THE EXPONENTIALLY CONVERGENT TRAPEZOIDAL RULE

LLOYD N. TREFETHEN* AND J. A. C. WEIDEMAN†

Abstract. It is well known that the trapezoidal rule converges geometrically when applied to analytic functions on periodic intervals or the real line. The mathematics and history of this phenomenon are reviewed and it is shown that far from being a curiosity, it is linked with computational methods all across scientific computing, including algorithms related to inverse Laplace transforms, special functions, complex analysis, rational approximation, integral equations, and the computation of functions and eigenvalues of matrices and operators.

Key words. trapezoidal rule, aliasing, quadrature, resolvent, Hankel contour, double exponential quadrature, FFT, Euler–Maclaurin formula, Faraday cage

AMS subject classifications. 65D32, 42A15, 65-02, 65R20

*Oxford U. Mathematical Institute, Oxford OX2 6GG, UK (trefethen@maths.ox.ac.uk). Supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013)/ERC grant agreement no. 291068. The views expressed in this article are not those of the ERC or the European Commission, and the European Union is not liable for any use that may be made of the information contained here.

†Dept. of Mathematical Sciences, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa (weideman@sun.ac.za). Supported by the National Research Foundation in South Africa.

Part I. Fundamentals

1. Introduction. Many people experienced in numerical computation have encountered a remarkable phenomenon: if the composite trapezoidal rule for integration is applied to a periodic integrand, or an integral over the whole real line, it is often exponentially accurate. Some have discovered this for themselves. Others learned it from a textbook or a colleague, though they may not quite recall which, or who.

It appears to have been Poisson, in the 1820s, who first identified this effect [137]. The example Poisson chose has remained a favorite ever since: the perimeter of an ellipse, which he took to have axis lengths $1/\pi$ and $0.6/\pi$, giving the integral

$$I = \frac{1}{2\pi} \int_0^{2\pi} \sqrt{1 - 0.36 \sin^2\theta}\, d\theta. \tag{1.1}$$

Poisson used this now-standard notation for definite integrals, but apparently it was not yet standard in the 1820s, for he comments that

> pour indiquer [les limites de l'intégrale] en même temps que l'intégrale, nous emploierons la notation très-commode que M. Fourrier a proposée.[1]

Perhaps in 1826 the spelling of Fourier's name wasn't yet standardized either!

The exact solution of (1.1) is

$$I = \frac{2}{\pi} E(0.36) = 0.90277992777219\ldots, \tag{1.2}$$

where E is the complete elliptic integral of the second kind. As trapezoidal rule approximations we can take

$$I_N = \frac{1}{N} \sum_{k=1}^{N} \sqrt{1 - 0.36 \sin^2(2\pi k/N)}$$

for any positive integer N, or equivalently, if N is divisible by 4, exploiting the fourfold symmetry as Poisson did,

$$I_N = \frac{4}{N} {\sum_{k=0}^{N/4}}' \sqrt{1 - 0.36 \sin^2(2\pi k/N)}.$$

The prime on the summation indicates that the terms with $k = 0$ and $k = N/4$ are multiplied by $1/2$. For $N = 4, 8, \ldots, 20$ we get the following rounded results, with digits that match the exact solution shown in boldface:

$N/4$	I_N
1	**0.90**00
2	**0.9027**69
3	**0.90277**986
4	**0.902779927**2
5	**0.902779277**67

[1] "To indicate the limits of integration at the same time as the integrand, we shall employ the very convenient notation proposed by Mr. Fourrier."

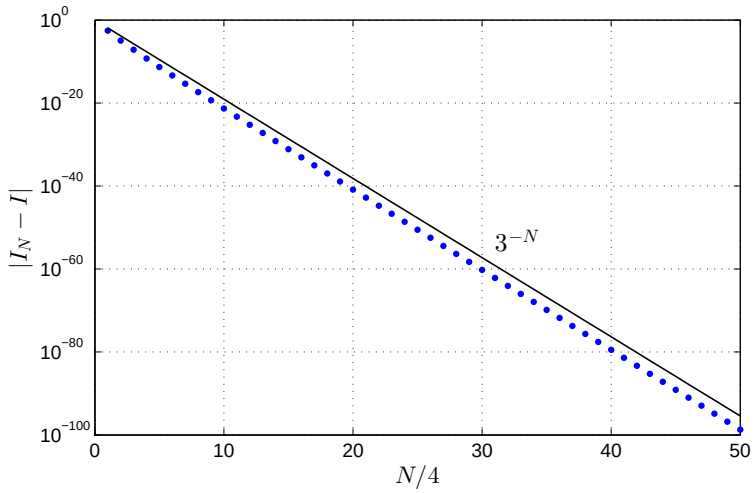


FIG. 1.1. *Convergence of the trapezoidal rule for computing the perimeter of an ellipse, with the error $|I_N - I|$ plotted against $N/4$ since symmetry permits evaluation of the integrand at just $N/4 + 1$ points. The convergence is geometric, although it is not clear whether Poisson recognized this in the* 1820*s.*

The convergence is extraordinarily fast. Poisson saw this, and he computed the result for $N = 16$, which requires just three nontrivial function evaluations, to ten digits. Confusingly, he reported these with a misprint in the second place,

$$\text{“la valeur approchée de } I \text{ sera } I = 0{,}9927799272\text{”}.$$

This was certainly just a misprint, since the following eight digits are right, and in fact, Poisson showed that his number must be in error by less than 4.84×10^{-6}. Using essentially the Euler–Maclaurin formula, he proved that $I_N - I$ decreases as $N \to \infty$ at least as fast as N^{-6}. In view of the style of mathematical exposition of his day, we can take this as the observation that the convergence is faster than N^{-m} for any $m > 0$. It is not clear whether Poisson realized that the convergence is actually geometric. In fact, it follows from Theorem 4.2 that for this problem we have

$$|I_N - I| = O(3^{-N}),$$

since the integrand has branch points in the complex θ-plane at a distance $\log(3)$ above and below the real axis. Thus each new row in the table brings an improvement by nearly two digits, a factor of 81, as confirmed in Figure 1.1. One could hardly ask for more, and indeed, the trapezoidal rule is an excellent method for computing complete elliptic integrals (see §18).

Our aim in the first half of this article is to present the mathematical foundations of the geometrically convergent trapezoidal rule (§2–§12). The geometric convergence applies to four canonical classes of analytic functions:

(a) on a circle in the complex plane,
(b) on an interval, periodic,
(c) on the real line, with sufficient decay at $\pm\infty$,
(d) on a Hankel contour wrapping around $(-\infty, 0]$ in the complex plane.

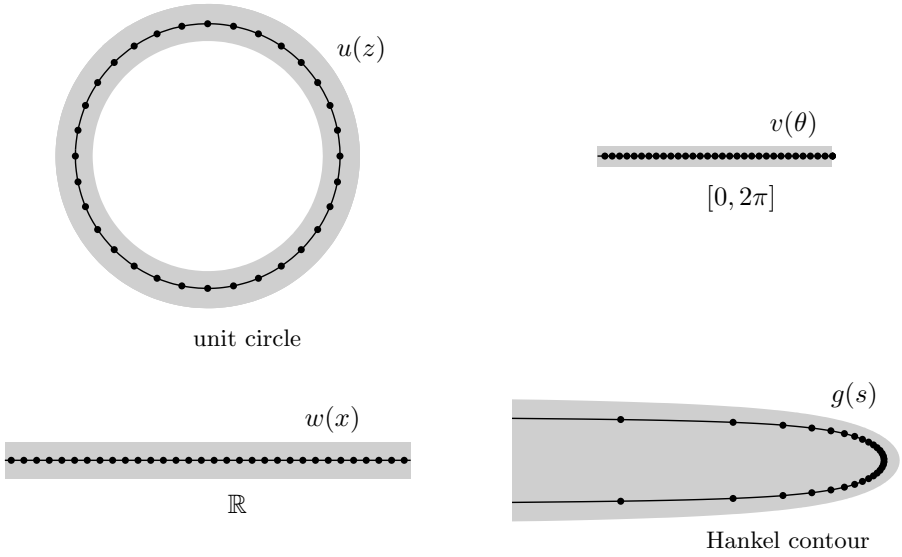


FIG. 1.2. *Four domains for the geometrically convergent trapezoidal rule and the variable names that are associated with them throughout this article. (For functions on* $[-1, 1]$ *or* $[0, 1]$*, we use* $y(\xi)$*.) The first three are the canonical domains considered in* §2–§8*. The fourth, equally important in practice, is the Hankel contour used for evaluation of integrals such as inverse Laplace transforms in the left half-plane* (§§16, 17)*. For Hankel contours, the trapezoidal rule is applied after a change of variables, an idea also used in other contexts including exponential and double exponential quadrature rules for nonperiodic functions on an interval* (§15)*.*

(Each of (a)–(c) further comes in a nonsymmetric or symmetric variant: disk or annulus; periodic half-plane or periodic strip; half-plane or infinite strip.) Theorems explaining the fast convergence have been based on four kinds of arguments:

(i) series and aliasing (or, for the real line, transforms and aliasing),
(ii) contour integrals and residue calculus,
(iii) interpolation,
(iv) the Euler–Maclaurin formula and related ideas.

These two lists suggest an array of related theorems and proofs. The case (d) of a Hankel contour is usually handled by reducing it to case (c), the real line, but even if we ignore (d), this still leaves twelve types of proofs that might be considered. In the early sections of this paper we shall give examples of nine of these twelve (see Table 12.1) as we set forth a collection of these ideas in as systematic a manner as possible. Our focus is always on geometric rates of convergence associated with analytic functions, not on the next-order algebraic estimates that depend on behavior at the edge of analyticity. We are not aware of any previous wide-ranging survey of this material.

Interspersed with the foundational material in the first half of the article are sections detailing a connection of the trapezoidal rule with the Faraday cage effect in electromagnetics (§3), a numerical example (§5), and a brief review of the history (§12). Who first noticed the geometric convergence of the trapezoidal rule, and what theorems were proved along the way? We do not know all the answers, but there are some clear post-Poisson milestones such as the papers of Turing in 1943 and Davis in 1959. We summarize what we have found in Table 12.1.

Up to this point one might imagine that the geometrically convergent trapezoidal rule is a rather isolated topic, more or less a curiosity in the field of numerical analysis. Our aim in the second half of this article is to show that in fact, it plays a significant role in scientific computing. We focus on eight areas of application, from Cauchy integrals in complex analysis to the solution of integral equations and time-dependent PDE (§13–§20).

A word about terminology. In the context of periodic or infinite domains, there is little difference between the "trapezoidal" and the "midpoint" or "rectangle" rules, and we use the first of these terms mainly because it has become more standard.[2] If one makes a distinction, sometimes the trapezoidal variant may have the advantage that it more readily reuses function values as N is doubled, whereas the midpoint variant has slightly simpler symmetry properties and no inconvenient divisions by 0 at endpoints after certain changes of variables (§16). For the purposes of this article, however, these distinctions are not very important, and we will just speak of the trapezoidal rule.

