

# Theory and Applications of an Orthogonal Rational Basis Set<sup>1</sup>

J.A.C. Weideman

Department of Mathematics  
University of Stellenbosch

Stellenbosch 7600  
South Africa<sup>2</sup>

**Abstract.** This paper reviews the theory and applications of the functions  $\phi_n(x) = (1+ix)^n/(1-ix)^{n+1}$ , which form a complete and orthogonal basis set for  $L_2(\mathbb{R})$ . It is shown how these functions may be used for the computation of integral transforms and certain special functions. The numerical solution of some differential equations is also discussed.

**1. Introduction.** Many techniques in scientific computing are based on series expansions of the form

$$f(x) = \sum_n a_n \phi_n(x), \quad x \in D. \quad (1)$$

The basis set  $\{\phi_n\}$  is typically complete and orthogonal in some appropriate space defined on the domain  $D$ . For example, if  $D$  is a finite interval Chebyshev or Legendre expansions are commonly used. If periodicity of  $f(x)$  may be assumed a Fourier series is more appropriate. For the semi-infinite interval  $D = [0, \infty)$  the Laguerre functions are candidates, and on the real line  $D = (-\infty, \infty)$  the Hermite or sinc (cardinal) functions may be used. There exists, however, a lesser-known basis set which may be used as alternative to Hermite and sinc functions on the real line. This basis set has some attractive features which we would like to review in this paper.

The basis set under consideration consists of rational functions, defined for all  $x \in \mathbb{R}$  by

$$\phi_n(x) = \frac{(1+ix)^n}{(1-ix)^{n+1}}, \quad n \in \mathbb{Z}, \quad i^2 = -1. \quad (2)$$

These functions are complete and orthogonal in  $L_2(\mathbb{R})$ ; see [10] for a proof of completeness and Sect. 2 for a proof of orthogonality. In Fig. 1 we show a few of the functions  $\phi_n(x)$ . Observe that the functions are oscillatory, with an increase in frequency as  $|n|$  increases. The oscillations are modulated by the envelope  $1/\sqrt{1+x^2}$ , as indicated by the dashed curve.

As for the history of the functions (2), they could be traced to Wiener's book on time series, published in 1947 [21]. The use of these functions for computational

---

<sup>1</sup>To appear in the Proceedings of the South African Numerical Mathematics Symposium 1994, University of Natal, Durban, South Africa.

<sup>2</sup>Permanent address: Department of Mathematics, Oregon State University, Corvallis 97331, USA.

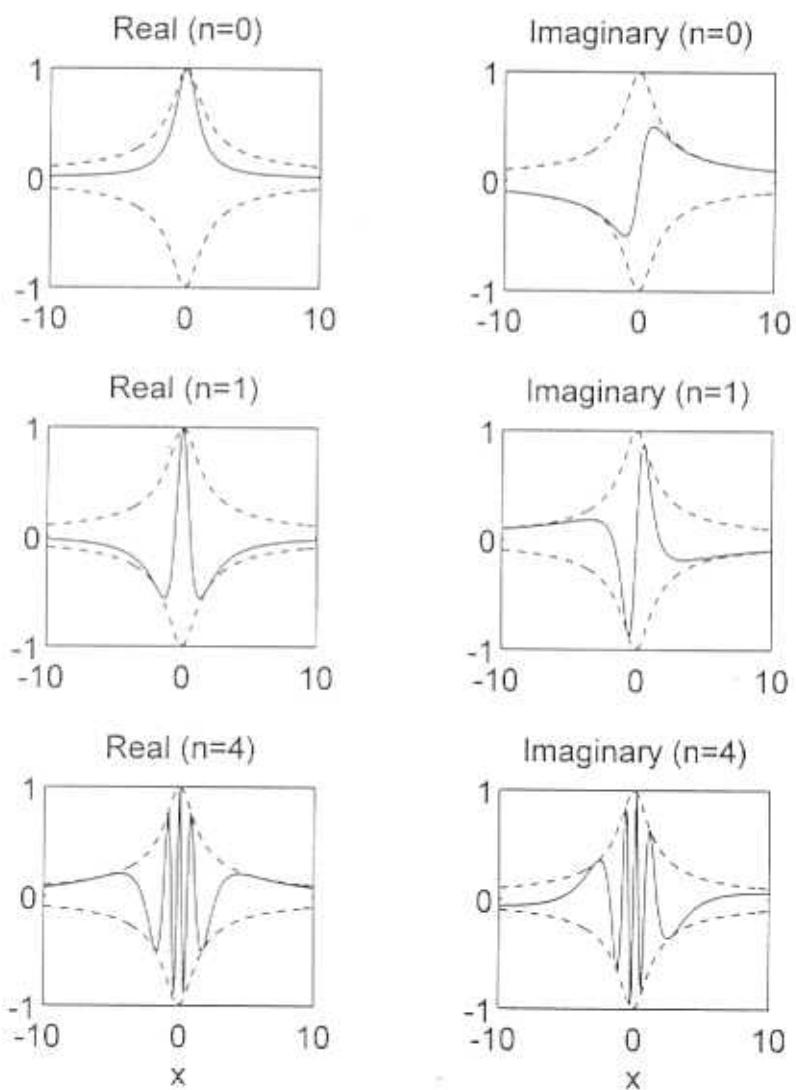


Figure 1: The basis functions  $\phi_n(x)$ .

purposes dates to the 1980s, and it is associated with the names of Christov [5,6] and Boyd [2,3,4]. In addition to the work of these two authors the present author has made some contributions in this area. This paper is a summary of these results, collected from the papers [11,17,18,19,20].

