Evaluating of Dawson's Integral by solving its differential equation using orthogonal rational Chebyshev functions

John P. Boyd
Department of Atmospheric, Oceanic and Space Science and Laboratory for Scientific Computation, University of Michigan, 2455 Hayward Avenue, Ann Arbor, MI 48109, United States

**Abstract**

Dawson's Integral is \( u(y) = \exp(-y^2) \int_0^y \exp(z^2) \, dz \). We show that by solving the differential equation \( du/dy + 2yu = 1 \) using the orthogonal rational Chebyshev functions of the second kind, \( SB_n(y; L) \), which generates a pentadiagonal Petrov–Galerkin matrix, one can obtain an accuracy of roughly \( (3/8)N \) digits where \( N \) is the number of terms in the spectral series. The \( SB \) series is not as efficient as previously known approximations for low to moderate accuracy. However, because the \( N \)-term approximation can be found in only \( O(N) \) operations, the new algorithm is the best arbitrary-precision strategy for computing Dawson's Integral.

**1. Introduction**

Dawson's Integral is defined by

\[
u(y) = \exp(-y^2) \int_0^y \exp(z^2) \, dz.
\]

This transcendental arises in many applications. Our motive is that Dawson's Integral is the Hilbert Transform of a Gaussian, and is therefore essential in developing an adaptive numerical method for solving the Benjamin–Ono equation by Gaussian radial basis functions [19,10]:

\[
\mathcal{H}[\exp(-x^2)](y) \equiv \frac{1}{\pi} \mathcal{P} \int_{-\infty}^\infty \frac{\exp(-x^2)}{x-y} \, dx = \frac{2}{\sqrt{\pi}} u(y).
\]

Because of its importance in applications, there is already a literature on approximating Dawson's Integral including [14,12,11,9,6,7,15,17,18]. Dawson's Integral can be written as an error function of imaginary argument via

\[
u(y) = -\exp(-y^2) \frac{i}{\sqrt{\pi}} \text{erf}(iy).
\]

In addition, the “plasma dispersion function” or “Fadeeva function” is really Dawson's Integral, too. Many articles on evaluating the error function in the complex plane such as [17] are implicitly methods for Dawson's Integral, too. Our new method is not as efficient as that of Cody et al. [6], which uses four different approximations on four subintervals that span all real \( x \). It is, however, a single, global expansion that converges exponentially fast with the truncation \( N \) and can be computed for any \( N \) to any desired precision in \( O(N) \) operations.
2. A Galerkin rational function method to solve the differential equation for Dawson’s Integral

Dawson’s Integral solves:
\[
\frac{du}{dy} + 2yu = 1, \quad y \in [-\infty, \infty].
\]  
(4)

The boundary conditions, discussed further below, are that \(u(y)\) is bounded as \(|y| \to \infty\). The first step is the change of coordinate \(y = L \cot(t)\) which transforms the problem into

\[
sin^3(t)u_t - 2L^2 \cos(t)u = -L \sin(t),
\]  
(5)

where \(L > 0\) is a user-choosable constant “map parameter” whose role will be discussed further below.

It has long been known that Dawson’s Integral has the divergent asymptotic expansion [6]:

\[
u(y) \sim \frac{1}{2} \sqrt{\frac{L}{2}} \sum_{k=0}^{M} \frac{\Gamma(k+1/2)}{\Gamma(1/2)} \frac{1}{y^{2k}}.
\]  
(6)

By analogy with the known properties of the similar series (in powers of \(x\) rather than \(x^2\)) for the exponential integral, the optimal truncation of the Dawson’s series is \(M(x) \approx \text{round}(x^2)\) which yields an optimal error proportional to \(\exp(-x^2)\).

As explained in [2], when a function decays as odd powers of \(y\), as known to be true of this function, the correct choice of Fourier basis is to use a sine series. (The standard basis functions, denoted by \(\text{TB}_n(y; L)\), are the images of a cosine series in \(t\), but the TB series converges very slowly for Dawson’s Integral because the cosines/TB functions do not have the correct behavior at \(y = \pm \infty \leftrightarrow t = 0, \pi\).) Dawson’s Integral is antisymmetric in \(y\) and this implies that \(u(y; t)\) must be antisymmetric with respect to \(t = \pi/2\). Thus, it is sufficient to use only a basis of even sines in \(t\), which gives:

\[
u(y) = \sum_{n=1}^{\infty} b_n \sin(2nt[y]).
\]  
(7)

It is legitimate to describe this as a “rational” function expansion because, when written in terms of the original coordinate \(y\),

\[
sin(2nt) = \text{SB}_{2n}(y; L) = \frac{L}{\sqrt{L^2 + y^2}} U_{2n-1} \left( \frac{y}{\sqrt{L^2 + y^2}} \right),
\]  
(8)

where \(U_n\) is the usual Chebyshev polynomial of the second kind. The square roots combine to give a true rational function.