2. Integrals over a circle in the complex plane. From the user's point of view, integrals of functions periodic on an interval represent perhaps the most basic form of the exponentially convergent trapezoidal rule. Mathematically, however, integrals over circles in the complex plane are arguably even simpler. We therefore begin with this case, though the reader who chooses to move directly to the case of the periodic interval in §4 will not lose much. The circle is also the setting of the *z-transform,* an indispensable tool in digital signal processing [173].

Let u be a real or complex function defined on the unit circle, which we shall denote by the abbreviated expression $|z| = 1$, and, setting $z = e^{i\theta}$, define

$$I = -\int_{|z|=1} iz^{-1}u(z)\,dz = \int_0^{2\pi} u(e^{i\theta})\,d\theta. \tag{2.1}$$

(Here and throughout this paper, a complex contour such as $|z| = 1$ is always to be understood as traversed once in the counterclockwise direction.) Note that I can be interpreted as 2π times the mean of u over the unit circle. For any positive integer N, we define the trapezoidal rule approximation to I by

$$I_N = \frac{2\pi}{N}\sum_{k=1}^{N} u(z_k), \tag{2.2}$$

where $\{z_k\}$ are the Nth roots of unity, $z_k = e^{2\pi ik/N}$. (Because of the periodicity, no special factors of $1/2$ are needed at the endpoints as one would have with the non-periodic trapezoidal rule.) This approximation will be exponentially accurate if u is analytic in a neighborhood of $|z| = 1$, i.e., in an annulus. For the simplest argument we consider first the case in which it is analytic in a disk.

THEOREM 2.1. *Suppose u is analytic and satisfies $|u(z)| \le M$ in the disk $|z| < r$ for some $r > 1$. Then for any $N \ge 1$,*

$$|I_N - I| \le \frac{2\pi M}{r^N - 1}, \tag{2.3}$$

and the constant 2π is as small as possible.

[2]Another term often used is "trapezoid rule," and the British say "trapezium." Language aficionados may note that whereas English has distinct words trapezoid (in geometry) and trapeze (in a circus), other languages such as French and German use a single word for both meanings.

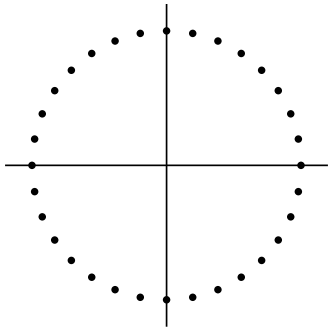


FIG. 2.1. *At the Nth roots of unity, z^N, z^{2N}, and so on are indistinguishable from the constant function 1—the phenomenon of aliasing. The error in the trapezoidal rule on the unit circle consists in counting a contribution of 2π from each such term in a series for $u(z)$ instead of the correct value of 0.*

We begin with the most elementary proof, summarized in Figure 2.1.

Proof of Theorem 2.1 *by Taylor series and aliasing.* The function u has the convergent Taylor series[3]

$$u(z) = \sum_{j=0}^{\infty} c_j z^j \tag{2.4}$$

with coefficients

$$c_j \;=\; \frac{1}{2\pi i}\int_{|z|=1} z^{-j-1}u(z)\,dz \;=\; \frac{1}{2\pi}\int_0^{2\pi} e^{-ij\theta}u(e^{i\theta})\,d\theta. \tag{2.5}$$

From (2.1) and (2.5) we have

$$I = -i\int_{|z|=1} z^{-1}u(z)\,dz = 2\pi c_0, \tag{2.6}$$

and from (2.2) and (2.4),

$$I_N \;=\; \frac{2\pi}{N}\sum_{j=0}^{\infty} c_j \sum_{k=1}^{N} z_k^j. \tag{2.7}$$

If j is a multiple of N, the numbers z_k^j are all equal to 1, and the second sum in this formula is equal to N. On the other hand if j is not a multiple of N, they are evenly spaced around the unit circle and sum to 0. This is the phenomenon of *aliasing*: z^{j+N}, z^{j+2N}, z^{j+3N}, and so on are indistinguishable from z^j at the Nth roots of unity. Thus (2.7) reduces to

$$I_N \;=\; 2\pi\sum_{j=0}^{\infty} c_{jN}, \tag{2.8}$$

[3]The Taylor series in this paper are centered at $z = 0$, so a more precise term would be *Maclaurin series.*

which together with (2.6) gives

$$I_N - I \;=\; 2\pi \sum_{j=1}^{\infty} c_{jN}. \tag{2.9}$$

To estimate the error, all we need is a bound on the coefficients c_{jN}, which must decay geometrically since u is analytic in a disk of radius $r > 1$. For a precise inequality we can deform the contour of integration in (2.5) to $|z| = r'$, where $r' < r$ is arbitrarily close to r; or if u is analytic in the closed disk we can simply take $r' = r$. This leads to the bound known as *Cauchy's estimate*,

$$|c_j| \le M r^{-j}, \quad j \ge 0. \tag{2.10}$$

Thus (2.9) implies

$$|I_N - I\,| \le 2\pi M \sum_{j=1}^{\infty} r^{-jN},$$

and summing this series gives (2.3). This completes the proof apart from the assertion that the constant 2π cannot be improved. For this it is enough to consider $u(z) = z^N$ and any $r > 1$, which gives $I = 0$, $I_N = 2\pi$, $M = r^N$, and thus $|I_N - I\,| = 2\pi M/r^N$. As $N \to \infty$, this is asymptotic to the bound (2.3). □

The proof just given represents one of the two most fundamental approaches to establishing theorems about exponential accuracy of the trapezoidal rule. The other, superficially a more surprising and ingenious method, is to use contour integrals and residue calculus. The idea, which can be traced at least in part to Cauchy in 1826,[4] goes as follows. We wish to compare an integral with a sum. We do this by converting the sum to another integral, so that we are left with the comparison of two integrals. How does one convert a sum to an integral? By multiplying by a function that has simple poles at the summation points, and then using residue calculus. In other words, every quadrature formula applied to an analytic function u is equivalent to a contour integral of $m(z)u(z)$ for some rational or meromorphic function m, known as a *characteristic function* for the quadrature formula. The nodes are the poles of m inside the contour, and the weights are the residues times $2\pi i$. We shall use arguments of this kind many times in this article, and the characteristic functions involved are summarized in Table 14.1.

Proof of Theorem 2.1 *by residue calculus.* The function

$$m(z) = \frac{-iz^{-1}}{1 - z^{-N}} \tag{2.11}$$

has simple poles at the Nth roots of unity with residues equal to $-i/N$, i.e., the weights of the trapezoidal rule (2.2) divided by $2\pi i$. By residue calculus, this implies

$$I_N \;=\; \int_{|z|=r'} m(z)u(z)\,dz \tag{2.12}$$

for any r' with $1 < r' < r$. Similarly, in (2.6), since u is analytic, we can enlarge $|z| = 1$ to the circle $|z| = r'$ without changing the value. Thus, combining (2.6) and

[4]For comments about the history, see [169, Chapter 11].

(2.12) gives a representation of $I_N - I$ as an integral,

$$I_N - I = \int_{|z|=r'} [m(z) + iz^{-1}]u(z)dz = \int_{|z|=r'} \frac{iz^{-1}}{1-z^N} u(z)dz. \tag{2.13}$$

From here (2.3) follows. □

Why would one want to integrate a function over a circle in the complex plane? As we shall discuss in §13, one reason is the Cauchy integrals and other kinds of contour integrals that are the basis of so much of complex function theory, with applications throughout applied mathematics. The simplest case of such an integral, already alluded to, is the mean value theorem for analytic functions: if u is analytic in a neighborhood of the unit disk, then in the notation of (2.1), $u(0) = I/2\pi$. We may illustrate this identity by an example. The function

$$u(z) = \frac{z}{e^z - 1} \tag{2.14}$$

has a removable singularity at $z = 0$; if we define $u(0) = 1$, it becomes analytic throughout the disk $|z| < 2\pi$. If we evaluate $u(z)$ for small values of z on a computer in 16-digit arithmetic, however, there is cancellation error and we may lose many digits of accuracy. By contrast, full precision is achieved by taking a mean over 18 points on the unit circle. We shall consider a more extreme example of this kind in §13.

If $u(0)$ is to be calculated by a mean over a circle, one may ask, why should that circle have radius 1? Any other radius that stays within the domain of analyticity is mathematically acceptable, and in the first integral of (2.1), this invariance is reflected in the presence of the scale-dependent term $-iz^{-1}$, which balances the scale-dependent contribution from dz. If $|z| = 1$ is replaced by $|z| = \rho$ for some $\rho > 0$, the change that results is that the powers z^N, z^{2N}, and so on of (2.7) sum to $N\rho^j$ instead of just N, and the summand in (2.8) and (2.9) changes from c_{jN} to $c_{jN}\rho^{jN}$. The end result is that (2.3) becomes

$$|I_N - I| \le \frac{2\pi M}{(r/\rho)^N - 1}. \tag{2.15}$$

Thus, at least if cancellation errors on a computer are not a concern, the trapezoidal rule converges geometrically as $N \to \infty$ at a rate determined by r/ρ, and it is advantageous to take ρ as small as possible. Of course, this conclusion only applies if u is analytic throughout a disk. For more about the choice of ρ, see §13 and [22, 49, 50].

Let us now turn to the case of a function $u(z)$ on the unit circle that may not be analytic throughout a disk, but just in an annulus. For example, if $u(z)$ is real and nonconstant on the circle, this will necessarily be the case. For simplicity, we suppose that the annulus of analyticity is symmetric about the unit circle, taking the form $r^{-1} < |z| < r$ for some $r > 1$. The assumption of boundedness in such an annulus is enough to give the same result as before, except that the factor 2π now doubles to 4π.

THEOREM 2.2. *Suppose u is analytic and satisfies $|u(z)| \le M$ in the annulus $r^{-1} < |z| < r$ for some $r > 1$. Then for any $N \ge 1$,*

$$|I_N - I| \le \frac{4\pi M}{r^N - 1}, \tag{2.16}$$

and the constant 4π is as small as possible.

Proof of Theorem 2.2 *by Laurent series and aliasing.* The proof of Theorem 2.1 by Taylor series generalizes readily to a proof of Theorem 2.2 by Laurent series. Equations (2.4)–(2.8) continue to hold, except now the limits of summation are $j = -\infty$ to ∞. Equation (2.9) becomes

$$I_N - I \;=\; 2\pi \sum_{j=1}^{\infty} (c_{jN} + c_{-jN}). \tag{2.17}$$

We can complete the proof by generalizing (2.10) to negative as well as positive j,

$$|c_j| \le M r^{-|j|}, \quad -\infty < j < \infty, \tag{2.18}$$

an inequality that follows in the case $j < 0$ from shrinking $|z| = 1$ to $1/r'$ for r' arbitrarily close to r. Finally, concerning the sharpness of the constant 4π, it is enough to consider the example $u(z) = z^N + z^{-N}$ and any $r > 1$, which gives $I = 0$, $I_N = 4\pi$, $M = r^N + r^{-N}$, and thus $|I_N - I| \sim 4\pi M/r^N$ as $N \to \infty$. ∎

An alternative approach to Theorem 2.2 is to note that any function u analytic in the annulus $r^{-1} < |z| < r$ can be split into $u = u^+ + u^-$, where the analytic part u^+ is analytic in $|z| < r$ and the coanalytic part u^- is analytic in $r^{-1} < |z| \le \infty$. Since u^+ and u^- can both be represented by Cauchy integrals, i.e., as appropriately weighted means of $u(z)$ on circles, they satisfy $|u^{\pm}(z)| \le M$ in the annulus, and from here, Theorem 2.2 can be obtained as a corollary of Theorem 2.1.

