), f(~select) = 1-F(-x(~select),N); end

function f = F(x,N)
h = sqrt(pi/(N+0.5));
t = h*((N:-1:1)-0.5);  AN = pi/h;
t2 = t.*t; t4 = t2.*t2; et2 = exp(-t2);
rooti = exp(i*pi/4);
z = rooti*x; x2 = x.*x; x4 = x2.*x2; z2 = i*x2;
S = (-et2(1)./(x4+t4(1))).*(z2+t2(1));
for n = 2:N
    S = S + (-et2(n)./(x4+t4(n))).*(z2+t2(n));
end
ez = exp((2*AN*i*rooti)*x);
f =  (i/AN)*z.*exp(z2).*S + ez./(ez+1);
```

Table 1 Matlab code to evaluate $F_N(x)$ given by (6), making use of (17) for $x < 0$.

It also depends on the integral representation [2, (7.1.4)] that

$$w(z) = \frac{\mathrm{i}}{\pi}\int_{-\infty}^{\infty} \frac{\mathrm{e}^{-t^2}}{z-t}\,\mathrm{d}t = \frac{\mathrm{i}z}{\pi}\int_{-\infty}^{\infty} \frac{\mathrm{e}^{-t^2}}{z^2-t^2}\,\mathrm{d}t, \quad \text{for } \mathrm{Im}(z) > 0. \tag{21}$$

Combining (20) and (21) gives an integral representation for $F(x)$, that

$$F(x) = \frac{x}{2\pi}\,\mathrm{e}^{\mathrm{i}(x^2+\pi/4)} \int_{-\infty}^{\infty} \frac{\mathrm{e}^{-t^2}}{x^2+\mathrm{i}t^2}\,\mathrm{d}t, \quad \text{for } x > 0. \tag{22}$$

The observation that the trapezium rule is exponentially convergent when applied to integrals of the form

$$\int_{-\infty}^{\infty} f(t)\mathrm{e}^{-t^2}\,\mathrm{d}t, \tag{23}$$

with $f(t)$ analytic in a strip surrounding the real axis, dates back at least to Turing [28] and Goodwin [12]. The derivation of this result uses contour integration and Cauchy's residue theorem; see §2 below. Applying the trapezium rule with step-length $h > 0$ to (22) leads to the approximation

$$F(x) \approx \frac{x}{2\pi}\,\mathrm{e}^{\mathrm{i}(x^2+\pi/4)}\,h \sum_{k=-\infty}^{\infty} \frac{\mathrm{e}^{-\tau_k^2}}{x^2+\mathrm{i}\tau_k^2} = \frac{xh}{\pi}\,\mathrm{e}^{\mathrm{i}(x^2+\pi/4)} \sum_{k=1}^{\infty} \frac{\mathrm{e}^{-\tau_k^2}}{x^2+\mathrm{i}\tau_k^2}, \quad \text{for } x > 0, \tag{24}$$

where

$$\tau_k := (k-1/2)h. \tag{25}$$

When $x > 0$ is large this approximation is very accurate, indeed is essentially identical to the approximation $F_N(x)$ with N large if the choice

$$h = \pi/A_N = \sqrt{\pi/(N+1/2)} \tag{26}$$

is made. However, this approximation becomes increasingly poor as $x > 0$ approaches zero.

In the context of developing methods for evaluating the complementary error function of complex argument (by (20), evaluating $F(x)$ for x real is just a special case of this larger problem), Chiarella and Riechel [7], Matta and Reichel [18], and Hunter and Regan [15] proposed modifications of the trapezium rule that follow naturally from the contour integration argument used to prove that the trapezium rule is exponentially convergent. The most appropriate form of this modification is that in [15] where the modified trapezium rule approximation

$$F(x) \approx \frac{xh}{\pi}\,\mathrm{e}^{\mathrm{i}(x^2+\pi/4)} \sum_{k=1}^{\infty} \frac{\mathrm{e}^{-\tau_k^2}}{x^2+\mathrm{i}\tau_k^2} + R(h,x), \quad \text{for } x>0, \tag{27}$$

is proposed. Here the correction term $R(h,x)$ is defined by

$$R(h,x) := \begin{cases} 1/(\exp(2\pi\mathrm{e}^{-\mathrm{i}\pi/4}x/h)+1), & \text{if } 0<x<\sqrt{2}\,\pi/h, \\ 0.5/(\exp(2\pi\mathrm{e}^{-\mathrm{i}\pi/4}x/h)+1), & \text{if } x=\sqrt{2}\,\pi/h, \\ 0, & \text{if } x>\sqrt{2}\,\pi/h. \end{cases}$$

The approximation (27) clearly coincides with $F_N(x)$ for $0 < x < \sqrt{2}\,\pi/h$ if the range of summation in (27) is truncated to $1, ..., N$ and the choice (26) for h is made. Hunter and Regan prove that the magnitude of the error in the approximation (27) is

$$\leq \frac{x\mathrm{e}^{-\pi^2/h^2}}{\sqrt{\pi}\left(1-\mathrm{e}^{-2\pi^2/h^2}\right)\,|x^2/2-\pi^2/h^2|}, \tag{28}$$

for $x > 0$, provided $x \neq \sqrt{2}\pi/h$. Similar estimates, it appears arrived at independently, are derived by Mori [19], in which paper the emphasis is on computing $\mathrm{erfc}(x)$ for real x.

It may be becoming clearer to the reader at this point what the contributions of this paper are, which are to improve and develop the work of Hunter and Regan [15] and the articles that precede it, as they apply to the computation of Fresnel integrals. The new contributions in this paper are essentially four. The first is to truncate the range of summation in (27) and to propose the choice (26) in order to balance the error arising from truncating the infinite sum to the finite range $1, ..., N$ with the error made in the trapezium rule approximation modified in the spirit of [15]. With this change the approximation (27) becomes identical to $F_N(x)$ for $0 < x < \sqrt{2}\,\pi/h$. The second contribution is to recognise that (27), while accurate, has the drawback that $R(x,h)$ is discontinuous, so that the entire function $F(x)$ is being approximated on the real line by a discontinuous function (albeit with small discontinuities). This contribution is further to realise that the approximation formula proposed on $0 < x < \sqrt{2}\,\pi/h$ in fact provides a smooth and accurate approximation on the whole real line. The third contribution, the contribution which is most substantial in terms of analysis, is to improve the error bound (28) of [15]. This error bound is unsatisfactory in that it blows up at $x = \sqrt{2}\,\pi/h$ in a way not seen in the numerical results in §4 (though there is some increase observed in the error when x is near this value; see Figure 3). The bounds we prove in §4 modify the arguments of [18] and [15] to establish rigorous bounds, on both the absolute and relative error in the approximation $F_N(x)$, that are uniform in x on the whole real line. (These rigorous bounds include, in particular, the bound (12).) The final contribution is to explore the implications of these results for the computation of $C(x)$ and $S(x)$, and, briefly, for the computation of $\mathrm{erfc}(x)$ for real and complex argument x.

As mentioned above, it is clear from (20) that numerical methods developed to evaluate $\mathrm{erfc}(z)$ for complex z can be applied in particular to compute $F(x)$ for $x \in \mathbb{R}$. Indeed the starting point for the new method we propose is the method in [15] for $\mathrm{erfc}(z)$. Since [15] was published, further algorithms for computing $\mathrm{erfc}(z)$ with complex argument z have been developed. The algorithm having the best combination of efficiency and accuracy for intermediate values of z (say in the range $1.5 \le |z| \le 5$), where computation is most challenging, appears to be that of Weideman [29]. We will make comparison with this algorithm in §4, showing that our algorithm appears to achieve substantially higher accuracy with fewer computations.

Of course, given the importance of Fresnel integrals in applications, there exist also many effective methods specifically designed to evaluate $C(x)$, $S(x)$ and/or $F(x)$ for real x. We will describe the best of these in §3.1, though we note that none of these approximations appear to have the combination of features of our new approximations (5), (8), and (9), namely that they are analytic in a strip around the real line, are exponentially convergent with provable error bounds, and are straightforward to implement.

We end this introduction by outlining the remainder of the paper. In §2 we derive the approximation (5) to $F(x)$ and prove rigorous bounds on $E_N(x)$, including (12). In §3 we deduce from this the approximations (8) and (9) and bounds on the errors $C(x) - C_N(x)$ and $S(x) - S_N(x)$, especially bounds for x small, and survey other methods for computing Fresnel integrals. In §4 we show numerical results, comparing our new approximations with the error bounds derived in the earlier sections and with certain rival methods for computing Fresnel integrals. In concluding remarks in §5 we explain how the methods we introduce in §2 are potentially of much wider application to computation of special functions, not least $\mathrm{erfc}(z)$ for complex z and other functions arising in computational acoustics. The appendix proves a sharp lower bound for $|F(x)|$ for $x > 0$, of independent interest and a key component in our theoretical bounds on relative errors. Maintaining a theme, the proof of this lower bound requires modified trazezium rule approximations and their error estimates.

2 The approximation of $F(x)$ and its error bounds

In this section we derive the approximation (5) to $F(x)$ and derive (12) and related error bounds for this approximation, these error bounds demonstrating that both the absolute and relative errors in the approximation $F_N(x)$ converge exponentially to zero as N increases, uniformly on the real line, and that $N = 12$ is large enough to achieve errors $< 10^{-15}$.

The first part of our derivation follows in large part Matta and Reichel [18] and Hunter and Regan [15]. From (22) we have that, for $x > 0$,

$$I := \int_{-\infty}^{\infty} f(t)\,\mathrm{d}t = F(x), \text{ where } f(t) := \mathrm{e}^{\mathrm{i}(x^2+\pi/4)}\,\frac{x}{2\pi}\,\frac{\mathrm{e}^{-t^2}}{x^2+\mathrm{i}t^2}, \tag{29}$$

and we have suppressed in our notation the dependence of $f(t)$ on x. Choose a step-size $h > 0$ for the trapezium rule and let

$$g(z) = \mathrm{i}\tan(\pi z/h),$$

which is a meromorphic function with simple poles at the points τ_k, defined by (25), which has the property that, for $z = X + \mathrm{i}H$ with $X \in \mathbb{R}$, $H > 0$,

$$|1 + g(z)| \le \frac{2\mathrm{e}^{-2\pi H/h}}{1 - \mathrm{e}^{-2\pi H/h}}. \tag{30}$$

The approximation (27) is obtained by considering the integral in the complex plane,

$$J = \int_{\Gamma} f(z)(1 + g(z))\,\mathrm{d}z,$$

where the path of integration is from $-\infty$ to ∞ along the real axis, except that the path makes small semicircular deformations to pass above each of the simple poles at the points τ_k, $k \in \mathbb{Z}$. Explicitly, the kth deformation is the semicircle $\gamma_k = \{\tau_k + \epsilon\mathrm{e}^{-\mathrm{i}\theta} : \pi \le \theta \le 2\pi\}$, with ϵ in the range $(0, h/2)$ small enough so that the simple pole singularity in $f(z)$ at $z = z_0 := \mathrm{e}^{\mathrm{i}\pi/4}x$ lies above Γ. Then, since $f(z)g(z)$ is an odd function, we see that

$$J = \int_{\Gamma} f(z)\,\mathrm{d}z + \int_{\Gamma} f(z)g(z)\,\mathrm{d}z = I + \sum_{k\in\mathbb{Z}} \int_{\gamma_k} f(z)g(z)\,\mathrm{d}z.$$

In the limit $\epsilon \to 0$, $\int_{\gamma_k} f(z)g(z)\,\mathrm{d}z \to -\pi\mathrm{i}\,\mathrm{Res}(fg, \tau_k) = -hf(\tau_k)$, where $\mathrm{Res}(g, \tau_k)$ denotes the residue of fg at τ_k. Thus $J = I - I_h$, where

$$I_h = h\sum_{k\in\mathbb{Z}} f(\tau_k) = 2h\sum_{k=1}^{\infty} f((k-1/2)h) \tag{31}$$

is a trapezium/midpoint rule approximation to I. On the other hand, where $\Gamma_H = \{x + \mathrm{i}H : x \in \mathbb{R}\}$, by the residue theorem,

$$J = \int_{\Gamma_H} f(z)(1 + g(z))\,\mathrm{d}z + \mathbf{H}\left(\sqrt{2}\,H - x\right)\, PC_h,$$

for $H > 0$, where $\mathbf{H}$ is the Heaviside step function and

$$PC_h = 2\pi\mathrm{i}\,\mathrm{Res}(f(1+g), z_0) = \frac{1}{2}\,(1 + g(z_0)) = \frac{1}{2}\left(1 + \mathrm{i}\tan\left(\mathrm{e}^{\mathrm{i}\pi/4} x\pi/h\right)\right).$$

Thus

$$I = I_h + \mathbf{H}\left(\sqrt{2}\,H - x\right)\,PC_h + \int_{\Gamma_H} f(z)(1 + g(z))\,\mathrm{d}z. \tag{32}$$

The point here is that the integral over Γ_H can be negligible so that a good approximation is obtained by the modified trapezium rule approximation, $I_h + \mathbf{H}\left(\sqrt{2}\,H - x\right)\,PC_h$. In particular, noting (30) and that, for $z = X + \mathrm{i}H$,

$$|x^2 + \mathrm{i}z^2| = |z_0 - z|\,|z_0 + z| \geq |x/\sqrt{2} - H|\,|x/\sqrt{2} + H| = |x^2/2 - H^2|$$

and that $\int_{-\infty}^{\infty} \mathrm{e}^{-t^2}\,dt = \sqrt{\pi}$, we see that

$$\left|\int_{\Gamma_H} f(z)(1 + g(z))\,\mathrm{d}z\right| \leq \frac{x\,\mathrm{e}^{H^2 - 2\pi H/h}}{\sqrt{\pi}\,|H^2 - x^2/2|\,\left(1 - \mathrm{e}^{-2\pi H/h}\right)}.$$

Choosing $H = \pi/h$, to minimise the exponent $H^2 - 2\pi H/h$, it follows that $I = I_h + \mathbf{H}\left(\sqrt{2}\,\pi/h - x\right)\,PC_h + e_h$ with

$$|e_h| \leq \delta_1(x) := \frac{x\,\,\mathrm{e}^{-\pi^2/h^2}}{\sqrt{\pi}\,|\pi^2/h^2 - x^2/2|\,\left(1 - \mathrm{e}^{-2\pi^2/h^2}\right)}. \tag{33}$$

Note that $I_h + \mathbf{H}\left(\sqrt{2}\,\pi/h - x\right)\,PC_h = I_h + R(h,x)$ is precisely the approximation (27), and that the above bound on e_h is precisely the bound (28) from [15].