The survey is organized as follows: In Sect. 2 we review implementation details and show how the Fast Fourier Transform (FFT) may be exploited for computational purposes. The computation of integral transforms is surveyed in Sect. 3. We focus on the author's method for the computation of the Hilbert transform [19]. However, we also show the connection with two other methods: Weber's method for the computation of the Fourier transform [15] and Weeks' method for inversion of the Laplace transform [16]. We believe the interconnection between these three methods is illuminated here for the first time. In Sect. 4 we discuss a series expansion for the computation of the complex error function. This expansion derives from the fact that the complex error function may be expressed as the Hilbert transform of the Gaussian  $\exp(-x^2)$ . In Sect. 5 we discuss the solution of a few differential equations, as well as some theoretical results involving the eigenvalues of differentiation operators based on (2).

**2. Implementation.** One attractive property of the basis set (2) is the fact that it can be implemented via the FFT. This gives it an edge over a competitor like Hermite expansions for which no practical fast transform is known to exist. The key is the coordinate transformation

$$e^{ix} = \frac{1+ix}{1-ix} \iff x = \tan \frac{1}{2}\theta$$

which maps the infinite line to a finite interval:

$$x \in [-\infty, \infty] \iff \theta \in [-\pi, \pi].$$

Thus the original series may be viewed as a Fourier series in the new variable  $\theta$ :

$$f(x) = \sum_{n=-\infty}^{\infty} a_n \phi_n(x) \iff f(x)(1-ix) = \sum_{n=-\infty}^{\infty} a_n e^{in\theta}.$$

This connection with Fourier series allows one to prove the orthogonality relation

$$\int_{-\infty}^{\infty} \phi_n(x) \overline{\phi_m(x)} dx = \pi \delta_{n,m}$$

where the overbar denotes complex conjugation, and  $\delta_{n,m}$  is the Kronecker delta. The orthogonality relation provides an integral formula for the expansion coefficients:

$$a_n = \frac{1}{\pi} \int_{-\infty}^{\infty} f(x) \overline{\phi_n(x)} dx. \quad (3)$$

For computational work the series is truncated to<sup>3</sup>

$$f(x) = \sum_{n=-N}^{N-1} a_n \phi_n(x), \quad \text{or} \quad f(x)(1-ix) = \sum_{n=-N}^{N-1} a_n e^{in\theta}. \quad (4)$$

---

<sup>3</sup>The term corresponding to  $n = N$  is omitted to ensure that all terms in the series appear as conjugate pairs.

For a given  $f(x)$  one needs to compute the expansion coefficients  $a_n$ . The integral formula (3) may be used, but the most practical procedure is that of *collocation*. We introduce the points

$$\theta_j = \frac{\pi j}{N}, \quad x_j = \tan \frac{1}{2} \theta_j, \quad j = -N, \dots, N-1,$$

and require that equality in (4) is achieved at each of these points. This yields the linear system

$$f(x_j)(1 - ix_j) = \sum_{n=-N}^{N-1} a_n e^{in\theta_j}, \quad j = -N, \dots, N-1. \quad (5)$$

Note that the point at infinity is included as a collocation point. Typically the function satisfies  $f(x) = o(|x|^{-1})$  as  $x \rightarrow \pm\infty$ , in which case the left-hand side is set to zero when  $j = -N$ . The series on the right is recognized to be nothing but a discrete Fourier transform. One may therefore solve for the  $a_n$  by taking the inverse transform, and this may be achieved with the FFT.

In summary, if the function  $f(x)$  is known, the coefficients  $a_n$  can be computed from (5) with an inverse FFT. On the other hand, if the coefficients  $a_n$  are given, the function  $f(x)$  may be evaluated at the collocation points  $x_j$  with the forward FFT. Both of these processes require  $O(N \log N)$  operations as opposed to the  $O(N^2)$  required by a direct summation of (5).

**3. Computation of Integral Transforms.** Many orthogonal functions are eigenfunctions of differential operators. For example, the relation  $\frac{d}{dx} e^{ikx} = ike^{ikx}$  shows that the Fourier polynomials  $\{e^{ikx}\}$  are eigenfunctions of the derivative operators  $\frac{d}{dx}, \frac{d^2}{dx^2}$ , etc. Chebyshev, Legendre, and Hermite functions may likewise be viewed as the eigenfunctions of certain linear differential operators. The question arises whether the rational functions (2) can be associated with some well-known linear operator in this manner. The answer is yes, and the operator in question is the Hilbert transform in  $L_2(\mathbb{R})$ . This transform is defined by the singular integral

$$F_{\mathcal{H}}(y) = \frac{1}{\pi} PV \int_{-\infty}^{\infty} \frac{f(x)}{x-y} dx,$$

where  $PV$  denotes the principal value. A quick residue calculation proves that the functions  $\{\phi_n\}$  are eigenfunctions of this operator, with eigenvalues  $\pm i$ . Details may be found in [19].

**Theorem 1** *For any  $n \in \mathbb{Z}$*

$$f(x) = \phi_n(x) \implies F_{\mathcal{H}}(y) = i \operatorname{sgn}(n) \phi_n(y).$$

*Note: we define  $\operatorname{sgn}(0) = 1$ .*

This theorem suggests a practical procedure for computing the Hilbert transform:

$$f(x) = \sum_{n=-\infty}^{\infty} a_n \phi_n(x) \implies F_{\mathcal{H}}(y) = \sum_{n=-\infty}^{\infty} i \operatorname{sgn}(n) a_n \phi_n(y). \quad (6)$$

As discussed in Sect. 2, these series are truncated to  $-N \leq n \leq N - 1$ , and the coefficients  $a_n$  are computed from (5) via the inverse FFT. This algorithm has two errors associated with it: first, the truncation of the infinite series, and second, the computation of the  $a_n$  via the FFT rather than the integral formula (3). These errors can be estimated separately, and this yields the following error bound, taken from [19]:

**Theorem 2** *Let  $F_{\mathcal{H}}(y)$  be as defined in (6), and let  $F_{\mathcal{H}}(y; N)$  be its finite truncation ( $-N \leq n \leq N - 1$ ), with the expansion coefficients  $a_n$  computed via (5). Assume that  $\sum_{n=-\infty}^{\infty} |a_n| < \infty$ . Then, for each real  $y$ ,*

$$|F_{\mathcal{H}}(y) - F_{\mathcal{H}}(y; N)| \leq \frac{4}{\sqrt{1+y^2}} \sum_{n=N}^{\infty} |a_n|.$$

This error bound confirms the intuitive notion that if the coefficients  $a_n$  decrease rapidly as  $n \rightarrow \pm\infty$ , the method is accurate, even with a small number of terms. It is however not possible to give a complete classification of the rate of decay in the  $a_n$  in terms of elementary properties of  $f(x)$ , but a few model functions have been analyzed in [19]. The results are as follows:<sup>4</sup>

$f(x)$	$a_n$
$1/(1+x^4)$	$O(r^n), r = \sqrt{2}-1$
$\exp(-x^2)$	$O(\exp(-\frac{3}{2}n^{2/3}))$
$\operatorname{sech}(x)$	$O(\exp(-2n^{1/2}))$
$\sin(x)/(1+x^2)$	$O(n^{-5/4})$
$\sin(x)/(1+x^4)$	$O(n^{-9/4})$
$\exp(- x )$	$O(n^{-2})$

The convergence rate is best if  $f(x)$  is a rational function of the form  $p(x)/q(x)$ , with  $\deg(q) > \deg(p)$ , and  $q(x)$  has no real zeros. Then it can be shown that the coefficients  $a_n$  decrease at least geometrically, i.e.,  $a_n = O(r^n)$  for some  $0 < r < 1$ . If, on the other hand,  $f(x)$  is infinitely differentiable and decays at least exponentially fast as  $x \rightarrow \pm\infty$  then integration by parts can be used to show that  $a_n = O(n^{-\ell})$  for each positive integer  $\ell$ . This is consistent with the second and third entries of the table, which were obtained by applying a steepest descent analysis to the integral (3). When the function is oscillatory as  $x \rightarrow \pm\infty$ , or if there exists a discontinuity in the derivative, the convergence is slowest of all as the last three entries in the table show.

<sup>4</sup>All the coefficients in the table satisfy  $|a_n| = |a_{-1-n}|$ ; therefore we consider only  $n > 0$ .

In order to improve the convergence rate it is recommended that one uses a rescaled basis set defined by

$$\tilde{\phi}_n(x) = \phi_n(x/p).$$

Here  $p$  is a real parameter that may be adjusted to maximize the accuracy. This rescaling stretches the collocation points to  $x_j = p \tan \frac{1}{2}\theta_j$ , but the implementation details discussed in Sect. 2, as well as Theorems 1 and 2, remain valid with only minor modifications.

To see how the accuracy depends on the parameter  $p$ , we consider the test problem

$$f(x) = \operatorname{sech}(x), \quad F_{\mathcal{H}}(x) = \tanh(x) + \frac{i}{\pi} \left[ \psi\left(\frac{1}{4} + \frac{ix}{2\pi}\right) - \psi\left(\frac{1}{4} - \frac{ix}{2\pi}\right) \right]$$

where  $\psi(z) = \Gamma'(z)/\Gamma(z)$  is the digamma function. Fig. 2 shows the error, defined by

$$\text{error} = \max_{|j| \leq N} |F_{\mathcal{H}}(x_j) - \mathcal{F}_{\mathcal{H}}(x_j; N)|, \quad (7)$$

as a function of  $p$ . With  $N = 16, 32$ , and  $64$  the optimal value of  $p$  appears to be  $3, 4$ , and  $5$ , roughly, and the best error improves from around  $10^{-6}$  to  $10^{-10}$  to  $10^{-15}$ . Even if one picks  $p$  non-optimally the accuracy can still be quite good. For example, fixing  $p$  at the value  $10$  the error improves from around  $10^{-2}$  to  $10^{-4}$  to  $10^{-9}$ . Unfortunately, finding the optimal relationship between  $p$  and  $N$  requires a non-trivial asymptotic estimate. We shall elaborate on this issue below. More comprehensive numerical tests, including a comparison with other methods for computing Hilbert transforms, may be found in [19].

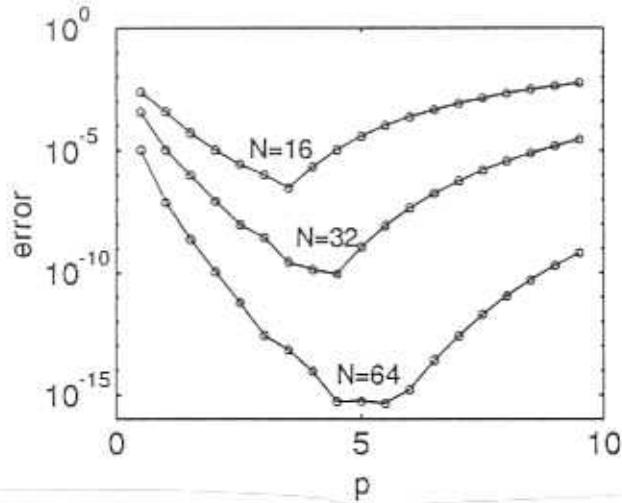


Figure 2: The error in the numerical computation of the Hilbert transform of  $\operatorname{sech}(x)$  as a function of  $p$  and  $N$ .

After working out the details of this method for computing Hilbert transforms the author discovered that it is in fact closely related to two other methods: Weeks'

method for the inversion of the Laplace transform [16], and Weber's method for Fourier transforms [15]. We believe the fact that all three of these methods are based on expansions involving the rational basis set (2) is pointed out here for the first time.