Another alternative is to prove Theorem 2.2 directly by residue calculus. This can be done following the pattern of the proof of Theorem 2.1 by residue calculus, but replacing the characteristic function $m(z)$ of (2.11) by the function listed in the (2,1) position of Table 14.1. Such an argument is spelled out in full in §4 for the equivalent problem of integration of a function periodic on a real interval.

Equation (2.17) (or Theorem 2.2 itself, after some easy estimates) has the following interesting corollary.

COROLLARY 2.3. *If u is a polynomial or Laurent polynomial of degree n (i.e., a linear combination of $z^{-n}, \dots, z^n$), the N-point trapezoidal rule* (2.2) *is exact for all $N > n$.*

There is a sharpening of Theorems 2.1 and 2.2 that is often applicable. If u is analytic but not bounded in the open disk $|z| < r$, Theorem 2.1 implies convergence at the rate $O((r-\varepsilon)^{-N})$ for any $\varepsilon > 0$ but not necessarily $O(r^{-N})$. However, if the singularities of u on the circle $|z| = r$ are just simple poles, the convergence rate is $O(r^{-N})$ after all. In the proof by Taylor series and aliasing, this follows from the fact that the Taylor coefficients associated with a pole $(z - z_0)^{-1}$ with $|z_0| = r$ are of size $O(r^{-n})$; compare [160, eq. (4.3)] and [169, Exercise 8.15]. In the proof by contour integrals, we can integrate over a slightly larger circle $|z| = r + \varepsilon$ for some $\varepsilon > 0$, while also including keyhole contributions from small circles around each pole with $|z| = r$; compare [40]. The $O(r^{-N})$ convergence of the trapezoidal rule also applies if there are a finite number of singularities on $|z| = r$ that are weaker than poles in the sense that u satisfies $u(z) = O(|z - \zeta|^{\alpha})$ for $|z| \le r$ with $\alpha > -1$ at each such singularity ζ. This can be proved by noting that since $|u|$ is integrable along $|z| = r$ across such a singularity, (2.5) leads to

$$|c_j| = \frac{1}{2\pi} \left| \int_{|z|=r} z^{-j-1} u(z)\, dz \right| \le \frac{1}{2\pi} \|z^{-j-1}\|_\infty \|u\|_1 = \frac{r^{-j-1}}{2\pi} \|u\|_1 \tag{2.19}$$

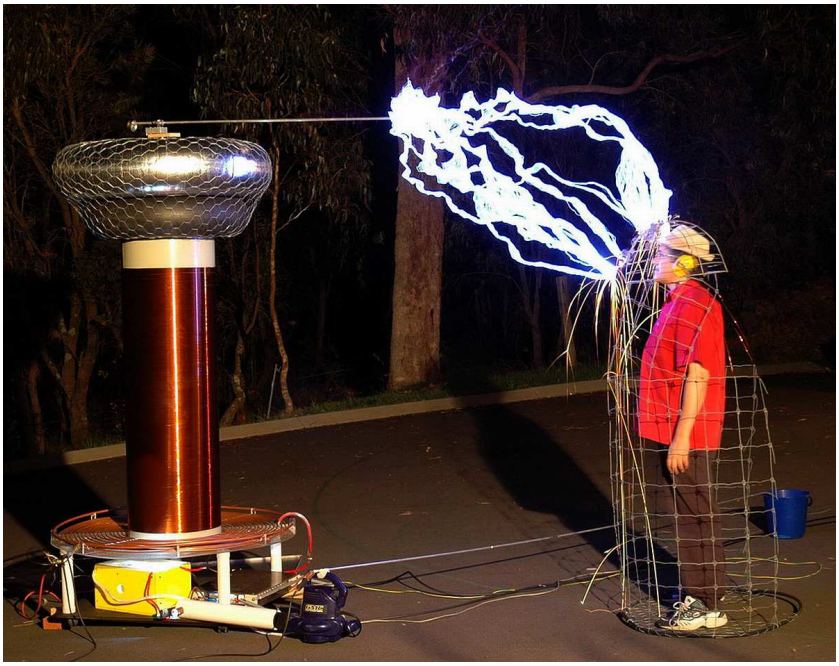


FIG. 3.1. *Electric fields are strongly attenuated inside a Faraday cage (photo from* `http://semperparatus.tk/the-poor-mans-faraday-cage/` [permission must be obtained or a replacement photo found]*). The mathematics of this effect is related to the rapid convergence of the trapezoidal rule for periodic functions.*

by Hölder's inequality, with the 1- and ∞-norms defined over $|z| = r$. By (2.9), this implies $|I_N - I| \le r^{-1}\|u\|_1/(r^N - 1)$. All these observations for Theorem 2.1 and the disk $|z| < r$ also apply with obvious modifications for Theorem 2.2 and the annulus $r^{-1} < |z| < r$.

In closing this section we note that in applications, there is often a symmetry that enables one to cut the number of function evaluations for the trapezoidal rule approximately in half. On the unit circle, this happens if $u(\overline{z}) = \overline{u(z)}$.

3. Physical interpretation: Faraday cage. A wire enclosure can be extremely effective at protecting its contents from electromagnetic fields, as illustrated in Fig. 3.1. This is the *Faraday cage effect,* discovered by Michael Faraday in 1836. This phenomenon is related to the rapid convergence of the periodic trapezoidal rule.

At a descriptive level, we can make the connection like this. It is well known that there can be no electric field inside a metal shell, for the electrons in the shell arrange themselves to achieve a constant potential, and the field is the gradient of the potential. What Faraday discovered is that a wire cage provides a good approximation to this effect, with fields inside that are very close to zero. Just as the sample points of the trapezoidal rule approximate the continuous interval of integration with high accuracy, the wires of the Faraday cage approximate a continuous shell with high accuracy.

To bring some mathematics into the analogy,[5] let us think of the N roots of

[5]We have been guided to some degree by presentations in [45, §7.5] and [188, §7.5.1]; the physics of the Faraday cage leads in many directions and is related for example to Babinet's principle [152]. In

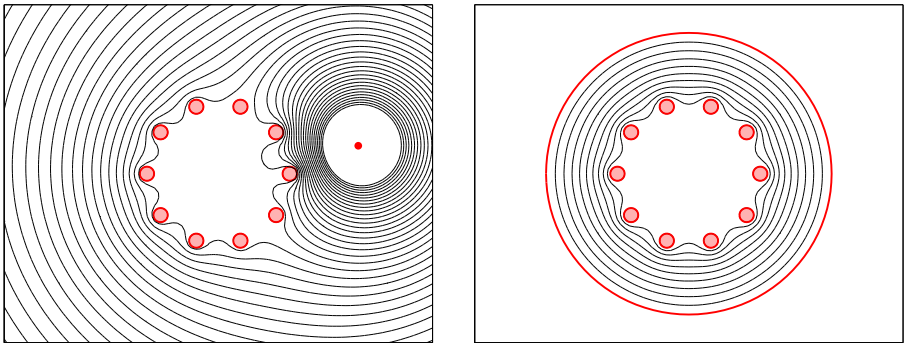


FIG. 3.2. *A 2D model of the Faraday cage effect. In both images,* 10 *disks of radius* 0.1 *are situated at roots of unity with fixed potential* $\varphi(z) = 0$*, and level curves of* $\varphi(z)$ *are plotted. The left image shows the potential generated by a point charge* $\log|z - z_0|$ *with* $|z_0| = 2$*, and the right image the potential generated by the constant value* $\varphi(z) = 1$ *along the circle* $|z| = 2$*. The values at the origin are* $\varphi(0) = 0.000022$ *on the left and* $\varphi(0) = 0.0335$ *on the right. Such values at* $z = 0$ *are mathematically related to the error in the trapezoidal rule at roots of unity.*

unity in Figure 2.1 as a 2D Faraday cage, imagining that these are cross-sections of wires that are connected in the third dimension and hence at equal potential. Now the potential is a real harmonic (as opposed to complex analytic) function $\varphi(z)$, satisfying Laplace's equation $\Delta\varphi = 0$. It does not make sense to specify the value of a harmonic function at a point; we cannot take the wires to have zero radius. Let us instead imagine that they are of a small radius $\sigma > 0$. In Figure 3.2 we see that for $N = 10$ wires with $\sigma = 0.1$, the potential in the interior is very close to constant. A mathematically well-posed question would be, if φ is a harmonic function in the region $|z| < r$ for some $r > 1$ minus disks of radius σ about the Nth roots of unity, with $\varphi(z) \le M$ in this region and $\varphi(z) = 0$ at the boundaries of the disks, then what is the maximum possible value of $\varphi(0)$? This is a question of *harmonic measure,* which could be answered by methods of conformal mapping. The maximum value is attained in the case shown in the right image of the figure, with $\varphi(z) = M$ on $|z| = r$.

The analogy to the trapezoidal rule on the unit circle goes as follows. Here we deal with analytic functions, whose values at points can be specified. Take u to be an integrand analytic and bounded in absolute value by M in $|z| < r$, as in (2.1), let p_{INTERP} be its polynomial interpolant of degree $N - 1$ in the Nth roots of unity, and set $f = p_{\text{INTERP}} - u$. Then f is an analytic function in $|z| < r$ with $f(z) = 0$ at the Nth roots of unity. This implies

$$f(z) = (1 - z^N)h(z) \tag{3.1}$$

for some function h also analytic in $|z| < r$. Now $|h|$ must be smaller on $|z| \le 1$ than on $|z| = r$, by the maximum modulus principle, so $|f|$ must be $O(r^N)$ times smaller throughout the disk $|z| \le 1$ than on $|z| = r$. In particular this applies to the value $f(0)$, which is exactly the error $I_N - I$ in the trapezoidal rule. We shall discuss such estimates further around (8.4) and (8.5).

---

the quadrature literature, a Faraday cage analogy has been mentioned by Korevaar and Meyers [93]. However, the development here is essentially our own and suggests a number of questions for fuller investigation. For example, as N increases with $\sigma \sim C/N$, $\varphi(z)$ appears to be only algebraically small but exponentially close to constant for $|z| < 1$.