Let $I_h^* := I_h + PC_h$ and $e_h^* := I - I_h^*$. Then (33) implies that

$$|e_h^*| \leq \delta_1(x), \;\text{ for } 0 < x < \sqrt{2}\,\pi/h. \tag{34}$$

Since, applying (30),

$$|PC_h| \leq \frac{\mathrm{e}^{-\sqrt{2}\,\pi x/h}}{1 - \mathrm{e}^{-\sqrt{2}\,\pi x/h}},$$

we see that

$$|e_h^*| \leq \delta_3(x) := \delta_1(x) + \frac{\mathrm{e}^{-\sqrt{2}\,\pi x/h}}{1 - \mathrm{e}^{-\sqrt{2}\,\pi x/h}}, \;\text{ for } x > \sqrt{2}\,\pi/h. \tag{35}$$

The bounds (34) and (35) both blow up as x approaches $\sqrt{2}\,\pi/h$. Continuing to choose $H = \pi/h$, select ϵ in the range $(0, H)$ and consider the case that

$$\left|\frac{x}{\sqrt{2}} - H\right| < \epsilon. \tag{36}$$

In this case we observe that the derivation of (32) can be modified to show that

$$e_h^* = \int_{\Gamma_H^*} f(z)(1 + g(z))\,\mathrm{d}z \tag{37}$$

where the contour Γ_H^* passes above the pole in f at z_0; precisely, Γ_H^* is the union of Γ' and γ, where $\Gamma' = \{z \in \Gamma_H : |z - z_0| > \epsilon\}$ and γ is the circular arc $\gamma = \{z_0 + \epsilon e^{i\theta} : \theta_0 \leq \theta \leq \pi - \theta_0\}$, where $\theta_0 = \sin^{-1}((H - x/\sqrt{2})/\epsilon) \in (-\pi/2, \pi/2)$. For $z \in \Gamma'$ it holds that

$$|x^2 + iz^2| = |z_0 - z|\,|z_0 + z| \geq \epsilon\,|x/\sqrt{2} + H|. \tag{38}$$

Thus, and applying (30), similarly to (34) we deduce that

$$\left|\int_{\Gamma'} f(z)(1 + g(z))\,dz\right| \leq \frac{x\ e^{-\pi^2/h^2}}{\sqrt{\pi}\,\epsilon|\pi/h + x/\sqrt{2}|\ \left(1 - e^{-2\pi^2/h^2}\right)}. \tag{39}$$

To bound the integral over γ we note that, for $z = X + iY = z_0 + \epsilon e^{i\theta} \in \gamma$, (38) is true and $Y \geq H$. Further, $|e^{-z^2}| = e^P$, where

$$P = Y^2 - X^2 = 2x\epsilon \sin(\theta - \pi/4) - \epsilon^2 \cos(2\theta) < 2x\epsilon + \epsilon^2 \leq 2\sqrt{2}H\epsilon + (2\sqrt{2}+1)\epsilon^2,$$

using (36). From these bounds and (30), defining $\alpha = \epsilon/H \in (0, 1)$, we deduce that

$$\left|\int_{\gamma} f(z)(1 + g(z))\,dz\right| \leq \frac{2x\ \exp((2\sqrt{2}\alpha + (2\sqrt{2}+1)\alpha^2 - 2)\pi^2/h^2)}{\epsilon|\pi/h + x/\sqrt{2}|\ \left(1 - e^{-2\pi^2/h^2}\right)}. \tag{40}$$

Choosing $\alpha = 1/4$, we can bound e_h^* using (37), (39), (40), and the triangle inequality, to get that

$$|e_h^*| \leq \delta_2(x) := \frac{4hx\ e^{-\pi^2/h^2}}{\pi^{3/2}|\pi/h + x/\sqrt{2}|\ \left(1 - e^{-2\pi^2/h^2}\right)}\left(1 + 2\sqrt{\pi}\,e^{-\beta\pi^2/h^2}\right), \tag{41}$$

for x in the range (36), where

$$\beta = 1 - 2\sqrt{2}\alpha - (2\sqrt{2}+1)\alpha^2 = 1 - \frac{\sqrt{2}}{2} - \frac{2\sqrt{2}+1}{16} \approx 0.0536. \tag{42}$$

Summarising up to this point, we see that we have shown that $I = F(x)$ can be approximated by the modified trapezium rule $I_h^* = I_h + PC_h$ with error $e_h^* = I - I_h^*$. This error satisfies $|e_h^*| \leq \Delta_h(x)$, for $x > 0$, where

$$\Delta_h(x) := \begin{cases} \delta_1(x), & 0 \leq \frac{x}{\sqrt{2}} \leq \frac{3}{4}H, \\ \delta_2(x), & \frac{3}{4}H < \frac{x}{\sqrt{2}} < \frac{5}{4}H, \\ \delta_3(x), & \frac{x}{\sqrt{2}} \geq \frac{5}{4}H, \end{cases} \tag{43}$$

with $H = \pi/h$. Here δ_1, δ_2, and δ_3 are defined by (33), (41), and (35), respectively.

The approximation $F_N(x)$, given by (5), that we propose for $I = F(x)$ is just $I_h^* = I_h + PC_h$ with a particular choice of h and with the range of summation in (31) reduced to the finite range $1, ..., N$. This induces an additional error,

$$T_N := 2h \sum_{m=N+1}^{\infty} f(\tau_m) \tag{44}$$

which, for $x > 0$, is bounded by

$$
\begin{aligned}
|T_N| &\le \frac{hx}{\pi} \sum_{m=N+1}^{\infty} \frac{\mathrm{e}^{-\tau_m^2}}{\sqrt{x^4+\tau_m^4}} \\
&\le \frac{x}{2\pi\sqrt{x^4+\tau_{N+1}^4}} \left(2h\mathrm{e}^{-\tau_{N+1}^2} + 2h \sum_{m=N+2}^{\infty} \mathrm{e}^{-\tau_m^2} \right) \\
&\le \frac{x}{2\pi\sqrt{x^4+\tau_{N+1}^4}} \left(2h\mathrm{e}^{-\tau_{N+1}^2} + 2 \int_{\tau_{N+1}}^{\infty} \mathrm{e}^{-t^2}\mathrm{d}t \right) \\
&\le \frac{x}{2\pi\sqrt{x^4+\tau_{N+1}^4}} \left(2h\mathrm{e}^{-\tau_{N+1}^2} + \frac{\mathrm{e}^{-\tau_{N+1}^2}}{\tau_{N+1}} \right) = \frac{(2h\tau_{N+1}+1)x}{2\pi\tau_{N+1}\sqrt{x^4+\tau_{N+1}^4}}\, \mathrm{e}^{-\tau_{N+1}^2}.
\end{aligned}
$$

To arrive at the last line we have used that, for $x > 0$,

$$
2\int_x^{\infty} \mathrm{e}^{-t^2}\mathrm{d}t = \frac{\mathrm{e}^{-x^2}}{x} - \int_x^{\infty} \frac{\mathrm{e}^{-t^2}}{t^2}\mathrm{d}t < \frac{\mathrm{e}^{-x^2}}{x}. \tag{45}
$$

The choice of h we make is designed to approximately equalise $\Delta_h(x)$ and this bound on T_N. We choose h so that $H = \pi/h = \tau_{N+1} = (N+1/2)h$, *i.e.*, we make the choice $h = \sqrt{\pi/(N+1/2)}$ given by (26), in which case $\tau_{N+1} = A_N = \sqrt{(N+1/2)\pi}$, and $\tau_k = t_k$, where t_k is defined by (7). With this choice of h it holds that

$$
E_N(x) = F(x) - F_N(x) = e_h^* + T_N
$$

and that

$$
|T_N| \le \frac{(2\pi+1)x}{2\pi A_N\sqrt{x^4+A_N^4}}\, \mathrm{e}^{-A_N^2}.
$$

Thus we arrive at our main pointwise error bound, that

$$
|E_N(x)| \le \eta_N(x) := \Delta_h(|x|) + \frac{(2\pi+1)|x|}{2\pi A_N\sqrt{x^4+A_N^4}}\, \mathrm{e}^{-A_N^2}, \tag{46}
$$

with $h = \sqrt{\pi/(N+1/2)}$ so that $H = \pi/h = A_N$. We have shown this bound for $x > 0$, but the symmetries (16) and (17) imply that $E_N(-x) = -E_N(x)$, so that (46) holds also for $x < 0$, and, by continuity, also for $x = 0$ (and in fact $E_N(0) = \eta_N(0) = 0$). Explicitly, for this choice of h we have that

$$
\Delta_h(x) = \begin{cases}
\dfrac{x\,\mathrm{e}^{-A_N^2}}{\sqrt{\pi}\,(A_N^2 - x^2/2)\,\left(1-\mathrm{e}^{-2A_N^2}\right)}, & 0 \le \dfrac{x}{\sqrt{2}} \le \frac{3}{4}A_N, \\
\dfrac{4x\,\mathrm{e}^{-A_N^2}\left(1+2\sqrt{\pi}\,\mathrm{e}^{-\beta A_N^2}\right)}{\sqrt{\pi}\,A_N(A_N + x/\sqrt{2})\,\left(1-\mathrm{e}^{-2A_N^2}\right)}, & \frac{3}{4}A_N < \dfrac{x}{\sqrt{2}} < \frac{5}{4}A_N, \\
\dfrac{x\,\mathrm{e}^{-A_N^2}}{\sqrt{\pi}\,(x^2/2 - A_N^2)\,\left(1-\mathrm{e}^{-2A_N^2}\right)} + \dfrac{\mathrm{e}^{-\sqrt{2}\,A_N x}}{1-\mathrm{e}^{-\sqrt{2}A_N x}}, & \dfrac{x}{\sqrt{2}} \ge \frac{5}{4}A_N.
\end{cases} \tag{47}
$$

We will compare $|E_9(x)|$ to the upper bound $\eta_9(x)$ in Figure 2 below.