The differential equation \(t\) is singular at both \(t = 0\) and \(\pi\), but this is actually a good thing. It is not necessary to explicitly impose numerical boundary conditions because the spectral series will automatically pick out the solution which is analytic and well-behaved at the singular points. Such “natural” boundary conditions are discussed at length in [4,1,2]. Usually, of course, a first order ordinary differential equation has a unique solution only with the explicit imposition of an initial condition such as \(u(\pi/2) = 0\). Dawson’s Integral actually satisfies this condition, and this is implicitly imposed by the basis of even sine functions. However, it is unusual to need to impose boundary conditions when the differential equation is imposed on an infinite interval (as in the original coordinate \(y\)), or is the image of such a problem under a coordinate mapping.

When the sine series is substituted into the transformed differential equation, the result is a series of odd sines. Collecting the terms and demanding that each sine coefficient individually equal zero gives

\[
r_1 = (-1 - L^2)b_1 + (1/2)b_2 = -L,
\]  
(9)

\[
r_3 = ([3]/4 - L^2)b_1 + ((-3/2) - L^2)b_2 + (3/4)b_3 = 0,
\]  
(10)

plus for \(m \geq 2\)

\[
r_{2m-1} = (-m - 2)/4b_{m-2} + (3m - 1/4 - L^2)b_{m-1} + (-3m/4 - L^2)b_m + ((3m + 1)/4)b_{m+1}, \quad m = 3, 4, \ldots
\]  
(11)

If we organize these conditions that \(r_m = 0\) for \(m = 1, 2, \ldots, N\) into an \(N\)-dimensional matrix equation where \(N\) is the truncation of the sine series for \(u(y; t)\) and the unknown vector is one whose elements are the sine coefficients \(b_n\), we obtain a pentadiagonal matrix with at most only four nonzero elements (instead of five, as typical for a general pentadiagonal matrix) in each row. The usual operation counts for a pentadiagonal matrix show that the solution can be obtained in only \(19N\) floating point operations.

In the language of mean-weighted residual methods, this is a “Petrov–Galerkin” method because the weight functions (odd sines) in the integrals that give the matrix elements are different from the basis functions themselves. We shall not bother to write down the usual Petrov–Galerkin integrals because we can obtain the elements by simply applying trigonometric identities and rearrangement of the odd-sine series for the residual.

Ideally, the map parameter \(L\) is chosen to optimize the rate of convergence. Weideman in a closely related problem [17] shows that the optimum \(L\) grows with the truncation proportional to \(\sqrt{N}\), but determining the proportionality constant from
theory is difficult. Fortunately, a little experimentation is usually sufficient because near the optimum, the accuracy is insensitive to the precise value of $L$.

Fig. 1 shows the SB/ $\sin(2t)$ coefficients as computed for several $L$. The optimum $L$ for $N = 40$ is about $L = 4$, but is smaller for smaller $N$: note that the curves for small $L$ are well below those of the dark circles for $L = 4$.

Because $|\sin(2nt)| \leq 1$ for all real $t$, it follows that the error in truncating a spectral series is bounded by the sum of the absolute value of all the neglected coefficients [4]. It is clear from the exponential decay of the coefficients that the error for $N = 40$ is very close to machine precision. The fact that the coefficients decay so close to machine precision is a numerical proof that the Petrov–Galerkin matrix problem is very well-conditioned.

On a finite domain, the usual rate of convergence for spectral series is “geometric”, that is, the coefficients asymptote to a straight line when plotted versus degree on a log-linear graph. However, the apparent geometric convergence evident in Fig. 1 is an illusion as explained in [17,1,2]. Because Dawson’s Integral is singular at $y = \infty$, theory predicts that the coefficients $b_n$ are proportional to $\exp(-qn^{2/3})$ for some constant $q > 0$ for fixed $L$. The concave upward curvature of the coefficients in Fig. 2, instead of the linear behavior of geometrical convergence, shows that the convergence rate is indeed “subgeometric”.

As noted by Weideman [17] and Boyd [1,2,5], however, the optimum $L$ increases with the truncation. For Dawson’s Integral, our experiments suggest, in agreement with the theory of Weideman [17], that

$$L \approx 2^{-1/4} \sqrt{N},$$

is near-optimum. The theory of these authors for closely related problems suggests that a geometric rate of decrease of the error with the truncation $N$ can be achieved by allowing $L$ to vary with $N$ thusly.

To confirm this, we performed multiple precision computations for various $N$ in Maple. Fig. 3 shows that the convergence is, choosing $L$ as in (12), indeed geometric. The dotted curve show that the error in the $L_\infty$ norm is well-approximated by

$$\max_{y \in [-\infty, \infty]} |u(y) - u_0(y)| \approx 0.07 \exp(-0.88N),$$

which is in good agreement with [17], which predicts that the constant in the exponential is $-\log(\sqrt{2} - 1)$. The maximum error is roughly $10^{-3/8N}$.

The algorithm is so simple that the entire Maple algorithm, missing only printing and plotting commands, is given as Table 1.