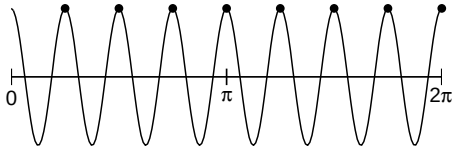


FIG. 4.1. *On an N-point equispaced grid in $[0, 2\pi]$, $e^{iN\theta}$, $e^{2iN\theta}$ and so on are indistinguishable from the constant function* 1. *(This image shows the real part.) The error in the trapezoidal rule for functions periodic on an interval comes from counting a contribution of 2π from each such term instead of the correct contribution of* 0.

It seems clear physically that the wires of a Faraday cage should not have to be exactly evenly spaced for it to be effective; the electric charge will adjust to compensate for irregularities. This suggests that it ought to be possible to generalize the periodic trapezoidal rule to non-equispaced points too, with appropriately modified weights, and still get rapid convergence. This is indeed the case, and such perturbations of the standard configuration are discussed in §10.

4. Integrals over a periodic interval. Having treated integrals over a circle in §2, we now turn to the case of an integral of a function periodic on an interval. This is essentially the same problem, and one can be reduced to the other by a change of variables.

Let v be a real or complex 2π-periodic function on the real line, and define

$$I = \int_0^{2\pi} v(\theta)\, d\theta. \tag{4.1}$$

For any positive integer N, the trapezoidal rule approximation now takes the form

$$I_N = \frac{2\pi}{N} \sum_{k=1}^{N} v(\theta_k), \tag{4.2}$$

where $\theta_k = 2\pi k/N$. The exponential convergence of such approximations is just as in Theorems 2.1 and 2.2. We begin with the analogue of Theorem 2.1, whose hypothesis of analyticity in a disk becomes analyticity in a half-plane. Like Theorem 2.1, this theorem pertains to functions that are not real, unless they are constant. If $v(\theta)$ is real, Theorem 4.2 is the relevant one.

THEOREM 4.1. *Suppose v is 2π-periodic and analytic and satisfies $|v(\theta)| \le M$ in the half-plane $\mathrm{Im}\,\theta > -a$ for some $a > 0$. Then for any $N \ge 1$,*

$$|I_N - I\,| \le \frac{2\pi M}{e^{aN} - 1}, \tag{4.3}$$

and the constant 2π is as small as possible.

Proof as a corollary of Theorem 2.1. The change of variables $z = e^{i\theta}$, $u(z) = v(\theta)$ transplants the problem to $0 < |z| < r$ with $r = e^a$. To apply Theorem 2.1, all we need is to verify that u can be extended to be analytic at $z = 0$, i.e., that the singularity at $z = 0$ is removable. This follows via Cauchy integrals from the boundedness of u; see e.g. [2, Theorem 7, Chapter 4] or [117, Theorem II.1.5]. □

Proof of Theorem 4.1 *by Fourier series and aliasing.* Another approach is to repeat the proof of Theorem 2.1 but now in the transplanted variables, with the

Taylor series turning into a Fourier series. We argue as follows. Since v is analytic (much less smoothness than this would suffice, such as Lipschitz continuity), it has the uniformly and absolutely convergent Fourier series

$$v(\theta) = \sum_{j=-\infty}^{\infty} c_j e^{ij\theta}, \tag{4.4}$$

with coefficients

$$c_j = \frac{1}{2\pi} \int_0^{2\pi} e^{-ij\theta} v(\theta)\, d\theta. \tag{4.5}$$

However, all the coefficients c_j with $j < 0$ are zero, so that we really have

$$v(\theta) = \sum_{j=0}^{\infty} c_j e^{ij\theta}. \tag{4.6}$$

To see that $c_j = 0$ for $j < 0$, we can consider the formula (4.5) for some $j < 0$ and shift the interval $[0, 2\pi]$ up a distance b into the upper half of the complex plane. The contributions from the vertical sides vanish by periodicity, and we get $|c_j| \le Me^{-|j|b}$ for any $b > 0$ and $j < 0$, hence $c_j = 0$. From (4.1) and (4.5) we now have

$$I = \int_0^{2\pi} v(\theta)\, d\theta = 2\pi c_0, \tag{4.7}$$

and from (4.2) and (4.6), interchanging summations as is permitted since the series (4.4) and (4.6) are absolutely convergent, we get

$$I_N = \frac{2\pi}{N} \sum_{j=0}^{\infty} c_j \sum_{k=1}^{N} e^{2\pi ikj/N}. \tag{4.8}$$

The second sum in this formula is equal to N when j is a multiple of N—this is aliasing in its more familiar context (see Fig. 4.1)—and equal to 0 otherwise. Thus (4.8) becomes

$$I_N = 2\pi \sum_{j=0}^{\infty} c_{jN}, \tag{4.9}$$

which together with (4.7) gives

$$I_N - I = 2\pi \sum_{j=1}^{\infty} c_{jN}. \tag{4.10}$$

To derive a bound on the coefficients c_j in this formula, we can shift the interval $[0, 2\pi]$ again in the complex plane. This time we shift it down a distance $a' < a$ into the lower half-plane, where a' may be arbitrarily close to a, and the resulting bound is

$$|c_j| \le Me^{-ja}, \quad j \ge 0. \tag{4.11}$$

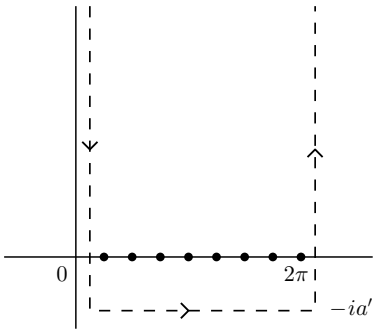


FIG. 4.2. *Integration contour* Γ *for the proof of Theorem* 4.1 *by residue calculus. The contributions from the vertical sides cancel, reducing the calculation to an integral from* $\pi/N - ia'$ *to* $2\pi + \pi/N - ia'$.

Thus (4.10) implies

$$|I_N - I| \le 2\pi M \sum_{j=1}^{\infty} e^{-jaN}, \tag{4.12}$$

and summing this series gives (4.3). The sharpness of the factor 2π can be established by considering the example $v(\theta) = e^{iN\theta}$. □

Proof of Theorem 4.1 *by residue calculus.* The characteristic function

$$m(\theta) = \frac{1}{1 - e^{-iN\theta}} \tag{4.13}$$

has simple poles at the equispaced points $\theta_k = 2\pi k/N$ with residues equal to $-i/N$, i.e., the weights of the trapezoidal rule (4.2) divided by $2\pi i$. By residue calculus, this implies

$$I_N = \int_\Gamma m(\theta) v(\theta) d\theta \tag{4.14}$$

if Γ is a positively oriented contour enclosing the poles in $(0, 2\pi]$. A convenient choice of Γ is the three-segment contour extending from $\pi/N + i\infty$ to $\pi/N - ia'$, then to $2\pi + \pi/N - ia'$, then to $2\pi + \pi/N + i\infty$, for any a' with $0 < a' < a$ (Figure 4.2). Subtracting I gives

$$I_N - I = \int_{\frac{\pi}{N} - ia'}^{2\pi + \frac{\pi}{N} - ia'} \left(\frac{1}{1 - e^{-iN\theta}} - 1 \right) v(\theta) d\theta = \int_{\frac{\pi}{N} - ia'}^{2\pi + \frac{\pi}{N} - ia'} \frac{v(\theta)}{e^{iN\theta} - 1} d\theta, \tag{4.15}$$

where the contributions from the sides of the infinite strip cancel since v is 2π-periodic, implying $|I_N - I| \le 2\pi M/(e^{a'N} - 1)$ for any $a' < a$, and this implies (4.3). □

The next theorem is an analogue of Theorem 2.2, involving an annulus, which now becomes a strip of half-width a.

THEOREM 4.2. *Suppose* v *is* 2π*-periodic and analytic and satisfies* $|v(\theta)| \le M$ *in the strip* $-a < \mathrm{Im}\,\theta < a$ *for some* $a > 0$. *Then for any* $N \ge 1$,

$$|I_N - I| \le \frac{4\pi M}{e^{aN} - 1}, \tag{4.16}$$

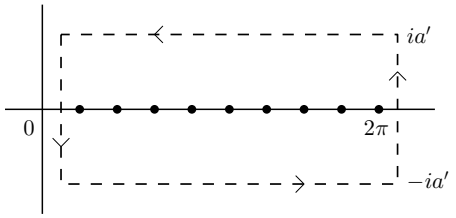


FIG. 4.3. *Integration contour* Γ *for the proof of Theorem* 4.2 *by residue calculus. The periodic half-plane of Figure* 4.2 *now becomes a periodic strip, and the characteristic function* $m(\theta)$ *changes as indicated in equation* (4.19) *and Table* 14.1.

and the constant 4π *is as small as possible.*

Proof by Fourier series and aliasing. The proof of Theorem 2.2 generalizes immediately. Equations (4.6)–(4.9) continue to hold with the limits of summation changed to $j = -\infty$ to ∞. Equation (4.10) becomes

$$I_N - I \;=\; 2\pi \sum_{j=1}^{\infty} (c_{jN} + c_{-jN}), \tag{4.17}$$

and we complete the proof by generalizing (4.11) to

$$|c_j| \le M e^{-|j|a}, \quad -\infty < j < \infty, \tag{4.18}$$

which follows in the case $j < 0$ from lowering $[0, 2\pi]$ to $[0, 2\pi] - ia'$ for a' arbitrarily close to a. □

Proof of Theorem 4.2 *by residue calculus.* The characteristic function

$$m(\theta) = -\tfrac{i}{2} \cot(N\theta/2) = \tfrac{1}{2} \frac{1 + e^{-iN\theta}}{1 - e^{-iN\theta}} \tag{4.19}$$

has simple poles at the equispaced points $\theta_k = 2\pi k/N$ with residues equal to $-i/N$, i.e., the weights of the trapezoidal rule (4.2) divided by $2\pi i$. By residue calculus, this implies

$$I_N = \int_\Gamma m(\theta) v(\theta) d\theta \tag{4.20}$$

if Γ is a positively oriented contour enclosing the poles in $(0, 2\pi]$. This time for Γ we choose the rectangle shown in Figure 4.3, extending from $\pi/N + ia$ to $\pi/N - ia'$, then to $2\pi + \pi/N - ia'$ and $2\pi + \pi/N + ia'$, for any a' with $0 < a' < a$. Over the same contour, the true integral I can be written

$$I = \int_\Gamma \mu(\theta) v(\theta) d\theta, \tag{4.21}$$

where $\mu(\theta)$ is defined by

$$\mu(\theta) = \begin{cases} -\frac{1}{2}, & \mathrm{Im}\,(\theta) > 0\ , \\ \;\;\frac{1}{2}, & \mathrm{Im}\,(\theta) < 0\ . \end{cases} \tag{4.22}$$

The contributions from the ends cancel by periodicity, so subtracting these two formulas gives

$$I_N - I = \tfrac{1}{2}\int_{\frac{\pi}{N}-ia'}^{2\pi+\frac{\pi}{N}-ia'} \left(\frac{1+e^{-iN\theta}}{1-e^{-iN\theta}} - 1\right) v(\theta)\,d\theta - \tfrac{1}{2}\int_{\frac{\pi}{N}+ia'}^{2\pi+\frac{\pi}{N}+ia'} \left(\frac{1+e^{-iN\theta}}{1-e^{-iN\theta}} + 1\right) v(\theta)\,d\theta,$$

that is,

$$I_N - I = -\int_{\frac{\pi}{N}-ia'}^{2\pi+\frac{\pi}{N}-ia'} \left(\frac{1}{e^{iN\theta}-1}\right) v(\theta)\,d\theta - \int_{\frac{\pi}{N}+ia'}^{2\pi+\frac{\pi}{N}+ia'} \left(\frac{1}{1-e^{-iN\theta}}\right) v(\theta)\,d\theta. \tag{4.23}$$

Each integral is readily bounded by $2\pi M/(e^{a'N}-1)$ for any $a' < a$, and this leads to (4.16). □

Theorems 4.1 and 4.2 assume that v has period 2π, but of course, this parameter can be altered. If v has period T instead, (4.16) changes to

$$|I_N - I| \le \frac{2TM}{e^{2\pi aN/T}-1}. \tag{4.24}$$

Equation (4.17) has a corollary concerning what engineers call *band-limited* functions.