Note the factor $\exp(-A_N^2) = \exp(-\pi(N+1/2))$ in each of the terms on the right hand side of (47), so that (46) implies that $F_N(x)$ is exponentially convergent as $N \to \infty$ for each x. Note also that $\Delta_h(x)$ is increasing on $[0, \frac{5}{4}\sqrt{2}\,A_N)$ and decreasing on $[\frac{5}{4}\sqrt{2}\,A_N, \infty)$, behaving asymptotically like $k_N x^{-1}$ as $x \to \infty$, where $k_N = 2\mathrm{e}^{-A_N^2}/(\sqrt{\pi}\,(1-\mathrm{e}^{-2A_N^2}))$. Further, where $\Delta_h(\frac{5}{4}\sqrt{2}\,A_N^-)$ denotes the limiting value of $\Delta_h(x)$ as $x \to \frac{5}{4}\sqrt{2}\,A_N$ from below, since $2A_N^{-1} > \mathrm{e}^{-A_N^2}$,

$$\begin{aligned}\Delta_h\left(\tfrac{5}{4}\sqrt{2}A_N^-\right) &= \frac{20\sqrt{2}\,\mathrm{e}^{-A_N^2}}{9\sqrt{\pi}\,A_N\,\left(1-\mathrm{e}^{-2A_N^2}\right)}\left(1+2\sqrt{\pi}\,\mathrm{e}^{-\beta A_N^2}\right)\\ &> \frac{20\sqrt{2}\,\mathrm{e}^{-A_N^2}}{9\sqrt{\pi}\,A_N\,\left(1-\mathrm{e}^{-2A_N^2}\right)} + \frac{\mathrm{e}^{-5A_N^2/2}}{1-\mathrm{e}^{-5A_N^2/2}} = \Delta_h\left(\tfrac{5}{4}\sqrt{2}\,A_N\right).\end{aligned}$$

Similarly, $x\Delta_h(x)$ is increasing on $[0, \frac{5}{4}\sqrt{2}\,A_N)$ and decreasing on $[\frac{5}{4}\sqrt{2}\,A_N, \infty)$, approaching the limit k_N as $x \to \infty$. Thus

$$\Delta_h(|x|) \le \Delta_h\left(\tfrac{5}{4}\sqrt{2}\,A_N^-\right) \quad \text{and} \quad x\Delta_h(|x|) \le \tfrac{5}{4}\sqrt{2}\,A_N \Delta_h\left(\tfrac{5}{4}\sqrt{2}\,A_N^-\right), \quad \text{for } x \in \mathbb{R}. \tag{48}$$

Moreover,

$$\frac{|x|}{\sqrt{x^4+A_N^4}} \le \frac{1}{\sqrt{2}\,A_N} \quad \text{and} \quad \frac{x^2}{\sqrt{x^4+A_N^4}} < 1, \quad \text{for } x \in \mathbb{R}. \tag{49}$$

Combining (46), (48), and (49), we see that

$$|F(x) - F_N(x)| = |E_N(x)| \le \eta_N(x) \le c_N \frac{\mathrm{e}^{-\pi N}}{\sqrt{N+1/2}}, \quad \text{for } x \in \mathbb{R}, \tag{50}$$

where

$$c_N = \mathrm{e}^{\pi N}\sqrt{N+1/2}\left[\Delta_h\left(\tfrac{5}{4}\sqrt{2}\,A_N^-\right) + \frac{(2\pi+1)}{2\sqrt{2}\,\pi}\,\frac{\mathrm{e}^{-A_N^2}}{A_N^2}\right]$$

is given by (13). Note that c_N decreases as N increases, with $c_1 \approx 0.8249$ and $\lim_{N\to\infty} c_N \approx 0.2080$ given by (14). Of course, as observed in the introduction, this simple, explicit bound shows exponential convergence of the approximation F_N to F, uniformly on the real line.

In the appendix it is shown that $|F(x)| \ge (2+2\sqrt{\pi}x)^{-1}$ for $x \ge 0$, and that $|F(x)| \ge 1/2$ for $x \le 0$. Combining these bounds with (46), (48), and (49), we see that

$$\frac{|F(x)-F_N(x)|}{|F(x)|} \le \frac{\eta_N(x)}{|F(x)|} \le \begin{cases} c_N^*\mathrm{e}^{-\pi N}, & \text{for } x \ge 0,\\ 2c_N\dfrac{\mathrm{e}^{-\pi N}}{\sqrt{N+1/2}}, & \text{for } x \le 0,\end{cases} \tag{51}$$

where

$$c_N^* = 2\mathrm{e}^{\pi N}\left[\left(1+\tfrac{5}{4}\sqrt{2\pi}A_N\right)\mathit{\Delta}_h\left(\tfrac{5}{4}\sqrt{2}\,A_N^-\right)+\frac{(2\pi+1)}{2\pi}\,\frac{\mathrm{e}^{-A_N^2}}{A_N}\left(\frac{1}{\sqrt{2}\,A_N}+\sqrt{\pi}\right)\right]$$
$$= \frac{10\sqrt{2}\left(4+5\sqrt{2\pi}A_N\right)\left(1+2\sqrt{\pi}\mathrm{e}^{-\beta A_N^2}\right)}{9\sqrt{\pi}\,\mathrm{e}^{\pi/2}\,A_N\left(1-\mathrm{e}^{-2A_N^2}\right)}+\frac{(2\pi+1)}{\pi\mathrm{e}^{\pi/2}A_N}\left(\frac{1}{\sqrt{2}\,A_N}+\sqrt{\pi}\right).$$

Note that c_N^* decreases as N increases, with $c_1^* \approx 10.4$ and $\lim_{N\to\infty} c_N^* = 100\mathrm{e}^{-\pi/2}/9 \approx 2.3$. The bound (51) shows exponential convergence of the relative error, $|F_N(x) - F(x)|/|F(x)|$, uniformly on the real line, in particular showing that the relative error is $\leq 1.6 \times 10^{-16}$ on the whole real line if $N = 12$ (see Figure 1 below).

The above estimates use (46) and (47) to bound the maximum absolute and relative errors in the approximation $F_N(x)$. Let us note that these inequalities, additionally, imply that $F_N(x)$ is particularly accurate for $|x|$ small. For $|x| \leq A_N/\sqrt{2} = \sqrt{(N+1/2)\pi/2}$, it follows from (46) and (47) that

$$|F(x) - F_N(x)| \leq \eta(x) \leq \tilde{c}_N|x|\frac{\mathrm{e}^{-\pi N}}{2N+1} \tag{52}$$

where

$$\tilde{c}_N = \frac{8}{3\pi^{3/2}\mathrm{e}^{\pi/2}\left(1-\mathrm{e}^{-2A_N^2}\right)}+\frac{(2\pi+1)}{\pi^2\mathrm{e}^{\pi/2}A_N}. \tag{53}$$

Note that $\tilde{c}_N$ decreases as N increases, with $\tilde{c}_1 \approx 0.17$ and $\lim_{N\to\infty}\tilde{c}_N = 8/(3\pi^{3/2}\mathrm{e}^{\pi/2}) \approx 0.10$.

In §1 we have made claims regarding the analyticity of the approximation $F_N(x)$, considered as a function of x in the complex plane. We justify these claims now. One attractive feature of the modified trapezium rule approximation I_h^* is that, in contrast to I_h, it is entire as a function of x. This is not immediately obvious: $I_h^* = I_h + PC_h$, and PC_h has simple pole singularities at $x = \mathrm{e}^{-\mathrm{i}\pi/4}\tau_k$, $k \in \mathbb{Z}$. But I_h also has simple poles at the same points and it is an easy calculation to see that the residues add to zero, so that the singularities cancel out. Since $F_N(x) = I_h^* - T_N$, with h given by (26), it follows that the singularities of $F_N(x)$ are those of T_N, *i.e.*, simple poles at $\pm\mathrm{e}^{-\mathrm{i}\pi/4}t_k$, for $k = N+1, N+2, \ldots$. Thus $F_N(x)$ is a meromorphic function and, in particular, is analytic in the strip $|\mathrm{Im}(x)| < A_N/\sqrt{2}$ and in the first and third quadrants of the complex plane.

We will note two consequences of this analyticity and the bounds that we have already proved. In these arguments we will use an extension of the maximum principle for analytic functions to unbounded domains, that if $w(z)$ is analytic in an open quadrant in the complex plane, let us say $Q = \{z \in \mathbb{C} : 0 < |\arg(z)| < \pi/2\}$, and is continuous and bounded in its closure, then

$$\sup_{z\in Q}|w(z)| \leq \sup_{z\in\partial Q}|w(z)|, \tag{54}$$

where ∂Q denotes the boundary of the quadrant. (This sort of extension of the maximum principle to unbounded domains is due to Phragmen and Lindelöf; see, *e.g.*, [25].)

The first consequence is that, from (11), (12), and (18), it follows that (12) holds if x is real or if $x = \mathrm{i}Y$ with $Y \in \mathbb{R}$, *i.e.*, the bound (12) holds on both the real and imaginary axes. Further, from (20) and the asymptotics of $\mathrm{erfc}(x)$ in the complex plane [2, (7.1.23)], it follows that $F(x) \to 0$, uniformly in $\arg(x)$, for $0 \le \arg(x) \le \pi/2$; moreover, it is clear from (6) that the same holds for $F_N(x)$ and hence for $E_N(x)$. Thus (54) implies that (12) holds for $0 \le \arg(x) \le \pi/2$, and (16) and (17) then imply that (12) holds also for $\pi \le \arg(x) \le 3\pi/4$.

It is clear from the derivations above that, if h is given by (26), then I_h^* also satisfies the bound (12), *i.e.*,

$$|F(x) - I_h^*| \le c_N \frac{\mathrm{e}^{-\pi N}}{\sqrt{N+1/2}}\,, \tag{55}$$

this holding in the first instance for real x, then for imaginary x, and finally for all x in the first and third quadrants. The bound (12) cannot hold in the second or fourth quadrant because $E_N(x) = F(x) - F_N(x)$ has poles there. This issue does not apply to $F(x) - I_h^*$, which is an entire function, but (55) cannot hold in the whole complex plane because this, by Liouville's theorem ([25]), would imply that $F(x) - I_h^*$ is a constant. What does hold is that $\mathrm{e}^{-\mathrm{i}x^2}(F(x) - I_h^*)$ is bounded in the second and fourth quadrants, this a consequence of the definition of I_h^* and the asymptotics of $\mathrm{e}^{z^2}\mathrm{erfc}(z)$ at infinity. Thus it follows from (54), and since $|\mathrm{e}^{-\mathrm{i}x^2}| = 1$ if x is real or pure imaginary, that

$$|F(x) - I_h^*| \le c_N \mathrm{e}^{-XY}\, \frac{\mathrm{e}^{-\pi N}}{\sqrt{N+1/2}}\,, \tag{56}$$

for $x = X + \mathrm{i}Y$ in the second and fourth quadrants.

We can use the bound (56) to obtain a bound on $E_N(x)$ in the second and fourth quadrants. Clearly, where T_N is defined by (44), with h given by (26), for $x = X + \mathrm{i}Y$ in the second and fourth quadrants,

$$|F(x) - F_N(x)| \le c_N \mathrm{e}^{-XY}\, \frac{\mathrm{e}^{-\pi N}}{\sqrt{N+1/2}} + |T_N|.$$

Further, arguing as below (44), if $|Y| \le A_N/(2\sqrt{2})$ so that

$$|x^2 + \mathrm{i}t_k^2| \ge \left(\frac{A_N}{\sqrt{2}} - |Y|\right)\left(\left(\frac{A_N}{\sqrt{2}} - |Y|\right)^2 + \left(\frac{A_N}{\sqrt{2}} + |X|\right)^2\right) \ge \frac{A_N}{2\sqrt{2}}\left(A_N^2/8 + |X|^2\right),$$

which implies that $|x^2 + \mathrm{i}t_k^2| \ge |x|A_N/(2\sqrt{2})$, then

$$|T_N| \le \mathrm{e}^{-XY}\, \frac{(2\pi+1)\sqrt{2}}{\pi A_N^2}\, \mathrm{e}^{-A_N^2} = \mathrm{e}^{-XY}\, \frac{\sqrt{2}(2\pi+1)}{\pi^{3/2}\exp(\pi/2)(N+1/2)}\, \mathrm{e}^{-\pi N}.$$

Thus, for $x = X+\mathrm{i}Y$ in the second and fourth quadrants with $|Y| \leq A_N/(2\sqrt{2})$,

$$|F(x) - F_N(x)| \leq \hat{c}_N \mathrm{e}^{-XY} \frac{\mathrm{e}^{-\pi N}}{\sqrt{N+1/2}} \tag{57}$$

where

$$\hat{c}_N := c_N + \frac{\sqrt{2}(2\pi+1)}{\pi^{3/2}\exp(\pi/2)\sqrt{N+1/2}}. \tag{58}$$

The sequence $\hat{c}_N$ is decreasing with $\hat{c}_1 \approx 1.14$ and $\lim_{N\to\infty} \hat{c}_N = \lim_{N\to\infty} c_N \approx 0.208$.