The Fourier and Laplace transforms are respectively defined by

$$F_{\mathcal{F}}(t) = \int_{-\infty}^{\infty} e^{itx} f(x) dx, \quad F_{\mathcal{L}}(t) = \int_0^{\infty} e^{-tx} f(x) dx, \quad t \in \mathbb{R}. \quad (8)$$

Both of these transforms may be computed by expanding  $f(x)$  in terms of  $\{\phi_n\}$ , followed by termwise integration. Before sketching the details, however, we choose to modify the basis set (2) temporarily to

$$\tilde{\phi}_n(x) = (-1)^n p^{-1} \phi_n(x/p).$$

This modification facilitates the comparison with the expansions found in [15,16]. It means we are now looking at expansions of the form

$$f(x) = \sum_{n=-\infty}^{\infty} a_n \frac{1}{ix + p} \left( \frac{ix - p}{ix + p} \right)^n. \quad (9)$$

For the computation of the Fourier transform we insert this series into the first integral in (8). Assuming that summation and integration can be interchanged, one obtains

$$F_{\mathcal{F}}(t) = \sum_{n=-\infty}^{\infty} a_n \psi_n(t), \quad (10)$$

where

$$\psi_n(t) = \int_{-\infty}^{\infty} \frac{e^{itx}}{ix + p} \left( \frac{ix - p}{ix + p} \right)^n dx.$$

It is a straightforward exercise to evaluate this integral using residues, with contours involving the usual semi-circles in the upper or lower half-planes, depending on the sign of  $t$ . The result is

$$\psi_n(t) = \begin{cases} 2\pi e^{-pt} L_n(2pt) & t \geq 0 \text{ and } n \geq 0 \\ 2\pi e^{pt} L_{-n-1}(-2pt) & t < 0 \text{ and } n < 0 \\ 0 & \text{otherwise,} \end{cases}$$

where  $L_n(x)$  represents the Laguerre polynomial of degree  $n$ .<sup>5</sup> This is identical to the formula in [15], but derived via a different route.

The algorithm proceeds as follows: Given  $f(x)$ , the expansion coefficients can be computed with the FFT from an expression similar to (5). This is followed by an evaluation of the Laguerre series (10). The numerical experiments reported in [16] have shown this procedure to be highly accurate. An indication of the high accuracy is provided by the model example  $f(x) = 1/(1+x^2)$ ,  $F_{\mathcal{F}}(t) = 2\pi e^{-|t|}$ : with  $p = 1$

<sup>5</sup>The Laguerre polynomials, defined by  $L_n(x) = \frac{1}{n!} e^x \frac{d^n}{dx^n} [e^{-x} x^n]$ , enter into the calculation in the evaluation of the residues at the  $n$ -th order poles.

the algorithm gives the exact result with only two terms in the series (10). Some additional unpublished numerical results, corresponding to the test example

$$f(x) = e^{-x^2}, \quad F_{\mathcal{F}}(t) = \sqrt{\pi} e^{-t^2/4},$$

are shown in Fig. 3. The figure shows level curves of the error, defined as in (7), as a function of  $p$  and  $N$ . The labels on the curves indicate the  $\log_{10}$  of the error. The dashed curve represents  $p = 2^{-\frac{1}{4}}\sqrt{N}$ , which was shown to be the optimal relationship between  $p$  and  $N$  for functions of the form  $\exp(-x^2) \times \text{polynomial}(x)$ ; see [18].

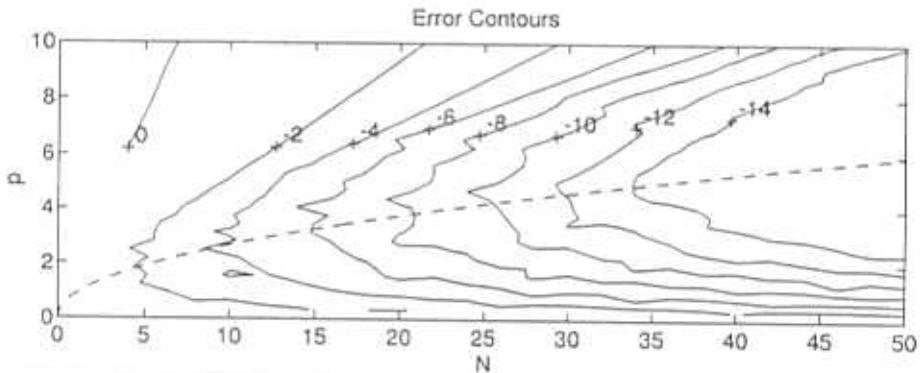


Figure 3: Level curves of the error in the numerical computation of the Fourier transform of  $\exp(-x^2)$ .

Turning to the inversion of the Laplace transform, this involves the computation of the Bromwich integral

$$f(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{zx} F_{\mathcal{L}}(z) dz, \quad x > 0,$$

where  $c > c_0$  (the Laplace convergence abscissa). Next, expand  $F_{\mathcal{L}}(z)$  on the line  $\text{Re}(z) = c$  as a rational series of the form (9):

$$F_{\mathcal{L}}(c+it) = \sum_{n=-\infty}^{\infty} a_n \frac{1}{it+p} \left( \frac{it-p}{it+p} \right)^p,$$

Inserting this expression into the Bromwich integral leads to a residue calculation similar to the one for the Fourier transform alluded to above. The result is

$$f(x) = e^{cx} \sum_{n=0}^{\infty} a_n e^{-px} L_n(2px),$$

which represents the celebrated method of Weeks, often quoted as one of the most efficient methods for inverting the Laplace transform. For more details of the Weeks method we refer to [12,16], and for comparisons with other algorithms we refer to [7,8].