COROLLARY 4.3. *If v is a trigonometric polynomial of degree n (a linear combination of $e^{-in\theta},\dots,e^{in\theta}$), the N-point trapezoidal rule* (4.2) *is exact for all $N > n$.*

This result can be interpreted as follows. If a uniform grid samples a periodic function finely enough that all Fourier components are sampled with more than 1 point per wavelength, then the trapezoidal rule is exact for that function. Why do we not need twice as many points, the famous "2 points per wavelength" condition known as the *Nyquist limit*? (We might also have asked this question in connection with Corollary 2.3.) The reason is that although waves that are sampled with between one and two points per wavelength are aliased incorrectly to smoother waves on the grid—e.g., $e^{1.5iN\theta}$ looks like $e^{-0.5iN\theta}$ on the N-point periodic grid in $(0,2\pi]$—these aliases still have nonzero wave number and hence integrate to the correct value of zero. We shall return to this matter in §9 and §13.

As explained in §2, the periodic trapezoidal rule achieves $O(e^{-aN})$ convergence even for unbounded analytic functions in the strip of half-width a, so long as any singularities with $|\mathrm{Im}\, z| = a$ are just simple poles or weaker.

Note that another way to write the exponential decay factor e^{-aN} is as $e^{-2\pi a/h}$, where $h = 2\pi/N$ is the grid spacing between adjacent sample points. This formulation highlights the crucial quantity controlling the exponential convergence of the periodic trapezoidal rule: a/h, the ratio of the half-width of the strip of analyticity to the grid spacing.

5. Example: integral of a periodic entire function. Here is an example of the periodic trapezoidal formula in action. Suppose we want to evaluate the integral

$$I = \int_0^{2\pi} e^{\cos(\theta)}\,d\theta,$$

whose exact value is $I = 2\pi I_0(1) = 7.95492652101284\dots$, where I_0 is a Bessel function. As trapezoidal rule approximations, following (4.2), we can take

$$I_N = \frac{2\pi}{N}\sum_{k=1}^{N} e^{\cos(2\pi k/N)} \tag{5.1}$$

for any positive integer N, and here are some numerical results (rounded):

N	I_N
1	17.1
2	9.7
3	8.23
4	**7.9**89
5	**7.95**83
6	**7.95**520
7	**7.9549**47
8	**7.95492**78
9	**7.9549265**90
10	**7.95492652**45
11	**7.954926521**17
12	**7.95492652101**94

(If symmetry is used, the number of evaluations can be cut approximately in half.) The convergence looks approximately exponential, with each increase of N yielding about one new digit. Actually, it is slightly better than exponential since, unlike the integrand of our opening example (1.1), $\exp(\cos\theta)$ is entire, that is, analytic throughout the complex plane. To apply Theorem 4.2, we note that $|e^{\cosh(\theta)}|$ takes a maximum in the strip $-a \le \mathrm{Im}\theta \le a$ at $a = \pm ia\theta$, where its value is $\cosh(a)$. The theorem accordingly tells us that the bound

$$|I_N - I| \le \frac{4\pi e^{\cosh(a)}}{e^{aN} - 1} \tag{5.2}$$

holds for any value of a. Calculus shows that the bound is minimized for a value of a close to $a = \log(2N)$, giving $\cosh(a) = N + 1/4N$ and

$$|I_N - I| \le \frac{4\pi e^{N+1/(4N)}}{(2N)^N - 1} \sim 4\pi \left(\frac{e}{2N}\right)^N, \quad N \to \infty. \tag{5.3}$$

For $N = 10$, for example, this estimate is 2.7×10^{-8}, not far from the actual error 3.5×10^{-9} in the table. This example is also considered in [178] and [24, Example 8.2.1].

6. Integrals over the real line. Now we turn to the case of integration over the whole real line. In part, this is a small step from periodic functions, with Fourier series replaced by Fourier transforms. New questions arise, however, because for the discussion to make sense, the integrand must decay sufficiently rapidly.

Let w be a real or complex function on the real line, and define

$$I = \int_{-\infty}^{\infty} w(x)\,dx. \tag{6.1}$$

For any $h > 0$, define the trapezoidal rule approximation by

$$I_h = h \sum_{k=-\infty}^{\infty} w(x_k), \tag{6.2}$$

where $x_k = kh$. For the theorems we consider w will be smooth and decay at infinity, so the existence of the integral and the sum will be assured.

Before proving a theorem, let us look at an example, a famous original example in this area:

$$I = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-x^2} dx = 1. \tag{6.3}$$

Here is what we find as h gently diminishes (rounded results):

h	I_h
$2\pi/1$	3.5
$2\pi/2$	**1**.8
$2\pi/3$	**1**.21
$2\pi/4$	**1.0**37
$2\pi/5$	**1.00**39
$2\pi/6$	**1.000**25
$2\pi/7$	**1.0000000**96
$2\pi/8$	**1.00000000**23
$2\pi/9$	**1.0000000000**32
$2\pi/10$	**1.000000000000**28
$2\pi/11$	**1.00000000000000**15
$2\pi/12$	**1.0000000000000000**44

The convergence is amazingly fast, with 4-digit accuracy for $h = 1$ and 16 digits for $h = 0.5$ (the precise rate is given below). The numbers in this table are in principle infinite sums, but since $\exp(-x^2)$ decays rapidly, the series (6.2) can be truncated at quite small $|k|$ with negligible effect. The approximation of the final line can be obtained by adding up just 23 numbers (12 if symmetry is used). The matter of how best to truncate (6.2) is considered in the next section, which analyzes the approximation

$$I_h^{[n]} = h \sum_{k=-n}^{n} w(x_k). \tag{6.4}$$

In §2 and §4 we began with nonsymmetric domains of analyticity (disk, periodic half-plane) before turning to symmetric ones (annulus, periodic strip). Here, we reverse the pattern and begin with the symmetric case.

THEOREM 6.1. *Suppose w is analytic in the strip $|\mathrm{Im}(x)| < a$ for some $a > 0$. Suppose further that $w(x) \to 0$ uniformly as $|x| \to \infty$ in the strip, and for some M, it satisfies*

$$\int_{-\infty}^{\infty} |w(x + ib)| dx \le M \tag{6.5}$$

for all $b \in (-a, a)$. Then, for any $h > 0$, I_h as defined by (6.2) exists and satisfies

$$|I_h - I\,| \le \frac{2M}{e^{2\pi a/h} - 1}, \tag{6.6}$$

and the constant $2M$ is as small as possible.

Note that from the assumption that w decreases to 0 as $|x| \to \infty$ in the strip, it follows by Cauchy integrals that the same holds for all the derivatives of w. It also follows that $w(x+ib)$ is not only in L^1 but also bounded and hence in L^2 too for each b with $|b| < a$. The essential idea of our first proof is related to one of the Paley–Wiener theorems [134]: under appropriate assumptions, a function analytic in a strip has an exponentially decaying Fourier transform.

Outline of proof by Fourier transform and aliasing. Since w is integrable, its Fourier transform

$$\widehat{w}(\xi) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\xi x} w(x)\,dx \tag{6.7}$$

is a uniformly continuous function of $\xi \in \mathbb{R}$. From (6.1) and (6.7) we have

$$I = 2\pi\widehat{w}(0). \tag{6.8}$$

The accuracy of the trapezoidal rule follows from the corresponding formula for I_h, known as the *Poisson summation formula*:

$$I_h = 2\pi \sum_{j=-\infty}^{\infty} \widehat{w}(2\pi j/h). \tag{6.9}$$

A discussion of the Poisson summation formula can be found in [72] and many other places, but generally with stricter hypotheses on w. We do not know a reference that asserts its validity in the present case, so we shall just take (6.9) as given; hence the heading "Outline of proof."

Subtracting (6.8) from (6.9) gives

$$I_h - I = 2\pi \sum_{\substack{j=-\infty \\ j\neq 0}}^{\infty} \widehat{w}(2\pi j/h). \tag{6.10}$$

For a bound on $\widehat{w}(\xi)$, again following §4, we shift the real axis up or down a distance $a' < a$ into the upper or lower half-planes, depending on whether $\xi < 0$ or $\xi > 0$. Here a' may be arbitrarily close to a. Hence one obtains from (6.5) and (6.7)

$$|\widehat{w}(\xi)| \le \frac{1}{2\pi} M e^{-|\xi| a} \tag{6.11}$$

(this is the Paley–Wiener result). Again the assumption $w(x) \to 0$ in the strip ensures that this shift does not change the value of (6.7). Thus (6.10) implies

$$|I_h - I| \le 2M \sum_{j=1}^{\infty} e^{-2\pi j a/h},$$

and summing this series gives (6.6).