We observe above that the bound (12) on $E_N(x) = F(x) - F_N(x)$ holds for all complex x in the first and third quadrants of the complex plane, and on the boundaries of those quadrants, the real and imaginary axes, while the bound (57) holds in the second and fourth quadrants for $|\mathrm{Im}(x)| \leq A_N/(2\sqrt{2})$. A significant implication of these bounds is that they imply that the coefficients in the Maclaurin series of $F_N(x)$ are close to those of $F(x)$. Precisely, at least for $|x| < A_N/\sqrt{2}$,

$$F(x) = \sum_{n=0}^{\infty} a_n x^n \quad \text{and} \quad F_N(x) = \sum_{n=0}^{\infty} b_n x^n,$$

with $a_n = F^{(n)}(0)/n!$, $b_n = F_N^{(n)}(0)/n!$. Thus, where $M_N = \sup_{|z|<\sqrt{\pi/2}} |E_N(z)|$, it follows from Cauchy's estimate [25, Theorem 10.26] and the bounds (12) and (57) that, for $N \geq 4$ so that $A_N/(2\sqrt{2}) \geq \sqrt{\pi/2}$,

$$|a_n - b_n| = \frac{|E_N^{(n)}(0)|}{n!} \leq M_N \left(\frac{2}{\pi}\right)^{n/2} \leq \hat{c}_N \left(\frac{2}{\pi}\right)^{n/2} \frac{\mathrm{e}^{-\pi(N-1/2)}}{\sqrt{N+1/2}}. \tag{59}$$

We will use this bound to derive (69) and (70) below.

3 Approximating $C(x)$ and $S(x)$

From (3) we see that, for x real,

$$C(x) = \mathrm{Re}\left(\sqrt{2}\,\mathrm{e}^{\mathrm{i}\pi/4}(\tfrac{1}{2} - F(\sqrt{\pi/2}\,x))\right), \;\; S(x) = \mathrm{Im}\left(\sqrt{2}\,\mathrm{e}^{\mathrm{i}\pi/4}(\tfrac{1}{2} - F(\sqrt{\pi/2}\,x))\right). \tag{60}$$

Clearly, given the approximation $F_N(x)$ to $F(x)$, these relationships can be used to generate approximations for the Fresnels integrals $C(x)$ and $S(x)$. These approximations are defined, for $x \in \mathbb{R}$, by

$$\begin{aligned} C_N(x) &= \mathrm{Re}\left(\sqrt{2}\,\mathrm{e}^{\mathrm{i}\pi/4}(\tfrac{1}{2} - F_N(\sqrt{\pi/2}\,x))\right), \\ S_N(x) &= \mathrm{Im}\left(\sqrt{2}\,\mathrm{e}^{\mathrm{i}\pi/4}(\tfrac{1}{2} - F_N(\sqrt{\pi/2}\,x))\right), \end{aligned} \tag{61}$$

and are given explicitly in (8) and (9). We note the similarity between (8) and (9) and the formulae [1, (7.5.3)-(7.5.4)]

$$C(x) = \tfrac{1}{2} + f(x)\sin\left(\tfrac{1}{2}\pi x^2\right) - g(x)\cos\left(\tfrac{1}{2}\pi x^2\right), \qquad (62)$$
$$S(x) = \tfrac{1}{2} - f(x)\cos\left(\tfrac{1}{2}\pi x^2\right) - g(x)\sin\left(\tfrac{1}{2}\pi x^2\right), \qquad (63)$$

which express $C(x)$ and $S(x)$ in terms of the auxiliary functions, $f(x)$ and $g(x)$, for the Fresnel integrals [1, §7.2(iv)]. Indeed, it follows from [1, (7.7.10)-(7.7.11)] that, for $x > 0$, $f(x)$ and $g(x)$ have the integral representations

$$f(x) = \frac{\sqrt{\pi}\,x^3}{2}\int_0^\infty \frac{\mathrm{e}^{-t^2}}{\left(\frac{\pi}{2}x^2\right)^2 + t^4}\,\mathrm{d}t \quad \text{and} \quad g(x) = \frac{x}{\sqrt{\pi}}\int_0^\infty \frac{t^2\mathrm{e}^{-t^2}}{\left(\frac{\pi}{2}x^2\right)^2 + t^4}\,\mathrm{d}t.$$

Recalling that A_N is linked to the quadrature step-size through (26), it is clear that, for $x > 0$, $\sqrt{\pi}\,xa_N\left(\frac{\pi}{2}x^2\right)/A_N$ and $\sqrt{\pi}\,xb_N\left(\frac{\pi}{2}x^2\right)/A_N$ can be viewed as quadrature approximations to these integrals.

Table 2 shows the Matlab code implementing (8) and (9) that we use for the computations in the next section. Some comments of explanation are in order regarding the evaluation of

$$\frac{\sinh t \pm \sin t}{\cosh t + \cos t}, \qquad (64)$$

with $t = \sqrt{\pi}\,A_N x$, in (8) and (9). An issue in floating point arithmetic for evaluation of (64) for larger values of $|t|$ is overflow; for $t \geq 711$ both $\sinh t$ and $\cosh t$ evaluate in the IEEE double precision arithmetic implemented in Matlab as `Inf`, the IEEE representation for $+\infty$, so that (64) is undefined. But, since u is indistinguishable from $u \pm v$ in IEEE double precision real arithmetic if $|v|/|u| \leq \varepsilon/2$, where $\varepsilon = 2^{-52} \approx 2.22 \times 10^{-16}$ (`eps` in Matlab) is the smallest number such that $1 + \varepsilon$ is distinguishable from 1, it makes sense to replace (64) by ± 1 well before this point. Precisely, for $|t| \geq 39$, the expressions $\cosh(t) + \cos(t)$ and $\exp(t)/2$ have the same value in double precision arithmetic, as do the expressions $\sinh t \pm \sin t$ and $\mathrm{sign}(t)\exp(t)/2$. Thus (64) evaluates as $\mathrm{sign}(t)$ in double precision arithmetic for $39 \leq |t| \lessapprox 710$. It makes sense then, both to avoid overflow and to reduce computation time, to evaluate (64) as $\mathrm{sign}(t)$ for $|t| \geq 39$ (which corresponds to $|x| \geq 39/(\sqrt{\pi}A_N) \approx 3.51$ for $N = 12$).

For small t there is an additional issue in evaluation of (64) when the negative sign is chosen, that $\sinh t \approx \sin t \approx t$ for small t, so that there is loss of precision in evaluating $\sinh t - \sin t$ for $|t|$ small. This is avoided in the code in Table 2 by using the power series $\sinh t - \sin t = 2t^3/3! + 2t^7/7! + \ldots$ for $|t| < 1$, truncating this after four terms as the fifth term is negligible in double precision arithmetic for $|t| < 1$.

The approximations (8) and (9) inherit the accuracy of $F_N(x)$ on the real line: from (60) and (61) we see that the bounds (15) hold for $x \in \mathbb{R}$, where $E_N(x) = F(x) - F_N(x)$. Thus the error bounds of the previous section can be

```
function [C,S] = fresnelCS(x,N)
% Evaluates approximations to the Fresnel integrals C(x) and S(x).
% x is a real scalar or matrix,
% N is a positive integer controlling accuracy (suggest N=12),
% C and S are the scalars/matrices of the same size as x approximating C(x) and S(x).
h = sqrt(pi/(N+0.5));
t = h*((N:-1:1)-0.5);  AN = pi/h; rootpi = sqrt(pi);
t2 = t.*t; t4 = t2.*t2; et2 = exp(-t2);
x2pi2 = (pi/2)*x.*x; x4 = x2pi2.*x2pi2;
a = et2(1)./(x4+t4(1)); b = t2(1)*a;
for n = 2:N
    term = et2(n)./(x4+t4(n));
    a = a + term;  b = b + t2(n)*term;
end
a = a.*x2pi2;
mx = (rootpi*AN)*x; Mx = (rootpi/AN)*x;
Chalf = 0.5*sign(mx); Shalf = Chalf;
select = abs(mx)<39;
if any(select)
    mxs = mx(select); shx = sinh(mxs); sx = sin(mxs);
    den = 0.5./(cos(mxs)+cosh(mxs));
    Chalf(select) = (shx+sx).*den;
    ssdiff = shx-sx;
    select2 = abs(mxs)<1;
    if any(select2)
        mxs = mxs(select2); mxs3 = mxs.*mxs.*mxs; mxs4 = mxs3.*mxs;
        ssdiff(select2) = mxs3.*(1/3 + mxs4.*(1/2520 ...
            + mxs4.*((1/19958400)+(0.001/653837184)*mxs4)));
    end
    Shalf(select) = ssdiff.*den;
end
cx2 = cos(x2pi2); sx2 = sin(x2pi2);
C = Chalf + Mx.*(a.*sx2-b.*cx2); S = Shalf - Mx.*(a.*cx2+b.*sx2);
```

Table 2 Matlab code to evaluate $C_N(x)$ and $S_N(x)$ given by (8) and (9). See §3 for details.

applied. In particular, from (50) and (52) it follows that both $|C(x) - C_N(x)|$ and $|S(x) - S_N(x)|$ are

$$\leq 2c_N \frac{\mathrm{e}^{-\pi N}}{\sqrt{2N+1}}, \quad \text{for } x \in \mathbb{R}, \tag{65}$$

and

$$\leq \sqrt{\pi}\,\tilde{c}_N |x|\, \frac{\mathrm{e}^{-\pi N}}{2N+1}, \quad \text{for } |x| \leq \sqrt{N+1/2}\,. \tag{66}$$

Here $c_N < 0.83$ and $\tilde{c}_N < 0.18$ are the decreasing sequences of positive numbers defined by (8) and (53), respectively.

These bounds show that $C_N(x)$ and $S_N(x)$ are exponentially convergent as $N \to \infty$, uniformly on the real line, so that very accurate approximations can be obtained with very small values of N ((65) shows that both $|C_N(x) - C(x)|$ and $|S_N(x) - S(x)|$ are $\leq 1.4 \times 10^{-16}$ on the real line for $N \geq 11$). In §4 we will confirm the effectiveness of these approximations by numerical experiments, checking the accuracy of (8) and (9) by comparison with the power series [1, §7.6(i)]

$$C(x) = \sum_{n=0}^{\infty} \frac{(-1)^n \left(\frac{1}{2}\pi\right)^{2n} x^{4n+1}}{(2n)!(4n+1)}, \quad S(x) = \sum_{n=0}^{\infty} \frac{(-1)^n \left(\frac{1}{2}\pi\right)^{2n+1} x^{4n+3}}{(2n+1)!(4n+3)}. \quad (67)$$

It follows from the analyticity of $F_N(x)$ in the complex plane, discussed in §2, that $F_N(x)$ has a Maclaurin series convergent in $|x| < A_N/\sqrt{2}$, and from (61) that $C_N(x)$ and $S_N(x)$ have convergent Maclaurin series representations in $|x| < A_N/\sqrt{\pi}$. From the observations below (19) it is clear that, echoing (67), these take the form

$$C_N(x) = \sum_{n=0}^{\infty} \mathfrak{c}_n x^{4n+1}, \quad S_N(x) = \sum_{n=0}^{\infty} \mathfrak{s}_n x^{4n+3}. \quad (68)$$

Further, it follows from (61) and (59) that the coefficients $\mathfrak{c}_n$ and $\mathfrak{s}_n$ are close to the corresponding coefficients of $C(x)$ and $S(x)$, with the difference having absolute value

$$\leq \sqrt{2}\,\hat{c}_N \frac{\mathrm{e}^{-\pi(N-1/2)}}{\sqrt{N+1/2}}, \quad (69)$$

for $N \geq 4$, where $\hat{c}_N \leq \hat{c}_4 < 0.77$ is the decreasing sequence of positive numbers given by (58). This implies that, near zero, where $C(x)$ has a simple zero and $S(x)$ a zero of order three, the approximations $C_N(x)$ and $S_N(x)$ retain small relative error. For $C_N(x)$ this follows already from (66) but to see this for $S_N(x)$ we need the stronger bound implied by (69) that, for $|x| < 1$,

$$|S(x) - S_N(x)| \leq \sqrt{2}\,\hat{c}_N \frac{\mathrm{e}^{-\pi(N-1/2)}}{\sqrt{N+1/2}} \sum_{n=0}^{\infty} |x|^{4n+3} = \frac{|x|^3}{1-|x|^4} \frac{\sqrt{2}\,\hat{c}_N\, \mathrm{e}^{-\pi(N-1/2)}}{\sqrt{N+1/2}}. \quad (70)$$

3.1 Other Methods for Computing Fresnel Integrals

Naturally, there exist already a number of effective schemes for computation of Fresnel integrals, and we briefly summarise now the best of these. An effective computational method for smaller values of $|x|$ is to make use of the power series (67). These converge for all x, and very rapidly for smaller x, and so are widely used for computation. For example, the algorithm in the standard reference [23] uses these power series for $|x| \leq 1.5$. For this range, after the first two terms, these series are alternating series of monotonically decreasing terms, and the error in truncation has magnitude smaller than the first neglected term. Thus, for $|x| \leq 1.5$, the errors in computing $C(x)$ and $S(x)$ by these power series truncated to N terms are $\leq 2 \times 10^{-16}$ and $\leq 2.3 \times 10^{-17}$, respectively, for $N = 14$.