**4. Computation of the Complex Error Function.** The complex error function, which arises routinely in astrophysics, is defined by [9,14]:

$$w(z) = e^{-z^2} \operatorname{erfc}(-iz), \quad \text{where } \operatorname{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^\infty e^{-t^2} dt, \quad z \in \mathbb{C}.$$

Important special cases include

$$w(ix) = e^{-x^2} \operatorname{erfc}(x), \quad \text{and } \operatorname{Im}(w(x)) = \frac{2}{\sqrt{\pi}} e^{-x^2} \int_0^x e^{t^2} dt, \quad x \in \mathbb{R},$$

the latter of which is referred to as Dawson's integral.

The function  $w(z)$  may also be represented by the integral

$$w(z) = \frac{i}{\pi} \int_{-\infty}^{\infty} \frac{e^{-t^2}}{z-t} dt, \quad (11)$$

valid for each  $z$  in the upper half-plane. We expand  $e^{-t^2}$  as a weighted series in  $\{\phi_n\}$ , followed by termwise integration. This yield a series expansion for  $w(z)$ , with some attractive computational properties. The details are as follows.

We consider the expansion

$$(p^2 + t^2)e^{-t^2} = \sum_{n=-\infty}^{\infty} a_n \left( \frac{p+it}{p-it} \right)^n, \quad (12)$$

with the coefficients given by

$$a_n = \frac{p}{\pi} \int_{-\infty}^{\infty} \frac{f(t)}{p^2 + t^2} \left( \frac{p-it}{p+it} \right)^n dt. \quad (13)$$

Next we insert (12) into (11). The validity of interchanging summation and integration was justified in [18], and thus

$$w(z) = \sum_{n=-\infty}^{\infty} a_n \psi_n(z),$$

where

$$\psi_n(z) = \frac{i}{\pi} \int_{-\infty}^{\infty} \frac{1}{p^2 + t^2} \left( \frac{p+it}{p-it} \right)^n \frac{1}{z-t} dt.$$

A residue calculation yields

$$\psi_n(z) = \begin{cases} \frac{1}{p} \frac{1}{p-iz} \left( \frac{p+iz}{p-iz} \right)^n & n = 0 \\ \frac{2}{p^2 + z^2} \left( \frac{p+iz}{p-iz} \right)^n & n > 0 \\ 0 & n < 0, \end{cases}$$

and so

$$w(z) = \frac{1}{\sqrt{\pi}(p-iz)} + \frac{2}{p^2 + z^2} \sum_{n=1}^{\infty} a_n \left( \frac{p+iz}{p-iz} \right)^n, \quad \operatorname{Im}(z) > 0. \quad (14)$$

As before the series (14) is truncated at  $n = N$ , and again the question is how to select a  $p$  for any given  $N$  such that the accuracy is optimized. In particular, we would like to maximize the rate decay over the coefficients retained in the expansion, namely  $a_0, \dots, a_N$ . One way to achieve this is to maximize the rate of decay in  $a_N$ , the last coefficient retained in the expansion, as  $N \rightarrow \infty$ . The details of this calculation are presented in [18], and the optimal relationship was shown to be  $p = 2^{-\frac{1}{4}}\sqrt{N}$ . (We have verified the validity of this estimate in a different context in Fig. 3.) Assuming this relationship between  $p$  and  $N$ , it was shown in [18] that

$$a_N = O(r^N), \quad \text{with } r = \sqrt{2} - 1, \quad (15)$$

suggesting high accuracy when (14) is truncated at  $n = N$ .

The algorithm for computing  $w(z)$  proceeds as follows: For a given  $N$  and  $M \gg N$  the coefficients  $a_n$  are computed with the aid of the FFT from

$$a_n = \frac{1}{2M} \sum_{j=-M+1}^{M-1} (p^2 + t_j^2) e^{-t_j^2} e^{-in\theta_j}, \quad j = 1, \dots, N,$$

where  $p = 2^{-\frac{1}{4}}\sqrt{N}$ , and

$$t_j = p \tan \frac{1}{2} \theta_j, \quad \theta_j = \frac{\pi j}{M}, \quad j = -M + 1, \dots, M - 1.$$

This summation formula is of course the trapezoidal rule approximation of the integral (13) with  $2M$  gridpoints. Once the coefficients have been calculated,  $w(z)$  may be computed from (14) by evaluating a polynomial of degree  $N$  in the variable  $(p + iz)/(p - iz)$ .

Two comments on this algorithm are in order: First, for any given  $N$  the coefficients  $a_n$  may be computed once and for all. Second, we recommend using  $M = 2N$  in practice. This ensures that the error in the trapezoidal rule approximation of the integral is on the same order of magnitude as the truncation error in the series.

As a numerical test of this method we offer Fig. 4, taken from [18]. We have computed the values of  $w(z)$  for  $z = 10^\ell e^{i\theta}$ , for several thousand values lying in the range  $-6 \leq \ell \leq 6$ ,  $0 \leq \theta \leq \pi/2$ . The relative error in each approximation was computed using the TOMS Algorithm 680, which computes  $w(z)$  accurately to at least 14 significant digits [13]. In Fig. 4 the level curves of the relative error are shown as a function of  $\ell$  and  $\theta$ . The label on each curve represents the  $\log_{10}$  of the relative error and is therefore an indication of the number of correct digits in the approximation.

An inspection of Fig. 4 reveals the following: There is a region near the imaginary axis where the accuracy is exceptionally high. The accuracy also improves significantly as  $|z| \rightarrow \infty$ . If one retains an additional eight terms in the series an extra three digits of accuracy, on average, are gained. This is consistent with the estimate (15).