To prove sharpness of the constant $2M$, for any h, consider the function

$$w(x) = \frac{\cos(2\pi x/h)}{x^2 + L^2}, \quad L > 0.$$

For this function, $I = (\pi/L)\exp(-2\pi L/h)$ and $I_h = (\pi/L)\coth(\pi L/h)$; hence $I_h - I \sim \pi/L$ as $h \to 0$. On the other hand, for any a with $0 < a < L$,

$$\int_{-\infty}^{\infty} |w(x \pm ia)|dx = \int_{-\infty}^{\infty} \sqrt{\frac{\cos^2(2\pi x/h)\cosh^2(2\pi a/h) + \sin^2(2\pi x/h)\sinh^2(2\pi a/h)}{(x^2 - a^2 + L^2) + 4a^2x^2}}\,dx$$
$$\le \cosh(2\pi a/h) \int_{-\infty}^{\infty} \frac{dx}{\sqrt{(x^2 - a^2 + L^2)^2 + 4a^2x^2}}.$$

The latter integral (which can be evaluated as an elliptic function) is bounded from above by π/L (which corresponds to $a = 0$), so we can take $M = (\pi/L)\cosh(2\pi a/h)$. The error bound in (6.6) now becomes

$$\frac{2M}{e^{2\pi a/h} - 1} = \frac{2(\pi/L)\cosh(2\pi a/h)}{e^{2\pi a/h} - 1},$$

and the right-hand side is asymptotic to π/L in the limits $L \to 0$, $h \to 0$, with $h = o(L)$. □

Proof of Theorem 6.1 *by residue calculus.* The function

$$m(x) = -\frac{i}{2}\cot\left(\frac{\pi x}{h}\right) \tag{6.12}$$

has simple poles at $x = 0, \pm h, \pm 2h, \ldots$, with residues all equal to $h/(2\pi i)$. By residue calculus, this implies that for any integer n,

$$I_h^{[n]} = \int_\Gamma m(x)w(x)\,dx, \tag{6.13}$$

where $I_h^{[n]}$ denotes the truncated trapezoidal rule (6.4) and Γ is a positively oriented contour enclosing the poles in $[-nh, nh]$. A convenient choice of Γ is the rectangular contour with vertices $\pm(n + \frac{1}{2})h + ia'$ and $\pm(n + \frac{1}{2})h - ia'$, for any a' with $0 < a' < a$ (Figure 6.1). Let Γ_+ and Γ_- be the parts of Γ in the upper and lower half-planes, respectively. By Cauchy's theorem, the integral of $w(x)$ on the interval $[-(n+\frac{1}{2})h, (n+\frac{1}{2})h]$ can be evaluated along Γ_- without changing its value; or equivalently, along Γ_+, but with negative orientation. Writing the integral as the average along these two paths and subtracting this from (6.13) gives

$$\begin{aligned} &h\sum_{k=-n}^{n} w(kh) - \int_{-(n+\frac{1}{2})h}^{(n+\frac{1}{2})h} w(x)\,dx \\ &= -\frac{1}{2}\int_{\Gamma_-} (1 + i\cot(\pi x/h))w(x)\,dx + \frac{1}{2}\int_{\Gamma_+} (1 - i\cot(\pi x/h))w(x)\,dx \\ &= -\int_{\Gamma_-} \frac{w(x)}{1 - e^{2\pi i x/h}}\,dx + \int_{\Gamma_+} \frac{w(x)}{1 - e^{-2\pi i x/h}}\,dx. \end{aligned} \tag{6.14}$$

The contributions from the sides of the rectangle Γ vanish in the limit $n \to \infty$. To see this, consider the following integral on one of the four half-sides:

$$\left| \int_{(n+\frac{1}{2})h}^{(n+\frac{1}{2})h+ia'} \frac{w(x)}{1 - e^{-2\pi i x/h}}\,dx \right| \le \frac{1}{2}\int_0^{a'} \left| w((n + \tfrac{1}{2})h + iy) \right| dy,$$

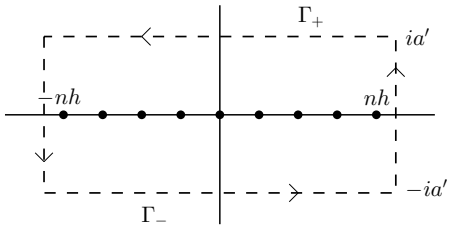


FIG. 6.1. *Integration rectangle* Γ *for the estimation of* $I_h - I$ *in* (6.15)*, with vertices* $\pm(n+\frac{1}{2})h + ia'$ *and* $\pm(n+\frac{1}{2})h - ia'$*. The contributions from the vertical sides vanish in the limit* $n \to \infty$*, reducing the calculation to two integrals, one from* $-\infty - ia'$ *to* $\infty - ia'$ *and the other from* $-\infty + ia'$ *to* $\infty + ia'$*. If* w *is real on the real axis, these two integrals can be combined into one as in* (6.16)*.*

since on the segment of integration, $|1 - \exp(-2\pi ix/h)| = 1 + \exp(2\pi \mathrm{Im}(x)/h) \geq 2$. This vanishes as $n \to \infty$ under the decay assumptions on w. Taking the limit as $n \to \infty$ in (6.14) leads to

$$I_h - I = -\int_{-\infty-ia'}^{\infty-ia'} \frac{w(x)}{1-e^{2\pi ix/h}}\,dx - \int_{-\infty+ia'}^{\infty+ia'} \frac{w(x)}{1-e^{-2\pi ix/h}}\,dx, \tag{6.15}$$

and from here (6.6) follows.

We note that the assertions of the theorem concern the series (6.2), which is properly defined not by a limit of symmetric sums from $-n$ to n as $n \to \infty$, as we have considered, but by a limit of arbitrary sums from $-n_-$ to n_+ as $n_-, n_+ \to \infty$. However, our argument did not use the symmetry, so it establishes the theorem as stated. □

It is straightforward to generalize the error bounds of this theorem to the case where w is analytic in a strip unsymmetric about the real axis, or to specialize it to the case where w is real on the real axis. In the former case, assuming analyticity in the strip $-a_- < \mathrm{Im}(x) < a_+$ for some $a_+, a_- > 0$, one obtains from (6.15)

$$|I_h - I| \leq \frac{M_+}{e^{2\pi a_+/h} - 1} + \frac{M_-}{e^{2\pi a_-/h} - 1}$$

where

$$\int_{-\infty}^{\infty} |w(x + ib_-)|dx \leq M_-, \quad \int_{-\infty}^{\infty} |w(x + ib_+)|dx \leq M_+,$$

for all $b_- \in (-a_-, 0)$ and $b_+ \in (0, a_+)$. We shall use this form of the error bound when we discuss the numerical inversion of the Laplace transform, for there the analytic properties of the integrand in the upper and lower half-planes are quite different (§16).

On the other hand, when w is real on the real axis, symmetry can be used to simplify (6.15) to

$$I_h - I = -2\,\mathrm{Re}\left\{\int_{-\infty+ia'}^{\infty+ia'} \frac{w(x)}{1-e^{-2\pi ix/h}}\,dx\right\}, \tag{6.16}$$

which is the error formula given in [118]. The use of asymptotic methods for estimating this integral was proposed in [65, 172]. For example, for the famous integral (6.3),

it was shown in [65] that the absolute error for small h is approximately $2e^{-\pi^2/h^2}$, which agrees perfectly with the results in the table below (6.3). See also §7.

We state the half-plane variant of Theorem 6.1 without proof.

THEOREM 6.2. *Suppose w is analytic in the half-plane $\mathrm{Im}(x) > -a$ for some $a > 0$, with $w(x) \to 0$ uniformly as $|x| \to \infty$ in that half-plane, and for some M, it satisfies*

$$\int_{-\infty}^{\infty} |w(x+ib)|dx \le M \tag{6.17}$$

for all $b > -a$. Then for any $h > 0$, I_h exists and satisfies

$$|I_h - I| \le \frac{M}{e^{2\pi a/h} - 1}, \tag{6.18}$$

and the constant M is as small as possible.

Here is the analogue for the real line of Corollaries 2.3 and 4.3 concerning band-limited functions.

COROLLARY 6.3. *If w is a function satisfying the conditions of Theorem* 6.1 *whose Fourier transform $\widehat{w}$ has compact support in $[-2\pi/H, 2\pi/H]$ for some $H > 0$, the trapezoidal rule* (6.2) *is exact for all $h \le H$.*

This section has considered integrals over the real line as defined by (6.1). A simple yet powerful variation on this theme is the Hilbert transform integral

$$I = \text{p.v.} \int_{-\infty}^{\infty} \frac{w(x)}{x} dx, \tag{6.19}$$

where p.v. denotes the principal value, which arises in numerous applications [74, Chapter 14]. The trapezoidal rule converges at essentially the same rate for (6.19) as for (6.1), provided the singularity at $x = 0$ is placed midway between two grid points, so that the trapezoidal rule becomes the midpoint rule. Kress and Martensen [100] prove a theorem to this effect by subtracting $w(0)\exp(-x^2)/x$ from the integrand to remove the singularity. The Hilbert transform on the unit circle is considered in [52, 145].

7. Optimal step sizes and convergence rates for the real line. Theorems 6.1 and 6.2 are based on the infinite sum (6.2), whose evaluation in principle involves an infinite amount of work. Provided the integrand decays rapidly, it is common practice to truncate the sum at a value of $|k|$ determined by the rate of decay of $w(x)$ as $|x| \to \infty$. To determine the optimal truncation point, we define the truncated approximation $I_h^{[n]}$ as in (6.4) and write

$$|I_h^{[n]} - I| \le |I_h - I| + |I_h - I_h^{[n]}|.$$

The *discretization error* $|I_h - I|$ can be estimated with Theorems 6.1 or 6.2, and decreases as $h \to 0$. The *truncation error* $|I_h - I_h^{[n]}|$ can be estimated as the maximum of the magnitudes of the integrand at the truncation points $x = \pm hn$, provided the terms in the trapezoidal sum decay sufficiently rapidly. Estimates for the optimal truncation point can accordingly be obtained by balancing the orders of magnitude of the discretization and truncation error estimates. Of course it is not necessary to truncate the infinite sum symmetrically at $\pm nh$ as in (6.4), but we consider just this case for simplicity.

TABLE 7.1

Optimal convergence rates and step sizes in the numerical calculation of $\int_{-\infty}^{\infty} w(x)\,dx$ *by the truncated trapezoidal rule* (6.4)*, for three typical functions* $w(x)$*. These are prototypical examples of functions that are analytic in a strip with exponential decay, analytic in a strip with Gaussian decay, and entire with Gaussian decay.*

$w(x)$	Disc. Error	Trunc. Error	Optimal h	Optimal rate
$(1+x^2)^{-1}\exp(-x\tanh x)$	$O(e^{-2\pi/h})$	$O(e^{-nh})$	$(2\pi/n)^{1/2}$	$O(e^{-(2\pi n)^{1/2}})$
$(1+x^2)^{-1/2}\exp(-x^2)$	$O(e^{-2\pi/h})$	$O(e^{-(nh)^2})$	$(2\pi/n^2)^{2/3}$	$O(e^{-(2\pi n)^{2/3}})$
$\exp(-x^2)$	$O(e^{-\pi^2/h^2})$	$O(e^{-(nh)^2})$	$(\pi/n)^{1/2}$	$O(e^{-\pi n})$

It is interesting to highlight three special cases of interest, as summarized in Table 7.1. The first two of these functions both have singularities at $x = \pm i$, and hence the theory of §4 can be used to establish a discretization error of order $O(e^{-2\pi/h})$ in each case. The third function has no singularities in the finite complex plane, and finding the discretization error involves a minimization argument already used in §5. Brief details are given below. For the second and third examples in the table, $w(x)$ exhibits Gaussian decay as $|x| \to \infty$, while the first decays only exponentially.

By balancing the estimates for the discretization and truncation errors in Table 7.1, formulas for the optimal step size and the optimal rate of convergence can be derived. Note how the final column shows increasingly fast convergence, with exponents proportional to $n^{1/2}$, $n^{2/3}$, and n as the integrands grow better behaved. Figure 7.1, based on high precision computations in Maple, confirms these predictions.