For $|x| > 1.5$, [23] recommends computation using (60) and (20) and the continued fraction representation for $e^{z^2}\mathrm{erfc}(z) = w(\mathrm{i}z)$ given as [1, (7.9.2)].

Methods for evaluation of $w(z)$ based on continued fraction representations for larger complex z (which can be applied to evaluate $F(x)$ and hence $C(x)$ and $S(x)$) are also discussed in Gautschi [11] and are finely tuned, to form TOMS "Algorithm 680", in Poppe and Wijers [21,22], which achieves relative errors of 10^{-14} over "nearly all" the complex plane by using Taylor expansions of degree up to 20 in an ellipse around the origin, convergents of up to order 20 of continued fractions outside a larger ellipse, and a more expensive mix of Taylor expansion and continued fraction calculations in between.

Weideman [29] presents an alternative method of computation (the derivation starts from the integral representation (21)) which approximates $w(z)$ by the polynomial

$$w_M(z) = \frac{2}{L^2 + z^2} \sum_{n=0}^{M} a_n Z^n \tag{71}$$

in the transformed variable $Z = (L + \mathrm{i}z)/(L - \mathrm{i}z)$. Here $L = \sqrt{M/\sqrt{2}}$ and the coefficients a_n can be viewed as Fourier coefficients and efficiently computed by the FFT. We will see in §4 that a polynomial degree $M = 36$ in (71) suffices to compute $F(x) = \mathrm{e}^{\mathrm{i}x^2} w(\mathrm{e}^{\mathrm{i}\pi/4}x)/2$ with relative error $\leq 10^{-15}$ uniformly on the real line. Weideman [29] argues carefully and persuasively that, in terms of operation counts, the work required to compute $w(z)$ with the 10^{-14} relative accuracy of Algorithm 680 [22] is much smaller using the approximation (71) for intermediate values of $|z|$ (values in approximately the range $1.5 \leq |z| \leq 5$ for the case $\arg(z) = \pi/4$ which we require).

All these approximations described above are polynomial or rational approximations (or piecewise polynomial/rational approximations, proposing different approximations on different regions). Many other authors describe approximations of these types for computing the Fresnel integrals specifically with real arguments. The best of these in terms of accuracy is Cody [8], where numerical coefficient values are given for piecewise rational approximations to $C(x)$ and $S(x)$ for $0 \leq x \leq 1.6$, and for piecewise rational approximations to $f(x)$ and $g(x)$ in (62) and (63), for $x \geq 1.6$. These approximations, in their respective regions of validity, achieve relative errors $\leq 10^{-15.58} \approx 2.7 \times 10^{-16}$, this using rational approximations which are ratios of polynomials of degree ≤ 6; in total five different approximations are used on different subintervals of the positive real axis. Single rational approximations, based on a "polar" version of (62) and (63), are computed in [14], but these are of limited accuracy (absolute errors $\leq 4 \times 10^{-8}$).

4 Numerical Results and Comparison of Methods

In this section we show the results of numerical computations that confirm and illustrate the theoretical error bounds in §2 and §3, and that explore the accuracy and efficiency of our new methods, through qualitative and quantitative comparisons with certain of the other computational methods described in §3.1.

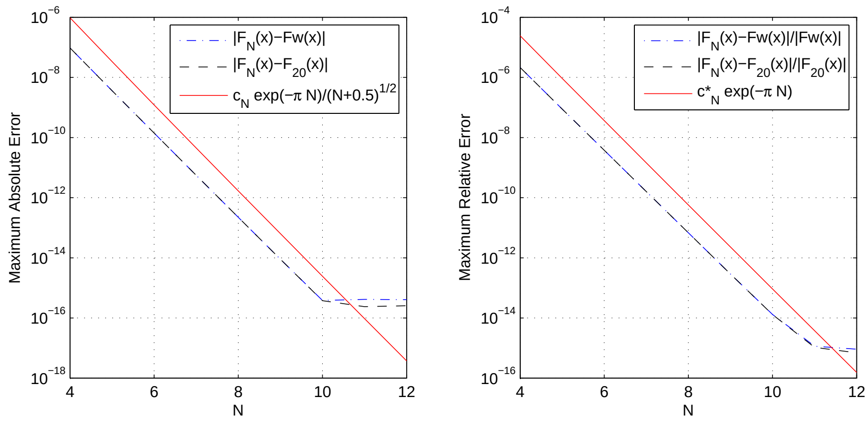


Fig. 1 Left hand side: maximum error, $\max_{x\geq 0}|F(x)-F_N(x)|$, and its upper bound (12) (—), plotted against N, where $F(x)$ is approximated by $Fw(x) := e^{ix^2}w_{36}(e^{i\pi/4}x)/2$ with $w_{36}(z)$ given by (71) (—·—·), and where $F(x)$ is approximated by $F_{20}(x)$ (−−). Right hand side: maximum relative error, $\max_{x\geq 0}|(F(x)-F_N(x))/F(x)|$, and its upper bound (51) (—), plotted against N, where $F(x)$ is approximated as on the left hand side.

Turning first to the approximation (5) to $F(x)$, in Figure 1 we plot against N the maximum values of the absolute and relative errors, $|E_N(x)| = |F(x) - F_N(x)|$ and $|F(x) - F_N(x)|/|F(x)|$, on $x \geq 0$, approximating these maximum values on $[0, \infty)$ by computing at 40,000 equally spaced points between 0 and 1,000 and replacing $F(x)$ either by $F_{20}(x)$ or by $Fw(x) := e^{ix^2}w_{36}(e^{i\pi/4}x)/2$ with $w_{36}(z)$ given by (71). (We compute $F_N(x)$ in Matlab using the code in Table 1, and $Fw(x)$ by `exp(i*x.^2).*cef(exp(i*pi/4)*x,36)/2`, where `cef.m` is the function in Table 1 of [29] which evaluates (71). The choice $M = 36$ in (71) is made because this is the smallest value which appears to achieve relative errors $\leq 10^{-15}$ uniformly on the positive axis, and increasing M further in the range $M \leq 40$ appears to lead to no increase in accuracy by comparison with the same approximation with $M = 50$; see Figure 2 below.) We show in the same plots the upper bounds which are the right hand sides of (12) and (51). It can be seen that the exponential convergence predicted by the bounds (12) and (51) is achieved, indeed these bounds overestimate their respective maximum errors by at most a factor of 10. Further, with N as small as 12 it appears that we achieve maximum absolute and relative errors in $F_N(x)$ which are $< 2.9\times 10^{-16}$ and $< 9.3\times 10^{-16}$, respectively; these values are upper bounds whichever of the two methods for approximating $F(x)$ accurately is used. (We should add a note of caution here: the different approximations agree to high accuracy, but the accuracy of each approximation is limited, for large x, by the accuracy with which e^{ix^2} is computed.)

The plots in Figure 1, in addition to shedding light on the accuracy of $F_N(x)$, provide independent verification of the high accuracy of the approximation (71) for $w(z)$ proposed in [29], at least for $\arg(z) = \pi/4$ and provided M

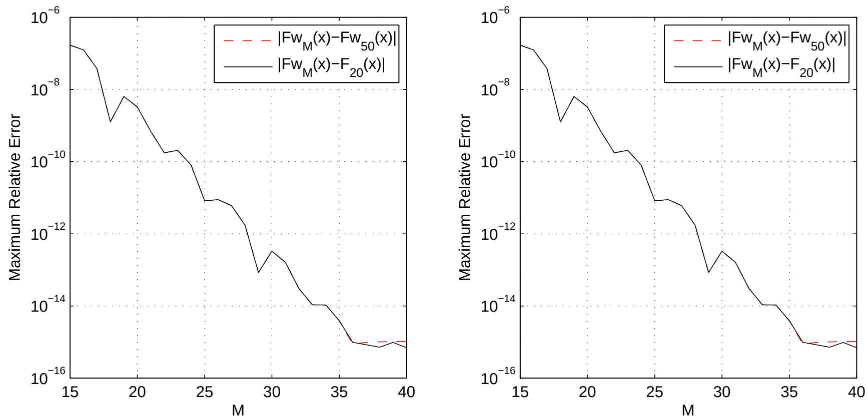


Fig. 2 Left hand side: maximum error, $\max_{x\geq 0}|F(x) - Fw(x)|$, where $Fw(x) := \mathrm{e}^{\mathrm{i}x^2} w_M(\mathrm{e}^{\mathrm{i}\pi/4}x)/2$ with $w_M(z)$ given by (71), plotted against M. Right hand side: same, but maximum relative error, $\max_{x\geq 0}|(F(x) - Fw(x))/F(x)|$, is plotted against M. In each plot the two curves correspond to different methods for approximating the exact value of $F(x)$, either $F(x) \approx F_{20}(x)$ given by (5) (−), or $F(x) \approx Fw(x)$ with $M = 50$ (−−).

is large enough in (71). Exploring this in more detail, in Figure 2 the maximum absolute and relative errors in the approximation $Fw(x) = \mathrm{e}^{\mathrm{i}x^2} w_M(\mathrm{e}^{\mathrm{i}\pi/4}x)/2$ for $F(x)$, with $w_M(z)$ given by (71), are plotted against M. (The maxima, as in Figure 1, are taken over 40,000 equally spaced points between 0 and 1,000.) In each of the plots in Figure 2 the trend is one of exponential convergence, but the convergence is not monotonic and is slower than that in Figure 1.

In Figure 3 we plot against x the absolute and relative errors in $F_N(x)$ for $N = 9$. On the same graphs we plot the upper bounds $\eta_N(x)$ and $2(1 + \sqrt{\pi}\,x)\eta_N(x)$, respectively, with $\eta_N(x)$ defined by (46). We see that the theoretical error bounds are upper bounds as claimed, and that these bounds appear to capture the x-dependence of the errors fairly well, for example that $E_N(x) = O(x)$ as $x \to 0$, $= O(x^{-1})$ as $x \to \infty$, and that $E_N(x)$ reaches a maximum at about $x = \sqrt{2}\,A_N = \sqrt{\pi(2N+1)}$ (≈ 7.7 when $N = 9$).

The above figures explore the accuracy of the approximation $F_N(x)$. Let us comment now on efficiency. Most straightforward is a comparison of the Matlab function `F(x,N)` in Table 1 with computation of $F(x)$ via the Matlab code `exp(i*x.^2).*cef(exp(i*pi/4)*x,36)/2` that uses `cef.m` from [29] which implements (71). Both `F(x,N)` and `cef(x,M)` are optimised for efficiency when `x` is a large vector. Assuming that the time for computation in `cef` of the coefficients a_n in (71) is negligible, the main cost in computation of $F(x)$ via `cef` when x is a large vector is a complex vector exponential (for $\mathrm{e}^{\mathrm{i}x^2}$), slightly more than M complex vector multiplications and M additions, and 2 complex vector divisions (all vector operations componentwise). The major part of this computation is that required to evaluate the polynomial (71) of degree M us-

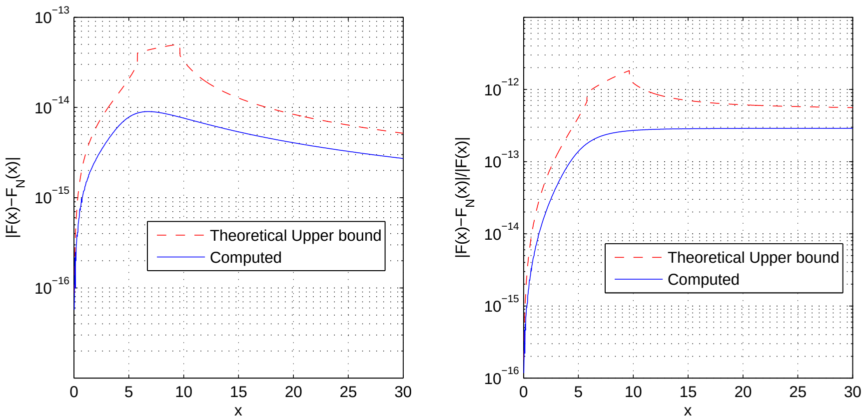


Fig. 3 Left hand side: absolute error, $|F(x) - F_N(x)|$ (—), and its upper bound $\eta_N(x)$ given by (46) (−−), plotted against x. Right hand side: relative error, $|F(x) - F_N(x)|/|F(x)|$ (—), and its upper bound $2(1 + \sqrt{\pi}\,x)\eta_N(x)$ (−−), plotted against x. In both plots $N = 9$ and $F(x)$ is approximated by $F_{20}(x)$.