The fact that the expansion (14) gives reasonably high accuracy more-or-less uniformly in the first quadrant gives it an edge over other expansions. A Taylor

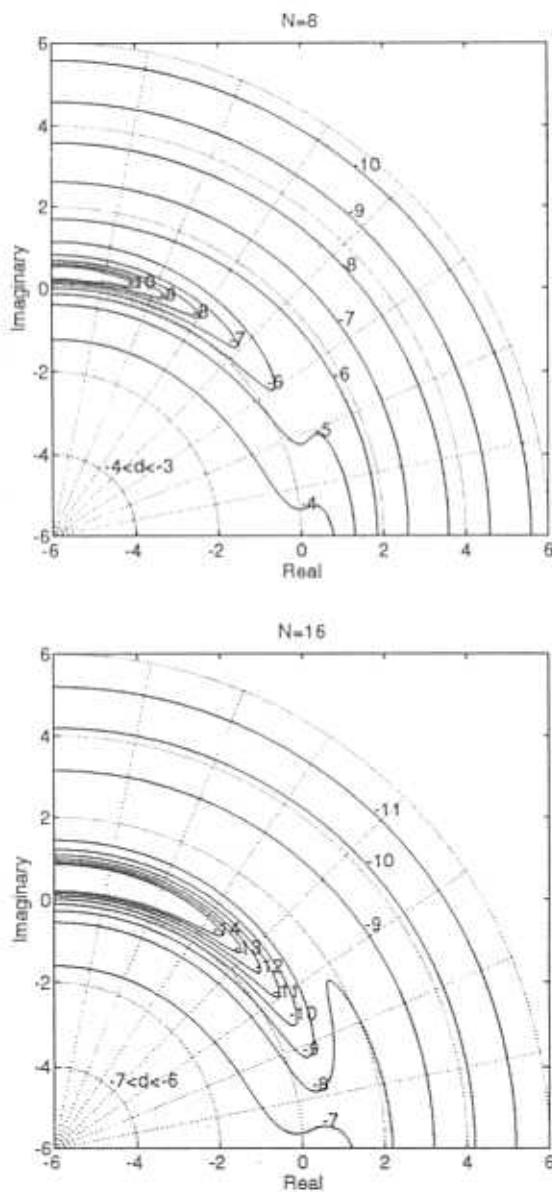


Figure 4: Level curves of the relative error in the series expansion of  $w(z)$  when truncated at  $n = N$ . The label on the coordinate axes refers to  $\ell$ , where  $z = 10^\ell e^{i\theta}$ . The label on the level curves refers to the  $\log_{10}$  of the relative error.

series does well near the origin, but breaks down for large  $|z|$ . Continued fractions, by contrast, yield high accuracy for large  $|z|$  but become inefficient near  $z = 0$ . Algorithm 680, which was used to compute the relative error in Fig. 4, consists of a combination of several such expansions tailored for various regions in the complex plane. Our method, being based on a single expansion, is therefore more efficient in vectorized implementations such as the Matlab code given in [18].

The expansion (14) not only represents an efficient way of computing  $w(z)$ , but also of its integral  $\int_0^z w(\zeta) d\zeta$ . In particular, the antiderivative relation

$$-\frac{i}{2pn} \frac{d}{dz} \left( \frac{p+iz}{p-iz} \right)^n = \frac{1}{p^2 + z^2} \left( \frac{p+iz}{p-iz} \right)^n$$

can be invoked to integrate (14) term-by-term:

$$\int_0^z w(\zeta) d\zeta = \frac{i}{\sqrt{\pi}} \log \left( \frac{p-iz}{p} \right) + \frac{i}{p} g(p) - \frac{i}{p} \sum_{n=1}^{\infty} \frac{a_n}{n} \left( \frac{p+iz}{p-iz} \right)^n, \quad \text{Im}(z) > 0, \quad (16)$$

where

$$g(p) = \sum_{n=1}^{\infty} \frac{a_n}{n} = p \left( \int_0^p e^{y^2} \operatorname{erfc}(y) dy - \frac{1}{\sqrt{\pi}} \log(2) \right).$$

Using the antiderivative to compute the integral  $\int_0^z w(\zeta) d\zeta$  is justified by the fact that  $w(z)$  defines an analytic function in the upper half-plane. A table of values of  $g(p)$  for various values of  $N$  (with  $p = 2^{-\frac{1}{4}}\sqrt{N}$ ) may be found in [20].

Also represented in [20] are numerical tests on the following two important integrals which are special cases of (16):

$$I_1(r) = \int_0^r e^{-x^2} \int_0^x e^{t^2} dt dx, \quad I_2(r) = \int_0^r e^{y^2} \int_y^{\infty} e^{-t^2} dt dy.$$

In [14] these two integrals were tabulated for  $0 \leq r \leq 6$  to ten digits of accuracy. Our expansion was found to yield identical results with only 24 terms in the series.

We conclude this section with a caveat: the expansion (16) may be numerically unstable for small  $|z|$ . A remedy is suggested in [20].

**5. Solution of Differential Equations.** To solve differential equations with the basis set (2) we need to construct discrete differential operators. These operators are obtained via

$$f^{(k)}(x) = \sum_n a_n \phi_n^{(k)}(x) = \sum_n a_n^{(k)} \phi_n(x), \quad k = 1, 2, \dots,$$

which defines a relationship:

$$a_n^{(k)} = D_k a_n.$$

To construct  $D_1$  we apply logarithmic differentiation to (2):

$$\frac{d\phi_n}{dx} = \phi_n \left[ \frac{in}{1+ix} + \frac{i(n+1)}{1-ix} \right] = \frac{1}{2} i [n\phi_{n-1} + (2n+1)\phi_n + (n+1)\phi_{n+1}],$$

where we have used the identities  $2\phi_n/(1 \pm ix) = \phi_n + \phi_{n\mp1}$ . This defines the first derivative operator  $D_1$  as

$$a_n^{(1)} = \frac{i}{2}[na_{n-1} + (2n+1)a_n + (n+1)a_{n+1}] \equiv D_1 a_n.$$

$D_2$  is similarly obtained:

$$\begin{aligned} a_n^{(2)} = & -\frac{1}{4}[n(n-1)a_{n-2} + 4n^2a_{n-1} + (6n^2 + 6n + 2)a_n \\ & + 4(n+1)^2a_{n+1} + (n+2)(n+1)a_{n+2}] \equiv D_2 a_n. \end{aligned}$$

When the series is truncated to  $-N \leq n \leq N-1$ , the infinite operators  $D_k$  turn into finite matrices, with  $D_1$  skew-hermitian and tridiagonal, and  $D_2$  real symmetric and pentadiagonal.