3. Low order approximations

When $N$ is small, the SB series has a rather large relative error for large $|y|$ because the Galerkin method does not enforce the asymptotic behavior:

![Fig. 1. The coefficients $b_n$ versus $n$ for six different values of $L$. The dark circles show the best choice for $N = 40$, which is $L = 4$.](image-url)
From the observation that, with $t = \arccot \left( \frac{y}{L} \right)$,

$$u(y) \sim \frac{1}{2} \frac{1}{y} + O \left( \frac{1}{y^2} \right).$$

From the observation that, with $t = \arccot(y/L)$,
\[
\sin(2nty) \sim (2n/L)y + O(1/y^3),
\]

it follows that:
\[
\sum_{n=1}^{\infty} 2nb_n = \frac{1}{2L}.
\]

We can correct a truncated series to enforce this condition by which gives:
\[
b_{N+1} = \frac{1/(2L) - \sum_{n=1}^{N} 2nb_n}{2(N+1)}.
\]

For example, the 4-term approximated derived this way has a maximum pointwise error of only 0.0064, with
\[
t = \arccot(x/[4/3]):
\]
\[
u(y) \approx 0.45365 \sin(2t) - 0.14638 \sin(4t) - 0.01806 \sin(6t) + 0.02020 \sin(8t)
\]
or equivalently in terms of \(y\) as
\[
u(y) \approx \frac{0.91714y + 1.97920y^3 + 0.29856y^5 + 0.050056y^7}{(1 + (9/16)y^2)^4}, \quad y \in [-\infty, \infty].
\]

This is excellent accuracy for such a low order approximation, but see [13] for comparison. Lether has computed a near-minimax approximation of the same order (his \(R_{3,4}\)) which has a global error that is smaller by a factor of 10.

4. Computing derivatives of Dawson’s Integral

From the differential equation satisfied by Dawson’s Integral and derivatives of that equation, one obtains
\[
\frac{du}{dy} = 1 - 2yu(y),
\]
\[
\frac{d^2u}{dy^2} = -2y + (4y^2 - 2)u(y),
\]
and so on. The only drawback of these formulas is that since \(u(y) \sim (1/2)/y\) as \(|y| \to \infty\), it follows that there will be cancellation for large \(|u|\) in the sense that the first derivative, which is \(O(1/y^2)\), is the difference of two terms which are \(O(1)\). This cancellation is so mild that it is not necessary to have special routines to evaluate derivatives.

5. Previous spectral methods for Dawson’s Integral

Hummer [9] has expanded Dawson’s Integral as a Chebyshev polynomial series, but his approximations are valid only for a finite range in \(x\). Coleman [7] applies the Lanczos tau-method to solve the same differential equation as here with a tau
forcing that is a Faber polynomial to obtain good accuracy within a sector in the complex plane. However, he does not employ a change of coordinate as done here, and his expansions are restricted to a finite domain.

Weideman [17] is the nearest to our approach. He also obtains an expansion in orthogonal rational functions. However, he derives the coefficients by inserting a rational orthogonal series under the alternative integral representation:

\[ u(y) = \frac{1}{\sqrt{2\pi}} \text{PV} \int_{-\infty}^{\infty} \frac{\exp(-z^2)}{y-z} \, dz, \]  

where “PV” denotes the principal value of the integral. His approach is a strength because his approximation (to the error function) is accurate for general complex values of \( y \). He derives asymptotic approximations to the spectral coefficients, explicitly displaying the subgeometric rate of convergence which is true of the SB series, too. However, he does not obtain the coefficients in the series by solving banded matrix problems derived from the differential equation, but instead by numerical quadrature of each coefficient integral, which is considerably slower.

Weideman’s orthogonal rational basis functions are complex-valued:

\[ \sigma_n(y) \equiv \left( \frac{L + iy}{L - iy} \right)^n, \quad n = -\infty, \ldots, \infty. \]  

There is a close connection between his rational functions and those used here [3], but the interrelationships of various species of rational orthogonal functions are too intricate to recapitulate here. Suffice it to say the theoretical properties are very similar. However, Dawson’s Integral is a real-valued function of a real-valued argument in most applications. It is obviously convenient to use a rational basis that requires only real arithmetic as presented here.

6. Summary

When a special function satisfies a simple ODE with polynomial coefficients, as true of Dawson’s Integral, a spectral method will usually yield a banded matrix of low bandwidth to determine the spectral coefficients. Denoting the number of terms retained in the spectral series by \( N \), the Petrov–Galerkin spectral method thereby furnishes an \( O(N) \) method to evaluate the function at a point to any desired degree of precision. In that sense this present work is merely the most recent continuation of a long line of work that is well-reviewed in [8,16].

One novelty here is the use of the rational functions \( \text{SB}_n(y; L) \). These have also been used to solve the “Yoshida jet” problem in oceanography [4].

Optimized minimax rational approximations are more accurate than ours. However, the minimax computations are too costly to employ for arbitrary precision evaluation, but are more suitable for constructing fixed precision library subroutines.

In contrast, to obtain roughly \( d \) decimal digits of accuracy where \( d \) is arbitrary, it is sufficient to apply the SB method derived here with

\[ N \approx (8/3)d. \]  

Acknowledgements

This work was supported by the National Science Foundation through Grants OCE 0451951 and ATM 0723440. I thank Andre’ Weideman for helpful comments and for reprints. I also thank the reviewer for useful corrections.

References