ing Horner's algorithm. In comparison, evaluation of $F(x)$ using the function `F(x,N)` in Table 1 requires 2 complex vector exponentials, one complex vector division, and slightly more than N real vector multiplications/divisions, real vector additions, complex vector multiplications, and complex vector additions. From Figures 1 and 2 we read off that to achieve absolute and relative errors below 10^{-8} requires $N = 6$ and $M = 18$; to achieve errors below 10^{-15} requires $N = 12$ and $M = 36$. Thus it seems clear that computing $F(x)$ via `F(x,N)` requires a substantially lower operation count than computing via `cef`. (We note, moreover, as discussed in §3.1 and in §7 of [29], that, at least for intermediate values of x ($1.5 \le x \le 5$), the operation counts via `cef` are lower than those required via the method for $w(z)$ of [21,22].)

To test whether `F(x,N)` is faster we have compared computation times in Matlab (version 7.8.0.347 (R2009a), running on a laptop with dual 2.4GHz P8600 Intel processors) between `exp(i*x.^2).*cef(exp(i*pi/4)*x,36)/2` and `F(x,12)` when `x` is a length 10^7 vector of equally spaced numbers between 0 and 1,000. The elapsed times (average of 10 executions) were 11.1 and 15.6 seconds, respectively, so that `F(x,12)` is a little less than 50% faster.

Turning to $C(x)$ and $S(x)$, these can of course be computed using `F(x,N)` to calculate $F_N(x)$, and then using (61) which incurs negligible additional computation. This is entirely satisfactory except for small x, where this method fails to maintain small relative errors. As discussed in §3, the Matlab function `fresnelCS.m` in Table 2 directly implements (8) and (9), taking care in the evaluation of $\sinh - \sin$ in (9) so as to achieve the high accuracy of $S_N(x)$ for small $|x|$ predicted in (70). To test the efficiency and accuracy of the implementation in Table 2 we have compared evaluation of $C_{12}(x)$ and $S_{12}(x)$ via `fresnelCS` with their evaluation via `F(x,12)` and (61), computing

$C_{12}(x)$ and $S_{12}(x)$ at 10^7 equally spaced x-values between 0 and 20. The values of $C_{12}(x)$ and $S_{12}(x)$ computed by these slightly different methods differ by $\leq 4.5 \times 10^{-15}$; this good but not perfect agreement is because there is a difference between $\exp(\mathrm{i}(\sqrt{\pi/2}\,x)^2)$ and $\exp(\mathrm{i}\pi x/2)$ in floating point arithmetic. In this test `fresnelCS` requires only 67% of the computation time of computing via `F(x,12)`, this because the real arithmetic in `fresnelCS` is faster and because the expressions (64), with $t = \sqrt{2}\,A_N x$, are evaluated (efficiently and accurately) in `fresnelCS` as $\mathrm{sign}(t)$ when $|t| \geq 39$ (corresponding to $x \geq 3.51$ for $N = 12$), as discussed in §3.

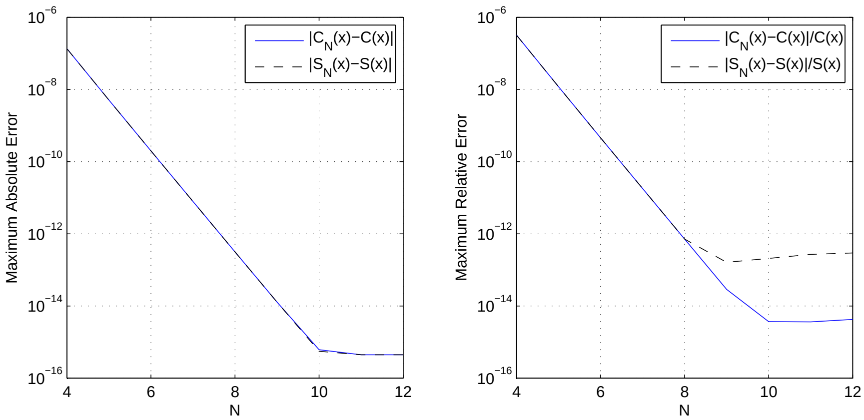


Fig. 4 Left hand side: maximum values of $|C_N(x)-C(x)|$ and $|S_N(x)-S(x)|$ on $0 \leq x \leq 20$. Right hand side: maximum values of $|C_N(x) - C(x)|/C(x)$ and $|S_N(x) - S(x)|/S(x)$ on $0 \leq x \leq 20$.

These small absolute errors in $C_N(x)$ and $S_N(x)$, evaluated by `fresnelCS`, do not guarantee small relative errors near the only zero of $C(x)$ and $S(x)$ at $x = 0$. Near zero, from (67), $C(x) \approx x$ and $S(x) \approx \pi x^3/6$, so very accurate calculations are needed to maintain small relative errors. The bounds (66) and (70) do, in fact, guarantee small relative errors near zero in infinite precision arithmetic, moreover predicting that these relative errors should decrease approximately in proportion to $\mathrm{e}^{-\pi N}$ as N increases. To see if this convergence is achieved in floating point arithmetic by `fresnelCS`, in Figure 4 we have plotted the maximum values, on $(0, 20]$, of the absolute errors, $|C_N(x)-C(x)|$ and $|S_N(x)-S(x)|$, and the relative errors, $|C_N(x)-C(x)|/C(x)$ and $|S_N(x) - S(x)|/S(x)$. In this figure we compute $S_N(x)$ and $C_N(x)$ by `fresnelCS` and approximate $C(x)$ and $S(x)$ by $C_{20}(x)$ and $S_{20}(x)$ for $x > 1.5$. For $0 < x < 1.5$ (as recommended in [23]) we approximate by the power series (67) truncated after 15 terms, and evaluating the resulting polynomials by the usual Horner algorithm. Exponential convergence in both the absolute and relative errors is seen in Figure 4, but we note that, while the absolute errors are $\leq 4.5 \times 10^{-16}$ for $N \geq 11$, the maximum relative error in $C_N(x)$ is $\approx 3.6 \times 10^{-15}$ for $N = 11$ and that in $S_N(x)$ as large as 2.7×10^{-13}. These errors may be entirely acceptable, but the truncated power series (67) must achieve smaller errors for small x, and may be cheaper to evaluate. (In fact, evaluating at 10^7 equally spaced points between 0 and 1.5 takes 2.9 times longer in Matlab with `fresnelCS` than evaluating 15 terms of both the series (67) via Horner's algorithm.)

5 Extensions and Concluding Remarks

To conclude, we have presented in this paper new approximations for the Fresnel integrals, derived from and inspired by modified trapesium rule approximations previously suggested for the complementary error function of complex argument in [18, 15]. These approximations are simple to implement (Matlab codes are included in Tables 1 and 2): the computation of $F_N(x)$ requires a couple of complex exponentiations and a short summation to compute a quadrature sum, and that of $C_N(x)$ and $S_N(x)$ evaluation of trigonometric and hyperbolic functions and a similar short summation. The numerical methods are proven to converge exponentially (in absolute and relative error), approximately in proportion to $\exp(-\pi N)$ where N is the number of quadrature points used. Simple explicit error bounds are provided, and the predicted exponential convergence is precisely observed in practice. The approximation $F_N(x)$ with $N = 12$ quadrature points achieves close to double precision machine precision uniformly on the real line (with the proviso that this precision is necessarily limited by the accuracy with which $\mathrm{e}^{\mathrm{i}x^2}$ can be calculated). The approximations $C_N(x)$ and $S_N(x)$ with $N = 11$ are similarly accurate, except that the relative error in $S_N(x)$ increases to 2.7×10^{-13} near $x = 0$ where $S(x)$ has a zero of order 3.

Operation counts and timings carried out suggest that $F_N(x)$ with $N = 12$ may be faster than previous methods, at least for intermediate values of $|x|$. In particular, the Matlab function in Table 1 outperforms that in Table 1 of [29] for this application. The code for $S_N(x)$ and $C_N(x)$ is faster still, but the power series (67), truncated after 15 terms, are more accurate and efficient on the interval $[0, 1.5]$, this conclusion endorsing recommendations in [23].

Part of the motivation for this paper was a remark in Weideman [29] regarding the modified trapezium rule methods of [18, 15] for computing $\mathrm{erfc}(z)$, that they are "very accurate, provided for given z and N [the finite number of quadrature points retained] the optimal stepsize h is selected. It is not easy, however, to determine this optimal h a priori." At least as far as computing $\mathrm{erfc}(z)$ for $\arg(z) = -\pi/4$ is concerned (which, by (20), is the same as computing $F(x)$) this problem is solved in this paper, so that the effectiveness of the modified trapezium rule methods of [18, 15, 29] is clearly demonstrated. We hope that the methodology and positive results of this paper will inspire further applications of this truncated, modified trapezium rule method.

With respect to this hope, most obviously the results in this paper suggest a revisit of the methods of [18, 15] for $\mathrm{erfc}(z)$. Clearly, (20) suggests $\mathrm{erfc}_N(z) := 2F_N(\mathrm{e}^{\mathrm{i}\pi/4}z)$, given explicitly as

$$\mathrm{erfc}_N(z) = \frac{2}{\mathrm{e}^{2A_N z}+1} + \frac{2z}{A_N}\,\mathrm{e}^{-z^2}\sum_{k=1}^{N}\frac{\mathrm{e}^{-t_k^2}}{z^2+t_k^2}, \tag{72}$$

as an approximation for $\mathrm{erfc}(z)$. (For $0 < \mathrm{Re}(z) < A_N$, this is precisely the approximation of [15] truncated to N quadrature points and with the particular choice (26) for h made.) The results of §2 show that (12) holds for $0 \le \arg(x) \le \pi/2$ and for $\pi \le \arg(x) \le 3\pi/4$ which implies that

$$|\mathrm{erfc}(z) - \mathrm{erfc}_N(z)| \le 2c_N\,\frac{\mathrm{e}^{-\pi N}}{\sqrt{N+1/2}} < 2\frac{\mathrm{e}^{-\pi N}}{\sqrt{N+1/2}}, \tag{73}$$

for $|\arg(z)| \le \pi/4$ and $3\pi/4 \le \arg(z) \le 5\pi/4$. This is a strong result in $3\pi/4 \le \arg(z) \le 5\pi/4$, where it is known that $|\mathrm{erfc}(z)| \ge 1$ so that (73) is a bound on both the absolute and relative error. However, in $|\arg(z)| < \pi/4$ the bound (73) is less satisfactory. In particular, since $\mathrm{erfc}(x) \sim \mathrm{e}^{-x^2}/(\sqrt{\pi}x)$ as $x \to +\infty$, for larger $x > 0$ (73) does not guarantee small relative errors. Indeed, $\mathrm{erfc}_N(x) \sim 2\mathrm{e}^{-2A_N x}$ has the wrong asymptotic behaviour as $x \to +\infty$.

A large part of a possible fix and analysis is already in [18] and [15] (see equations (7)-(8) in [15] and cf. (27), [19]), namely to discard the first term in (73) for $\mathrm{Re}(z) > A_N$, so that $\mathrm{erfc}(z)$ is approximated by

$$\mathrm{erfc}'_N(z) = R_N(z) + \frac{2z}{A_N}\,\mathrm{e}^{-z^2}\sum_{k=1}^{N}\frac{\mathrm{e}^{-t_k^2}}{z^2+t_k^2}, \tag{74}$$

where

$$R_N(z) := \begin{cases} 2/(\mathrm{e}^{2A_N z}+1), & \mathrm{Re}(z) \le A_N, \\ 0, & \mathrm{Re}(z) > A_N. \end{cases}$$

Figure 5 plots the supremum, on $0 \le x \le 25$, of the absolute and relative errors in $\mathrm{erfc}'_N(x)$ against N, computing $\mathrm{erfc}(x)$ with the inbuilt Matlab function `erfc`, and with erfc'_{20}. Clearly, both plots show exponential convergence at a rate approximately proportional to $\mathrm{e}^{-\pi N}$. The absolute error in erfc'_N is $< 4.5 \times 10^{-16}$ for $N \ge 10$ and the relative error $< 6.7 \times 10^{-16}$ for $N = 12$, while the maximum relative error of the standard Matlab function is limited to about 5.7×10^{-14}. We have not computed any theoretical upper bounds for these errors, but the fact that $\mathrm{erfc}'_N(x)$ is discontinuous at A_N implies lower bounds: in particular, the supremum of the relative error in any neighbourhood of A_N must be $\ge 1/(\mathrm{erfc}(A_N)(\exp(2A_N^2)+1))$, *i.e.*, half the discontinuity at A_N divided by $\mathrm{erfc}(A_N)$. From (45), $\mathrm{erfc}(x) < \mathrm{e}^{-x^2}/(\sqrt{\pi}x)$ for $x > 0$, so we see that

$$\sup_{0\le x\le 1+A_N}\frac{|\mathrm{erfc}'_N(x) - \mathrm{erfc}(x)|}{\mathrm{erfc}(x)} \ge \frac{\sqrt{\pi}\,A_N\mathrm{e}^{A_N^2}}{\mathrm{e}^{2A_N^2}+1} \approx \sqrt{\pi}\,A_N\mathrm{e}^{-A_N^2}. \tag{75}$$

In the right hand figure we plot this lower bound which accurately predicts the maximum error, this suggestive that the small size of error present is associated with the discontinuity in erfc'_N.