The third function in the table, $w(x) = \exp(-x^2)$, is a famous example in this context [65, 125]. Its discretization error can be derived by noting that for any $a > 0$,

$$|I_h - I| = O(e^{a^2 - 2\pi a/h}), \quad h \to \infty,$$

where we have used the fact that the quantity M in Theorem 6.1 is equal to $\exp(a^2)$. Following reasoning as in (5.2)–(5.3), we see that the choice of a that yields the tightest estimate is $a = \pi/h$, leading to $|I_h - I| = O(e^{-\pi^2/h^2})$, as listed in the table.

If the integrand $w(x)$ does not decay at least exponentially as $|x| \to \infty$, the truncated trapezoidal rule will be less efficient. For example, for the Runge function $w(x) = (1+x^2)^{-1}$, it takes thousands of points in (6.4) to achieve a few digits of accuracy. In such cases there are two courses of action. First, it is often possible to convert a slowly decaying integrand into a quickly decaying one by an appropriate change of variables. Such transformation methods are discussed in §15. A particular one, a sinh mapping, is mentioned at the end of §15 and applied to solve a PDE in §17. An alternative strategy for dealing with slowly decaying integrands is to apply convergence acceleration to the sequence of partial sums of (6.4) [68]. This effectively eliminates the truncation error, so that the total error is dominated by the discretization error.

8. Polynomial, trigonometric, and sinc interpolation. We have followed two routes to the exponentially convergent trapezoidal rule: series (or transforms, in the case of the real line), and contour integrals. A third idea is to make use of the following property: *the trapezoidal rule approximation is the exact integral of an interpolant through the given data values.* On the unit circle, the interpolant is a polynomial or a Laurent polynomial. On the periodic interval $[-\pi, \pi]$, it is a one-sided or two-sided trigonometric polynomial, i.e., a linear combination of functions

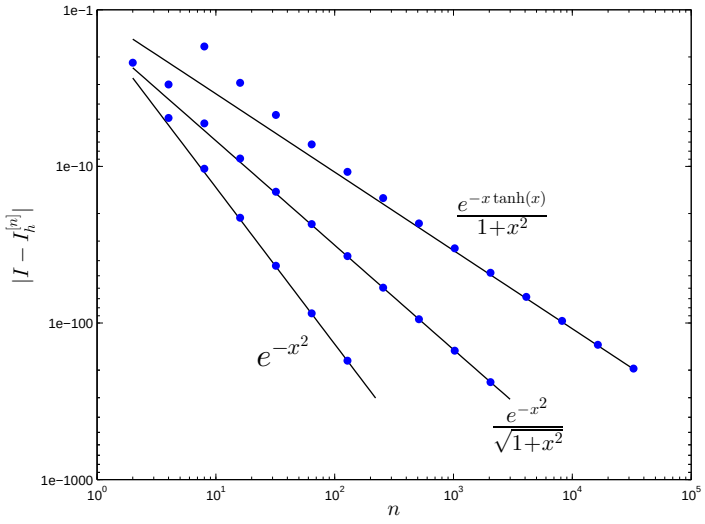


FIG. 7.1. *Actual (dots) and estimated (solid) convergence curves for the three examples of Table* 7.1.

e^{ikx} for various integers k either of a single sign or both positive and negative. On the real line, it is an infinite series of translates of sinc functions. These ideas are sketched in Figure 8.1.

This section is unfortunately rather dry, as it seems desirable to set down six versions of the formulas corresponding to three variants of the trapezoidal rule (complex unit circle, periodic interval, and real line) and two types of domains of analyticity (nonsymmetric and symmetric). Though individual results of this kind can be found here and there in the literature, we do not know any previous publications where such information is collected systematically.

As usual, we organize the presentation by beginning with the case of a function on the unit circle. Consider the context of Theorem 2.1, a function $u(z)$ analytic and bounded in absolute value by M in the disk $|z| < r$ with $r > 1$. One approximation by a polynomial of degree $\le N-1$ is obtained by truncating the Taylor series (2.4),

$$p_{\text{TAYLOR}}(z) = \sum_{j=0}^{N-1} c_j z^j. \tag{8.1}$$

It follows from the estimate (2.10) that the error in this approximation is bounded by

$$\|p_{\text{TAYLOR}} - u\| \le \frac{Mr^{-N}}{1 - r^{-1}}, \tag{8.2}$$

where $\|\cdot\|$ denotes the supremum norm over $|z| = 1$ (or $|z| \le 1$, by the maximum modulus principle). Another approximation of u would be the unique polynomial of degree $\le N-1$ that interpolates u in the roots of unity $\{z_k\}$. It is possible to write down an explicit formula for the coefficients of this interpolant by noting that it can be derived by aliasing z^N, $z^{2N}, \ldots$ to z^0; z^{N+1}, $z^{2N+1}, \ldots$ to z^1; and so on. The result is this formula:

$$p_{\text{INTERP}}(z) = \sum_{j=0}^{N-1} \left(\sum_{k=0}^{\infty} c_{j+kN} \right) z^j. \tag{8.3}$$

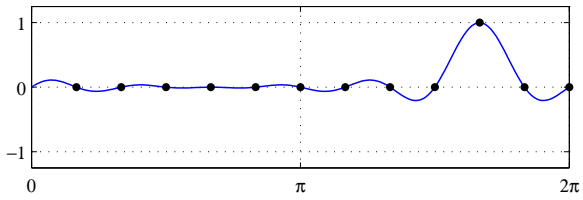


FIG. 8.1. *Any set of data at equispaced points can be interpolated by a trigonometric polynomial, and if the trapezoidal rule is applied to the data, the result I_N is equal to the integral of the interpolant. Here the data correspond to a Kronecker delta function on a grid with $N = 12$, the interpolant is a periodic sinc function, and the integral is equal to $h = 2\pi/N$. Interpolants to more general data can be regarded as linear combinations of translates of such periodic sinc functions. Analogous observations hold for the trapezoidal rule on the unit circle or the real line.*

For this approximation, (2.10) implies

$$\|p_{\text{INTERP}} - u\| \le \frac{2Mr^{-N}}{1 - r^{-1}}. \tag{8.4}$$

Examining (8.2) and (8.4), we see that both polynomial interpolants in roots of unity and truncated Taylor series are enough to establish the result that functions analytic on the closed unit disk can be approximated by polynomials with exponentially decreasing error as $N \to \infty$.[6] What is special about p_{INTERP} for our purposes is that when the trapezoidal rule is applied to u, the resulting approximation I_N is exactly the integral of p_{INTERP}, as can be derived from (2.8) and (8.3). Thus (8.4) yields another proof of the geometric convergence of the trapezoidal rule: if we multiply (8.4) by the circumference 2π, we find

$$|I_N - I| \le \frac{4\pi M}{r^N - r^{N-1}}. \tag{8.5}$$

This bound is similar to (2.3) but weaker by a factor slightly greater than 2.

What happens if u is analytic just in the annulus $r^{-1} < |z| < r$? Now the natural approximations to consider are *Laurent polynomials,* involving both negative and positive powers of z. Assuming N is odd for simplicity, the analogue of the Taylor projection p_{TAYLOR} is the Laurent projection

$$p_{\text{LAURENT}}(z) = \sum_{j=(1-N)/2}^{(N-1)/2} c_j z^j, \tag{8.6}$$

in which terms of degree $|k| \ge N/2$ of the Laurent series of u are dropped, satisfying

$$\|p_{\text{LAURENT}} - u\| \le \frac{2Mr^{-(N+1)/2}}{1 - r^{-1}}. \tag{8.7}$$

If these terms of the Laurent series are aliased to $|k| < N/2$ rather than dropped, we obtain a Laurent polynomial interpolant analogous to p_{INTERP},

$$p_{\text{INTERP2}}(z) = \sum_{j=(1-N)/2}^{(N-1)/2} \left(\sum_{k=-\infty}^{\infty} c_{j+kN} \right) z^j, \tag{8.8}$$

[6] Convergence of polynomial interpolants in roots of unity goes back to Runge [140, §II.15, pp. 136–137]. For Chebyshev polynomial analogues of these developments on $[-1, 1]$, see [169, Chapters 4 and 8].

with

$$\|p_{\text{INTERP2}} - u\| \le \frac{4Mr^{-(N+1)/2}}{1 - r^{-1}}. \tag{8.9}$$

Multiplying (8.9) by the circumference 2π yields

$$|I_N - I| \le \frac{8\pi M}{r^{(N+1)/2}(1 - r^{-1})}. \tag{8.10}$$

Perhaps it is worth noting explicitly that when the trapezoidal rule is applied to a set of data at roots of unity, the formula involved has no notion of whether the data come from a function that is analytic in a disk, analytic in an annulus, or not analytic at all. The quadrature result I_N is just a number computed from the data, and the bounds (8.5) and (8.10) differ not because I_N differs in the two cases but because different conclusions about its accuracy can be reached based on different hypotheses.

In passing from (8.2), (8.4), and (8.5) to (8.7), (8.9), and (8.10), the convergence rate has approximately halved, from r^{-N} to $r^{-(N+1)/2}$, so that twice as many points are needed for a good approximation of u on an annulus as on a circle. Yet from Theorems 2.1 and 2.2 we know that the actual rate of convergence of $|I_N - I|$ to zero is the same in both cases, r^{-N}. The distinction is genuine: moving from the disk to the annulus halves the rate of convergence of the interpolants without halving the rate of convergence of the trapezoidal rule. This sounds paradoxical, but it can be explained by returning to the observation mentioned in §4 in connection with the Nyquist limit of 2 points per wavelength. If u is analytic just in an annulus rather than a disk, then we need negative powers of z as well as positive ones for accurate interpolation, halving the rate of convergence of the interpolants as $N \to \infty$. However, the components $z^{(N+1)/2}, z^{(N+3)/2}, \dots, z^{N-1}$ of $u(z)$ that are now mishandled by our interpolant p_{INTERP2} contribute nothing to the integral, so the rate of convergence of the trapezoidal rule is unaffected.

For the remainder of this section we spell out analogues of the above results, first for functions periodic on an interval, then for functions on the real line. It is worthwhile to record these formulas, but the reader with no special interest in them is encouraged to turn to the next section.