The matrices  $D_k$  are used to solve differential equations in the following manner. Consider, for example, the problem

$$-u''(x) + u(x) = f(x), \quad -\infty < x < \infty.$$

Assuming that  $u(x), f(x) \in L_2(\mathbb{R})$ , the problem may be approximated by

$$(-D_2 + I)\mathbf{a} = \mathbf{b}.$$

Here  $I$  is the  $2N \times 2N$  identity matrix, and  $\mathbf{a}$  and  $\mathbf{b}$  are the vectors of expansion coefficients of  $u(x)$  and  $f(x)$  respectively. The matrix  $-D_2 + I$  is symmetric and banded (and also positive definite as will follow from Theorem 4 below), and therefore the linear system may be efficiently solved. Numerical results for the case  $u(x) = \text{sech}(x)$ ,  $f(x) = 2\text{sech}^3(x)$ , have been presented in [17]. More realistic differential equations have been solved in [3,5,6].

In addition to stationary problems such as the one above, evolution equations like the one dimensional wave and diffusion equations may also be solved. These equations are respectively defined by

$$u_t = u_x, \quad u_t = u_{xx}, \quad -\infty < x < \infty,$$

together with suitable initial conditions  $u(x, 0) \in L_2(\mathbb{R})$ . By considering

$$u(x, t) = \sum_n a_n(t) \phi_n(x),$$

the PDEs may be reduced to systems of ODEs:

$$\frac{da_n}{dt} = D_1 a_n, \quad \frac{da_n}{dt} = D_2 a_n. \quad (17)$$

When these systems are truncated to  $-N \leq n \leq N-1$ , any of the standard methods for initial-value problems may be used to integrate with respect to time. This idea was applied to the Benjamin-Ono equation

$$u_t + uu_x + F_{\mathcal{H}}\{u_{xx}\} = 0, \quad -\infty < x < \infty,$$

which describes weakly nonlinear waves on the interface between two fluids of different density [11]. The Hilbert transform  $F_{\mathcal{H}}$  was calculated by the method described in Sect. 3. This approach was found to yield highly accurate results, provided the waves do not move too far from the origin. As the waves move out the increase in the grid-spacing causes a lack of resolution and the accuracy deteriorates (recall  $x_j = \tan \pi j/2N$ ).

We conclude this section with some results of a theoretical nature. In many applications, particularly those involving stability analyses, it is necessary to have information on the eigenvalues of the differentiation operators  $D_k$ . The simplest example is the linear systems (17). When an explicit method is used to integrate these systems, it is necessary that all of the eigenvalues of  $D_1$  (or  $D_2$ ), multiplied with the time step  $\Delta t$ , lie within the region of absolute stability of the method. If not, the solution grows unboundedly as  $t \rightarrow \infty$ . This stability constraint poses a restriction on the maximum allowable time step.

To find the eigenvalues of  $D_k$ , it is perhaps best to look at the explicit display of the tridiagonal matrix  $D_1$ :

$$D_1 = \frac{i}{2} \begin{pmatrix} 1-2N & 1-N & & & & & & \\ 1-N & \ddots & & & & & & \\ & -5 & -2 & & & & & \\ & -2 & -3 & -1 & & & & \\ & & -1 & -1 & & & & \\ & & & & 1 & 1 & & \\ & & & & & 1 & 3 & 2 \\ & & & & & & 2 & 5 \\ & & & & & & & \ddots & N-1 \\ & & & & & & & & N-1 & 2N-1 \end{pmatrix}. \quad (18)$$

It suffices to find the eigenvalues of the sub-matrix in the lower right-hand block. This is done in the usual manner, by finding a recurrence formula for the characteristic polynomial  $P_N(\lambda)$ . One obtains

$$P_n(\lambda) = (2n-1-\lambda)P_{n-1}(\lambda) - (n-1)^2 P_{n-2}(\lambda), \quad n = 1, \dots, N,$$

which is recognized to be the recurrence relation for the Laguerre polynomials. Therefore  $P_N(\lambda) = L_N(\lambda)$ . We have thus proved the following theorem, first published in [17]:

**Theorem 3** *The eigenvalues of the matrix  $D_1$  are distinct, pure imaginary, and given by*

$$\lambda_j = \pm \frac{1}{2}i\mu_j, \quad j = 1, \dots, N,$$

where the  $\mu_j$  are the  $N$  distinct roots of the Laguerre polynomial  $L_N(x)$ .

Turning to the second derivative infinite operator it follows that  $D_2 = D_1^2$ , so the eigenvalues of  $D_2$  are the squares of the eigenvalues of  $D_1$ . However, due to the

truncation  $D_2 \neq D_1^2$  for the *finite* matrices. Nevertheless, it is possible to bound the eigenvalues of the matrix  $D_2$  in terms of the roots of the Laguerre polynomials, as was demonstrated in [17].

**Theorem 4** *The eigenvalues of the matrix  $D_2$  occur in pairs, are real and negative, and given by*

$$\lambda_j = -\frac{1}{4}\mu_j^2, \quad j = 1, \dots, N,$$

where

$$\eta_j < \mu_j < \nu_j,$$

and  $\eta_j$ ,  $\nu_j$  are the  $j$ -th and  $j+1$ -st roots of  $L_N(x)$  and  $L_{N+1}(x)$  respectively.