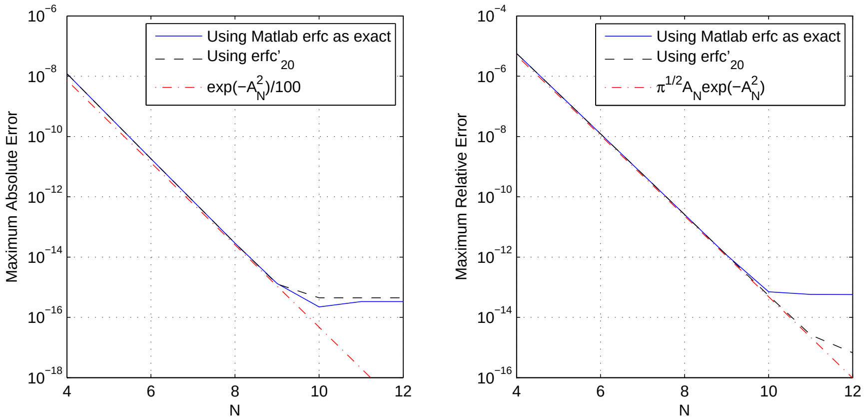


Fig. 5 Left hand side: supremum of $|\mathrm{erfc}'_N(x) - \mathrm{erfc}(x)|$ on $0 \le x \le 25$ and $\mathrm{e}^{-A_N^2}/100$ (−.). Right hand side: supremum of $|\mathrm{erfc}'_N(x) - \mathrm{erfc}(x)|/\mathrm{erfc}(x)$ on $0 \le x \le 25$ and its theoretical lower bound, $\sqrt{\pi}\, A_N \mathrm{e}^{-A_N^2}$ (−.). In each plot two of the curves differ only in how $\mathrm{erfc}(x)$ is approximated: the Matlab built in function (−) and erfc'_{20} (−−).

These are encouraging results, but further work is needed, revisiting [18, 15, 19], to develop a fully discrete and accurate modified trapezium rule method for $\mathrm{erfc}(z)$ for all z in the complex plane. Note, for example, that, while Figure 5 shows that $\mathrm{erfc}'_{12}(z)$ has relative error close to machine precision on the positive real axis, on the imaginary axis this approximation is singular at $z = \pm \mathrm{i} t_k$, for $k \ge 13$.

We finish by flagging that the modified trapezium rule method that we have used in this paper is applicable widely to the evaluation of integrals on the real line of functions that are analytic but with poles near the real axis. Indeed, general theories of the method are presented in Bialecki [3], Hunter [16] (and see [9], [24, §5.1.4]), and in the thesis of one of the authors [17], where the emphasis is on the particular case (23), where the analytic function $f(t) = O(1)$ as $t \to \pm\infty$. Integrals of the form (23) arise in probabilistic applications [9] and as representations in integral form of solutions to linear PDEs with constant coefficients, after solution by Fourier transform methods and deformation of the path of integration to a steepest descent path. One example which continues to be the subject of computational studies [6, 20, 13] is the Green's function for the Helmholtz equation $\Delta u + k^2 u = 0$ in a half-space with an impedance boundary condition, $\partial u/\partial n = \mathrm{i} k \beta u$. Representations for this Green's function in terms of a steepest descent path integral of the form

(23), in both the 2D and 3D cases, are given in [6], and the application of the truncated modified trapezium rule method is discussed in [17].

References

1. Digital Library of Mathematical Functions. National Institute of Standards and Technology, from http://dlmf.nist.gov/, release date: 2010-05-07 (2010)
2. Abramowitz, M., Stegun, I. A.: Handbook of Mathematical Functions, Dover, New York (1964)
3. Bialecki, B.: A modified sinc quadrature rule for functions with poles near the arc of integration, BIT **29**, 464–476 (1989)
4. Bowman, J. J., Senior, T. B. A., Uslenghi, P. L. E.: Electromagnetic and Acoustic Scattering by Simple Shapes, North-Holland, Amsterdam (1969)
5. Chandler-Wilde, S. N., Hewett, D. P., Langdon, S., Twigger, A.: A high frequency boundary element method for scattering by a class of nonconvex obstacles. University of Reading, Department of Mathematics and Statistics Preprint MPS-2012-04 (2012)
6. Chandler-Wilde, S.N., Hothersall, D. C.: Efficient calculation of the Green function for acoustic propagation above a homogeneous impedance plane, J. Sound Vib. **180**, 705-724 (1995)
7. Chiarella, C., Reichel, A.: On the evaluation of integrals related to the error function, Math. Comp. **22**, 137–143 (1968)
8. Cody, W. J.: Chebyshev approximations for the Fresnel integrals, Math. Comp. **22**, 450–453 + s1–s18 (1968)
9. Crouch, E. A. C., Spiegelman, D.: The evaluation of integrals of the form $\int_{-\infty}^{+\infty} f(t)\exp(-t^2)dt$: application to logistic-normal models, J. Amer. Stat. Assoc. **85**, 464–469 (1990)
10. Fettis, H. E., Caslin, J. C., Cramer, K. R.: Complex zeros of the error function and of the complementary error function, Math. Comp. **27**, 401–407 (1973)
11. Gautschi, W.: Efficient computation of the complex error function, SIAM J. Numer. Anal. **7**, 187–198 (1970)
12. Goodwin, E. T.: The evaluation of integrals of the form $\int_{-\infty}^{\infty} f(x)\mathrm{e}^{-x^2}dx$, Proc. Camb. Phil. Soc. **45**, 241–245 (1949)
13. Greengard, L., O'Neil, M., Pataki, A.: On the efficient representation of the half-space impedance Green's function for the Helmholtz equation. arXiv:1109.6708v1 (2011)
14. Heald, M. A.: Rational approximations for the Fresnel integrals, Math. Comp. **44**, 459–461 (1985)
15. Hunter, D. B., Regan, T.: A note on evaluation of the complementary error function, Math. Comp. **26**, 539–541 (1972)
16. Hunter, D. B.: The numerical evaluation of definite integrals affected by singularities near the interval of integration. In: Numerical Integration, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., **357**, pp. 111-120. Kluwer Acad. Publ., Dordrecht (1992)
17. La Porte, S.: Modified Trapezium Rule Methods for the Effcient Evaluation of Green's Functions in Acoustics. PhD Thesis, Brunel University, UK (2007)
18. Matta, F., Reichel, A.: Uniform computation of the error function and other related functions, J. Math. Phys. **34**, 298–307 (1956)
19. Mori, M.: A method for evaluation of the error function of real and complex variable with high relative accuracy, Publ. RIMS, Kyoto Univ. **19**, 1081–1094 (1983)
20. Durán, M., Hein, R., Nédélec, J.-C.: Computing numerically the Green's function of the half-plane Helmholtz operator with impedance boundary conditions, Numer. Math. **107**, 295-314 (2007)
21. Poppe, G. P., Wijers, C. M.: More efficient computation of the complex error function, ACM Trans. Math. Software **16**, 38–46 (1990)
22. Poppe, G. P., Wijers, C. M.: Algorithm 680 – Evaluation of the complex error functon, ACM Trans. Math. Software **16**, 47–47 (1990).
23. Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P.: Numerical Recipes 3rd Edition: The Art of Scientific Computing, Cambridge University Press (2007)

24. Kythe, P. M., Schäferkotter, M. R.: Handbook of Computational Methods for Integration, Chapman and Hall/CRC, Boca Raton, FL (2005)
25. Rudin, W.: Real and Complex Analysis, 3rd Edition, Mc-Graw Hill (1987)
26. Salzer, H.: Formulas for computing the error function of a complex variable, MTAC **5**, 67–70 (1951)
27. Strand, O.: A method for the computation of the error function of a complex variable, Math. Comp. **19**, 127–129 (1965)
28. Turing, A. M.: A method for the calculation of the zeta-function, Proc. London Math. Soc. **s2-48**, 180–197 (1945)
29. Weideman, J. A. C.: Computation of the complex error function, SIAM J. Numer. Anal. **5**, 1497–1518 (1994)

A Appendix

In this appendix we prove the bounds

$$\frac{1}{2} \geq |F(x)| \geq \frac{1}{2\left|\mathrm{e}^{\mathrm{i}\pi/4} + \sqrt{\pi}\,x\right|} = \frac{1}{2\sqrt{1+\sqrt{2\pi}\,x+\pi x^2}} \geq \frac{1}{2+2\sqrt{\pi}\,x}, \qquad (76)$$

for $x \geq 0$. The lower bounds in (76), which appear to be new, are used in §2 to prove an upper bound on the relative error in the approximation $F_N(x)$ to $F(x)$. From (76) and (16) we immediately deduce bounds for negative arguments which are also used in §2, that

$$\frac{3}{2} \geq |F(-x)| \geq \frac{1}{2}, \quad \text{for } x \geq 0. \qquad (77)$$

We can also immediately deduce bounds on the version of the Fresnel integral defined by (4), and on the complementary error function of argument $\pm\pi/4$ (via (20) and that $\overline{\mathrm{erfc}(z)} = \mathrm{erfc}(\bar{z})$), for example that

$$|\mathcal{F}(x)| \geq \frac{1}{\sqrt{2}+\pi x} \quad \text{and} \quad \left|\mathrm{erfc}\left(\mathrm{e}^{\pm\mathrm{i}\pi/4}x\right)\right| \geq \frac{1}{1+\sqrt{\pi}\,x}, \quad \text{for } x \geq 0. \qquad (78)$$

The remainder of this appendix is the proof of (76). But note first that both

$$L_1(x) := \frac{1}{2\sqrt{1+\sqrt{2\pi}\,x+\pi x^2}} \quad \text{and} \quad L_2(x) := \frac{1}{2+2\sqrt{\pi}\,x} \qquad (79)$$

are sharp lower bounds for $|F(x)|$ for $x = 0$ (since $|F(0)| = L_1(0) = L_2(0) = \frac{1}{2}$) and in the limit $x \to +\infty$ (since $|F(x)| \sim (2\sqrt{\pi}x)^{-1}$ as $x \to +\infty$, and $L_1(x)$ and $L_2(x)$ have the same asymptotic behaviour). In fact, numerical computations (approximating $|F(x)|$ by $|F_N(x)|$ with $N = 12$) suggest that $1 \leq |F(x)|/L_1(x) < 1.25$ and $1 \leq F(x)/L_2(x) < 1.35$, for $x \geq 0$, so that these are rather sharp lower bounds over the whole positive real axis.

To prove (76) we note first of all that, from (22), on substituting $t = ux$,

$$|F(x)| = \frac{1}{2\pi}\left|\int_{-\infty}^{\infty} \frac{\mathrm{e}^{-x^2u^2}(1-\mathrm{i}u^2)}{1+u^4}\,\mathrm{d}u\right| = \frac{1}{2\pi}\,|g_0(x) - \mathrm{i}g_1(x)| \qquad (80)$$

where, for $n = 0, 1, ...,$

$$g_n(x) := \int_{-\infty}^{\infty} \frac{e^{-x^2u^2}\, u^{2n}}{1+u^4}\, du.$$

Clearly, $g_n(x) > 0$ is well-defined for all n and all $x > 0$ by this definition, and also for $x = 0$ for $n = 0, 1$, with (this computation done, *e.g.*, by contour integration) $g_0(0) = g_1(0) = \pi/\sqrt{2}$. Further, for $x > 0$ (and $x = 0$ for $n = 0$),

$$g_n'(x) = -2xg_{n+1}(x) < 0, \tag{81}$$

so that

$$g_n(x) < g_n(0) = \frac{\pi}{\sqrt{2}}, \quad \text{for } x > 0 \text{ and } n = 0, 1. \tag{82}$$

Using this last inequality in (80) gives $|F(x)| \leq \frac{1}{2}$, for $x \geq 0$. Moreover, (80) implies that

$$\begin{aligned} 2\left|e^{i\pi/4} + \sqrt{\pi}\, x\right|\, |F(x)| \geq G(x) &:= \frac{1}{\sqrt{2}\,\pi}\mathrm{Re}\left(\left(1+\sqrt{2\pi}\, x + i\right)(g_0(x) - ig_1(x))\right) \\ &= \frac{1}{\sqrt{2}\,\pi}\left((1+\sqrt{2\pi}\, x)g_0(x) + g_1(x)\right). \end{aligned} \tag{83}$$

Clearly, (76) will follow if we can show that $G(x) \geq 1$ for $x \geq 0$.

Now, for $x > 0$,

$$g_0(x) + g_2(x) = \int_{-\infty}^{\infty} e^{-x^2u^2}\, du = \frac{\sqrt{\pi}}{x}, \tag{84}$$

so that, for $x \geq 0$,

$$g_1'(x) = -2xg_2(x) = -2\sqrt{\pi} + 2xg_0(x)$$

and

$$g_0(x) = g_0(0) + \int_0^x g_0'(t)\, dt = \frac{\pi}{\sqrt{2}} - 2\int_0^x tg_1(t)\, dt, \tag{85}$$

$$g_1(x) = g_1(0) + \int_0^x g_1'(t)\, dt = \frac{\pi}{\sqrt{2}} - 2\sqrt{\pi}x + 2\int_0^x tg_0(t)\, dt. \tag{86}$$

From (82) we see that $g_0(x) \leq \pi/\sqrt{2}$, and then from (86) that

$$g_1(x) \leq \frac{\pi}{\sqrt{2}} - 2\sqrt{\pi}x + \frac{\pi}{\sqrt{2}}x^2.$$

It follows from (85) that

$$g_0(x) \geq \frac{\pi}{\sqrt{2}} - \frac{\pi}{\sqrt{2}}x^2 + \frac{4}{3}\sqrt{\pi}\, x^3 - \frac{\pi}{2\sqrt{2}}\, x^4, \tag{87}$$

and then from (86) that

$$g_1(x) \geq \frac{\pi}{\sqrt{2}} - 2\sqrt{\pi}x + \frac{\pi}{\sqrt{2}}x^2 - \frac{\pi}{2\sqrt{2}}\, x^4 + \frac{8}{15}\sqrt{\pi}\, x^5 - \frac{\pi\sqrt{2}}{12}\, x^6, \tag{88}$$

all these bounds holding for $x \geq 0$. Using these lower bounds in (83) we see that, for $x \geq 0$,

$$\begin{aligned} G(x) &\geq 1 + \frac{1}{\sqrt{2\pi}}(\pi - 2)x - \frac{1}{\sqrt{2\pi}}\left(\pi - \frac{4}{3}\right)x^3 + \frac{5}{6}x^4 - \sqrt{\frac{2}{\pi}}\left(\frac{\pi}{4} - \frac{4}{15}\right)x^5 - \frac{x^6}{12} \\ &= 1 + \frac{x}{\sqrt{2\pi}}\, h_0(x), \end{aligned}$$

where

$$h_0(x) = \pi - 2 - \left(\pi - \frac{4}{3}\right)x^2 + \frac{5\sqrt{2\pi}}{6}\,x^3 - \left(\frac{\pi}{2} - \frac{8}{15}\right)x^4 - \frac{\sqrt{2\pi}}{12}\,x^5.$$

We will show now that $h_0(x) > 0$ for $0 \leq x \leq 1$ which will show that $G(x) \geq 1$ for $0 \leq x \leq 1$. To see this we observe that $h_0'(x) = xh_1(x)$ where

$$\begin{aligned} h_1(x) &= -2\left(\pi - \frac{4}{3}\right) + \frac{5\sqrt{2\pi}}{2}\,x - \left(2\pi - \frac{32}{15}\right)x^2 - \frac{5\sqrt{2\pi}}{12}\,x^3 \\ &< -\frac{10}{3} + 7x - 4x^2 = -\left(2x - \frac{7}{4}\right)^2 + \frac{1}{15} - \frac{1}{3} < 0, \end{aligned}$$

so that $h_0'(x) < 0$ for $x > 0$. Thus, for $0 \leq x \leq 1$,

$$h_0(x) \geq h_0(1) = \frac{3\sqrt{2\pi}}{4} - \frac{\pi}{2} - \frac{2}{15} > 0.$$

We have shown that $G(x) \geq 1$ for $0 \leq x \leq 1$. It remains to show that $G(x) \geq 1$ for $x > 1$. To see this we make use of (83) and (84) which give that, for $x > 0$,

$$\begin{aligned} G(x) &> \frac{1}{\sqrt{2}\,\pi}\,(1 + \sqrt{2\pi}\,x)\left(\frac{\sqrt{\pi}}{x} - g_2(x)\right) \\ &= 1 + \frac{1}{\sqrt{2\pi}\,x}\left(1 - \left(\frac{1}{\sqrt{\pi}} + \sqrt{2}\,x\right)xg_2(x)\right). \end{aligned} \tag{89}$$

A simple upper bound on $g_2(x)$ is

$$xg_2(x) < x\int_{-\infty}^{\infty} \mathrm{e}^{-x^2u^2}\, u^4\,\mathrm{d}u = \frac{3\sqrt{\pi}}{4x^4}, \tag{90}$$

where to obtain this last value we integrate by parts twice and then use (84). Thus, for $x > 0$,

$$\sqrt{2\pi}\,x(G(x) - 1) > 1 - \frac{3}{4x^4} - \frac{3\sqrt{2\pi}}{4x^3}.$$

Thus $G(x) \geq 1$ if x is large enough, in particular if $x \geq \sqrt{2}$.

To show that $G(x) \geq 1$ for $1 < x < \sqrt{2}$ we need a sharper upper bound on $g_2(x)$ than (90). To obtain this upper bound we write

$$xg_2(x) = I := \int_{-\infty}^{\infty} \frac{\mathrm{e}^{-t^2}\, t^4}{x^4 + t^4}\, \mathrm{d}t,$$

and approximate this integral by the trapezium rule as

$$I_h = h \sum_{n=-\infty}^{\infty} \frac{\mathrm{e}^{-n^2h^2}\, n^4h^4}{x^4 + n^4h^4} = 2h \sum_{n=1}^{\infty} \frac{\mathrm{e}^{-n^2h^2}\, n^4h^4}{x^4 + n^4h^4}.$$

Arguing as in §2 (or see [24, §5.1.4]), the error in this trapezium rule approximation is $I - I_h = PC_h + E_h^*$, where PC_h, a pole contribution, and E_h^* are given by

$$PC_h = 2\pi\mathrm{i}(r_0 + r_1) \text{ and } E_h^* = \int_{\Gamma_H} f(z)(1 + g(z))\, \mathrm{d}z.$$

Here, $f(z) = \mathrm{e}^{-z^2} z^4/(x^4 + z^4)$, $g(z) = -\mathrm{i}\cot(\pi z/h)$, and, for $m = 0, 1$, r_m is the residue of $f(z)(1 + g(z))$ at the pole z_m, with $z_0 = x\mathrm{e}^{\mathrm{i}\pi/4}$ and $z_1 = \mathrm{i}z_0$. (The difference in definition of g compared to §2 is because the trapezium rule here is shifted by $h/2$ compared to §2: note that $g(z + h/2) = \mathrm{i}\tan(\pi z/h)$ as in §2.) Further, Γ_H is the line $\Im z = H$ as in §2, and we assume that $H > x/\sqrt{2}$, so that the poles at z_0 and z_1 lie between Γ_H and the real axis.

Arguing as in §2, in particular using the bound (30), we see that

$$|E_h^*| \leq \sup_{z \in \Gamma_H} \left| \frac{z^4(1 + g(z))}{z^4 + x^4} \right| \int_{-\infty}^{\infty} \mathrm{e}^{H^2 - t^2} \mathrm{d}t \leq \frac{2\sqrt{\pi}\, H^4\, \mathrm{e}^{H^2 - 2\pi H/h}}{(H^4 - x^4)(1 - \mathrm{e}^{-2\pi H/h})},$$

provided that $H > x$. Further,

$$r_0 = \frac{\mathrm{e}^{-z_0^2} z_0^3}{2(z_0^2 + \mathrm{i}x^2)}(1 + g(z_0)) = \frac{1}{4}\mathrm{e}^{\mathrm{i}(\pi/4 - x^2)} x(1 + g(z_0)),$$

so that $|r_0| = x|1 + g(z_0)|/4$. Similarly, $|r_1| = x|1 + g(z_1)|/4$, so that, using (30),

$$|PC_h| \leq \frac{2\pi x \mathrm{e}^{-\sqrt{2}\,\pi x/h}}{1 - \mathrm{e}^{-\sqrt{2}\,\pi x/h}}.$$

Finally, using (45),

$$2h \sum_{n=N+1}^{\infty} \frac{\mathrm{e}^{-n^2h^2}\, n^4h^4}{x^4 + n^4h^4} < 2h \sum_{n=N+1}^{\infty} \mathrm{e}^{-n^2h^2} < 2\int_{Nh}^{\infty} \mathrm{e}^{-t^2}\, dt < \frac{\mathrm{e}^{-N^2h^2}}{Nh}.$$

Thus

$$\begin{aligned} I &\leq I_h + |PC_h| + |E_h^*| \\ &\leq 2h \sum_{n=1}^{N} \frac{\mathrm{e}^{-n^2h^2}\, n^4h^4}{x^4 + n^4h^4} + \frac{\mathrm{e}^{-N^2h^2}}{Nh} + \frac{2\pi x \mathrm{e}^{-\sqrt{2}\,\pi x/h}}{1 - \mathrm{e}^{-\sqrt{2}\,\pi x/h}} + \frac{2\sqrt{\pi}\, H^4\, \mathrm{e}^{H^2 - 2\pi H/h}}{(H^4 - x^4)(1 - \mathrm{e}^{-2\pi H/h})}, \end{aligned}$$

provided that $H > x$. In particular, for $1 < x < \sqrt{2}$, choosing $h = 1$, $H = \pi$, and $N = 2$, and noting that $x\mathrm{e}^{-\sqrt{2}\,\pi x}$ is decreasing on $x > 1$, we see that

$$\begin{aligned} xg_2(x) &< \frac{2}{\mathrm{e}(x^4+1)} + \frac{32}{\mathrm{e}^4(x^4+16)} + \frac{1}{2\mathrm{e}^4} + \frac{2\pi x\mathrm{e}^{-\sqrt{2}\,\pi x}}{1-\mathrm{e}^{-\sqrt{2}\,\pi x}} + \frac{2\sqrt{\pi}\,\pi^4\,\mathrm{e}^{-\pi^2}}{(\pi^4-x^4)(1-\mathrm{e}^{-2\pi^2})} \\ &< \frac{2}{\mathrm{e}(x^4+1)} + \delta \end{aligned}$$

where

$$\delta := \frac{32}{17\mathrm{e}^4} + \frac{1}{2\mathrm{e}^4} + \frac{2\pi\mathrm{e}^{-\sqrt{2}\,\pi}}{1-\mathrm{e}^{-\sqrt{2}\,\pi}} + \frac{2\sqrt{\pi}\,\pi^4\,\mathrm{e}^{-\pi^2}}{(\pi^4-4)(1-\mathrm{e}^{-2\pi^2})} \approx 0.119.$$

Thus

$$\left(\frac{1}{\sqrt{\pi}} + \sqrt{2}\,x\right) xg_2(x) < \mathcal{H}(x) := \left(\frac{1}{\sqrt{\pi}} + \sqrt{2}\,x\right)\left(\frac{2}{\mathrm{e}(x^4+1)} + \delta\right).$$

Now, for $1 < x < \sqrt{2}$, since $x^4/(1+x^4)^2$ is decreasing on $x > 1$,

$$\begin{aligned} \mathcal{H}'(x) &= \sqrt{2}\,\delta + \frac{2}{\sqrt{\pi}\,\mathrm{e}}\,\frac{\sqrt{2\pi} - 3\sqrt{2\pi}x^4 - 4x^3}{(1+x^4)^2} \\ &< \sqrt{2}\,\delta - \frac{2}{\sqrt{\pi}\,\mathrm{e}}\,\frac{3\sqrt{2\pi}x^4}{(1+x^4)^2} < \sqrt{2}\left(\delta - \frac{24}{25\mathrm{e}}\right) < 0. \end{aligned}$$

Thus, for $1 < x < \sqrt{2}$, it follows from (89) that

$$\begin{aligned} \sqrt{2\pi}\,x(G(x)-1) &> 1 - \left(\frac{1}{\sqrt{\pi}} + \sqrt{2}\,x\right) xg_2(x) > 1 - \mathcal{H}(x) \\ &> 1 - \mathcal{H}(1) = 1 - \left(\frac{1}{\sqrt{\pi}} + \sqrt{2}\right)\left(\mathrm{e}^{-1} + \delta\right) > 0. \end{aligned}$$

Thus $G(x) > 1$ for $1 < x < \sqrt{2}$ and the proof is complete.