**6. Conclusions and Open Questions.** We have shown the basis set  $\{\phi_n\}$  to be a highly effective tool in the computation of integral transforms and certain special functions related to the complex error function. In addition, we have surveyed the applications of this basis set to the solution of differential equations posed on the real line. We suspect that there are many potential applications of this approach awaiting discovery.

As for open questions, it is unclear under which conditions the rational basis set can be guaranteed to be superior to its main rivals, namely Hermite expansions or sinc functions. For example, for the test function  $f(x) = 1/(1+x^2)$  the rational basis set cannot be improved upon, since this function can equivalently be expressed as  $f(x) = \frac{1}{2}\phi_0(x) + \frac{1}{2}\phi_1(x)$ . In the language of Fourier analysis this function is *band-limited* with respect to the basis set  $\{\phi_n\}$ . By contrast Hermite or sinc functions require many terms to represent a slowly decaying function such as this one accurately. On the other hand, if the function is of the form  $e^{-x^2} \times \text{polynomial}(x)$ , Hermite functions will undoubtedly be superior.

A comprehensive comparison of the accuracy of expansions such as these will involve a study of the rate of decay in the expansion coefficients  $a_n$  as  $n \rightarrow \infty$ . A complete characterization of this rate of decay in terms of elementary properties of  $f(x)$  is out of reach, unfortunately, even in the case of ordinary Fourier series. We feel, however, that analyzing model functions such as the ones presented in the Table in Sect. 3 is a start. The list needs to be extended significantly however.

Related to the question of accuracy is the choice of the optimal parameter  $p$ . A specific theoretical estimate has been carried out in detail [18], the accuracy of which was verified in Fig. 3. But a general theory is nonexistent. One approach to finding the optimal parameter might be an experimental attack: presumably one could compute a few sets of expansion coefficients for various  $p$ , followed by some optimization strategies, the aim being to maximize the rate of decay over the expansion coefficients  $a_n$  retained in the expansion.

Whether or not these open questions can be resolved is unclear. What is clear, however, is that we have presented sufficient evidence that the rational basis set  $\{\phi_n\}$  is a powerful computational tool, and ready to take its rightful place among the better known basis sets appropriate for the real line.

## References

- [1] Boyd J. (1982): The optimization of convergence for Chebyshev polynomial methods in an unbounded domain. *J. Comp. Phys.* **45**, 43–79
- [2] Boyd, J. (1987): Spectral methods using rational basis functions on an infinite interval. *J. Comp. Phys.* **69**, 112–142
- [3] Boyd J. (1989): *Chebyshev & Fourier Spectral Methods*. Springer, Berlin
- [4] Boyd J. (1990): The orthogonal rational functions of Higgins and Christov and algebraically mapped Chebyshev polynomials. *J. Approx. Theory* **6**, 98–105
- [5] Christov C.I. (1982): A complete orthonormal system of functions in  $L^2(-\infty, \infty)$  space. *SIAM J. Appl. Math.* **42**, 1337–1344
- [6] Christov C.I. and Belyarov K.L. (1990): A Fourier-series method for solving soliton problems. *SIAM J. Sci. Stat. Comp.* **11**, 631–647
- [7] Davis B. and Martin B. (1979): Numerical inversion of the Laplace transform: A survey and comparison of methods. *J. Comp. Phys.* **33**, 1–32
- [8] Duffy D.G. (1993): On the numerical inversion of Laplace transforms: Comparison of three new methods on characteristic problems from applications. *ACM Trans. Math. Soft.* **19**, 333–359
- [9] Faddeeva V.N. and Terent'ev N.N. (1954): Tables of values of the function  $w(z) = e^{-z^2}(1 + \frac{2i}{\sqrt{\pi}} \int_0^z e^{t^2} dt)$  for complex argument. Gosud. Izdat. Teh.-Teor. Lit., Moscow; English transl.: Pergamon Press, New York
- [10] Higgins J.R. (1977): Completeness and basis functions of sets of special functions. Cambridge University Press, Cambridge
- [11] James R.L. and Weideman J.A.C. (1992): Pseudospectral methods for the Benjamin-Ono equation. In: Vichnevetsky R., Knight D. and Richter G. (eds.) *Advances in Computer Methods for PDES–VII*. IMACS, New Brunswick
- [12] Lyness J.N. and Giunta G. (1986): A modification of the Weeks method for Numerical Inversion of the Laplace Transform. *Math. Comp.* **47**, 313–322
- [13] Poppe G.P. and Wijers C.M. (1990): More efficient computation of the complex error function. *ACM Trans. Math. Software* **16**, 38–46
- [14] Rosser J.B. (1948): Theory and applications of  $\int_0^z e^{-x^2} dx$  and  $\int_0^z e^{-x^2} dy \int_0^y e^{-x^2} dx$ . Mapleton House, Brooklyn
- [15] Weber H. (1981): Numerical computation of the Fourier transform using Laguerre functions and the fast Fourier transform. *Numer. Math.* **36**, 197–209
- [16] Weeks W.T. (1966): Numerical inversion of Laplace transforms using Laguerre functions. *J. ACM* **13**, 419–429
- [17] Weideman J.A.C. (1992): The eigenvalues of Hermite and rational spectral differentiation matrices. *Numer. Math.* **61**, 409–431
- [18] Weideman J.A.C.: Computation of the complex error function. To appear in *SIAM J. Numer. Anal.*
- [19] Weideman J.A.C.: Computing the Hilbert transform on the real line. To appear in *Math. Comp.*
- [20] Weideman J.A.C.: Computing integrals of the complex error function. To appear in *Proceedings of Symposia in Applied Mathematics: Mathematics of Computation 1943–1993*
- [21] Wiener N. (1947): *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. M.I.T. Press, Cambridge