Here are the analogues of (8.1)–(8.10) for periodic functions on an interval. For a function $v(\theta)$ analytic and bounded by M in the half-plane $\text{Im}(\theta) > -a$ with $a > 0$, an approximation by a trigonometric polynomial of degree $\le N - 1$ is obtained by truncating the Fourier series (4.6),

$$q_{\text{TRIG}}(\theta) = \sum_{j=0}^{N-1} c_j e^{ij\theta}. \tag{8.11}$$

For this approximation, (4.11) gives

$$\|q_{\text{TRIG}} - v\| \le \frac{Me^{-aN}}{1 - e^{-a}}, \tag{8.12}$$

where $\|\cdot\|$ is the supremum norm over $[0, 2\pi]$. Alternatively, another approximation is the unique trigonometric polynomial of degree $\le N - 1$ that interpolates v in the points $\{\theta_k\}$, i.e., the trigonometric polynomial obtained by aliasing e^{iNx} to 1, $e^{i(N+1)x}$

to e^{ix}, and so on:

$$q_{\text{INTERP}}(\theta) = \sum_{j=0}^{N-1} \left(\sum_{k=0}^{\infty} c_{j+kN} \right) e^{ij\theta}, \tag{8.13}$$

with

$$\| q_{\text{INTERP}} - v \| \le \frac{2Me^{-aN}}{1 - e^{-a}}. \tag{8.14}$$

The trapezoidal rule applied to v will give the exact integral over $[0, 2\pi]$ of q_{INTERP}, implying

$$|I_N - I| \le \frac{4\pi M}{e^{aN}(1 - e^{-a})}, \tag{8.15}$$

as may be compared with (4.3). Again we may ask, what happens if v is analytic just in a strip $-a < \text{Im}(\theta) < a$? Now we must consider two-sided trigonometric polynomials. Assuming again for convenience that N is odd, we get

$$q_{\text{TRIG2}}(\theta) = \sum_{j=(1-N)/2}^{(N-1)/2} c_j e^{ij\theta}, \tag{8.16}$$

with

$$\| q_{\text{TRIG2}} - v \| \le \frac{2Me^{-a(N+1)/2}}{1 - e^{-a}}, \tag{8.17}$$

and

$$q_{\text{INTERP2}}(z) = \sum_{j=(1-N)/2}^{(N-1)/2} \left(\sum_{k=-\infty}^{\infty} c_{j+kN} \right) e^{ij\theta}, \tag{8.18}$$

with

$$\| q_{\text{INTERP2}} - v \| \le \frac{4Me^{-a(N+1)/2}}{1 - e^{-a}}, \tag{8.19}$$

implying

$$|I_N - I| \le \frac{8\pi M}{e^{a(N+1)/2}(1 - e^{-a})}, \tag{8.20}$$

to be compared with (4.16).

Finally there is the case of the trapezoidal rule on the real line. Here, given a function $w(x)$ analytic and with integrals bounded by M in the half-plane $\text{Im}(x) > -a$ with $a > 0$, as in Theorem 6.2, we may truncate the appropriate Fourier integral at $\xi = 2\pi/h$ to get

$$r_{\text{FOURIER}}(x) = \int_0^{2\pi/h} \widehat{w}(\xi) e^{i\xi x} d\xi, \tag{8.21}$$

for which (6.11) gives

$$\|r_{\text{FOURIER}} - w\| \le \frac{M}{2\pi a} e^{-2\pi a/h}, \tag{8.22}$$

where $\|\cdot\|$ is the supremum norm over $\mathbb{R}$. Alternatively, another approximation is the unique function band-limited to wave numbers $[0, 2\pi/h]$ that interpolates w in the points $\{x_k\}$:

$$r_{\text{SINC}}(x) = \int_0^{2\pi/h} \left(\sum_{k=0}^{\infty} \widehat{w}(\xi + 2\pi k/h) \right) e^{i\xi x} d\xi, \tag{8.23}$$

for which (6.11) implies

$$\|r_{\text{SINC}} - w\| \le \frac{M}{\pi a} e^{-2\pi a/h}. \tag{8.24}$$

Since $\mathbb{R}$ is of infinite length, however, this does not lead directly to a bound on $|I_h - I|$ to be compared with (6.6). If w is analytic just in the strip $-a < \text{Im}(x) < a$, as in Theorem 6.1, these formulas become

$$r_{\text{FOURIER2}}(x) = \int_{-\pi/h}^{\pi/h} \widehat{w}(\xi) e^{i\xi x} d\xi, \tag{8.25}$$

with

$$\|r_{\text{FOURIER2}} - w\| \le \frac{M}{\pi a} e^{-\pi a/h} \tag{8.26}$$

and

$$r_{\text{SINC2}}(x) = \int_{-\pi/h}^{\pi/h} \left(\sum_{k=-\infty}^{\infty} \widehat{w}(\xi + 2\pi k/h) \right) e^{i\xi x} d\xi, \tag{8.27}$$

with

$$\|r_{\text{SINC2}} - w\| \le \frac{2M}{\pi a} e^{-\pi a/h}. \tag{8.28}$$

Again there is no bound on $|I_h - I|$ that results directly from these estimates.

The function r_{SINC2} is known as the *cardinal interpolant* or *sinc interpolant* to f on the equispaced grid $h\mathbb{Z}$. This function can be obtained as a limit of periodic trigonometric interpolants on wider and wider intervals of periodicity, or as a series with coefficients derived from central difference formulas on wider and wider stencils, and its properties have been studied since E. T. Whittaker in 1915 [184]. A simple interpretation of $r_{\text{SINC2}}(x)$ is that it is the sum of an infinite series of translates of the sinc function $\sin(\pi x/h)/(\pi x/h)$. For a leisurely account of the mathematics, see [167, Chapter 2]. In the periodic case, one can interpret q_{TRIG2} in the same way, except that the series is a finite one involving periodic sinc functions, which take the form $\sin(\pi x/h)/((2\pi/h)\tan(x/2))$ when N is even as illustrated in Figure 8.1. See [96] and [167, Chapter 3].

9. Connections with Gauss and Clenshaw–Curtis quadrature. First in §4, and again repeatedly in §8, we have mentioned a certain factor of 2 property for the trapezoidal rule on a circle or a periodic interval: the formula requires just "one point per wavelength" for exactness, not two. Another way to put it is that the accuracy of the N-point quadrature formula is approximately equal to the accuracy of the trigonometric interpolant through not N but $2N$ points. We shall now sketch how this can be viewed as the same advantage enjoyed by Gauss (more precisely Gauss–Legendre) quadrature for a nonperiodic function defined on $[-1,1]$ in comparison with other methods for such functions such as Clenshaw–Curtis quadrature.

Consider for example the case of a periodic function v on $[0,2\pi]$. On an equispaced grid of N points, with N taken to be odd for simplicity, the trapezoidal rule can be interpreted as interpolating v by a trigonometric polynomial of degree $(N-1)/2$, and then integrating the interpolant exactly (§8). Thus if v is given by a convergent Fourier series, the terms in the series of orders $(N+1)/2$, $(N+3)/2, \dots$ can be viewed as aliased by the grid to orders $(-N+1)/2$, $(-N+3)/2, \dots$. However, though these terms are aliased, they are still integrated correctly by the trapezoidal rule to zero, until one reaches the term of order N, which aliases to 0 and erroneously gets a nonzero integral.

Gauss quadrature has an analogous property. Suppose we have a function $y(\xi)$ defined on $[-1,1]$. On an N-point Gauss grid in $[-1,1]$, i.e., the grid of roots of the Legendre polynomial P_N, Gauss quadrature can be interpreted as interpolating y by a polynomial of degree $N-1$, then integrating the interpolant exactly. Thus if y is given by a convergent series of Legendre polynomials, the terms in the series of degree N, $N+1$, and so on can be viewed as aliased to lower degrees, when sampled on the grid. More precisely, from the three-term recurrence relation associated with these orthogonal polynomials one can see that P_N is aliased to zero, P_{N+1} to a multiple of P_{N-1}, P_{N+2} to a linear combination of P_{N-2} and P_{N-1}, and in general, for any $k \le N$, P_{N+k} to a linear combination of $P_{N-k}, \dots, P_{N-1}$. Again, however, these aliases introduce no errors for $k < N$, because the alias of each Legendre polynomial integrates to the correct value of zero. This is because all Legendre polynomials other than P_0 have zero integrals over $[-1,1]$, since they are orthogonal to P_0. The first error is introduced at $k = N$, where we find that P_{2N} ought to integrate to zero but is aliased to a sum of Legendre polynomials that includes a nonzero component of P_0.

We can contrast this behavior with what happens with Clenshaw–Curtis quadrature on $[-1,1]$, that is, the formula obtained by applying the trapezoidal rule on $[0,2\pi]$ to the function $v(\theta)$ obtained by transplanting $y(\xi)$ under the change of variables $\xi = \cos(\theta)$. In the ξ variable on $[-1,1]$, this is equivalent to sampling at the Chebyshev points $\cos(\theta_k)$, interpolating by a sum of Chebyshev polynomials T_k, and integrating the interpolant exactly. As in the Gauss–Legendre case, T_{N+k} aliases to a linear combination of $T_{N-k}, \dots, T_{N-1}$; actually, the situation is even simpler, for T_{N+k} aliases to T_{N-k} alone. The difference is that this time, an integration error is introduced by this aliasing, since T_{N+k} and T_{N-k} have nonzero and distinct integrals (when $N+k$ is even). This is why Clenshaw–Curtis quadrature has half the polynomial order of accuracy as Gauss quadrature.[7]

[7] The full story is subtler than this, however, for although T_{N-k} has nonzero integral, the integral is of size $O(N^{-3})$ for fixed k as $N \to \infty$ and of size $O(N^{-2})$ in a certain sense averaged over all k. As a consequence Clenshaw–Curtis quadrature, despite its lower polynomial order of accuracy, is in practice nearly as accurate as Gauss quadrature in many cases, with errors often closer to the lower dots than the upper ones in plots like those of Figure 9.1. See [132, 168, 182, 186].

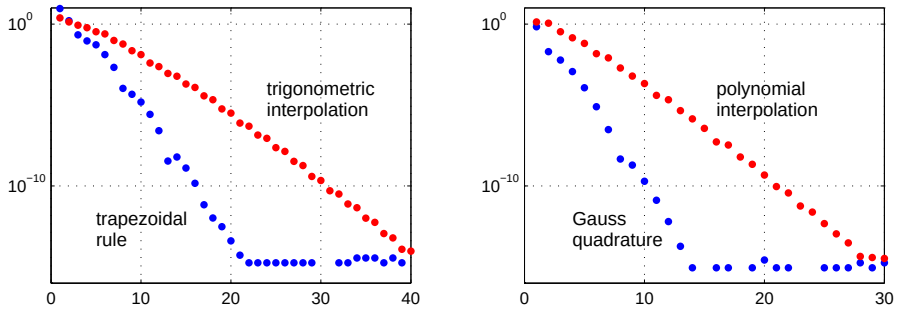


FIG. 9.1. *For analytic functions periodic on* $[0, 2\pi]$*, the trapezoidal rule for quadrature converges asymptotically at twice the rate of the underlying approximation by trigonometric interpolation. The image on the left shows this effect for the example* $v(\theta) = \exp(\cos(\theta) + \sin(2\theta)/4)$*. On the right, the analogous behavior for the function* $y(\xi) = \exp(\cos(\xi) + \sin(2\xi)/4)$ *on* $[-1, 1]$*, comparing Gauss quadrature with polynomial interpolation in Legendre points. Both computations are carried out in floating point arithmetic, so the errors level off at machine precision. Missing dots correspond to errors that cancel to zero in floating point.*

In summary, for $|z| = 1$ or $\theta \in [0, 2\pi]$, the trapezoidal rule *is* Gauss quadrature: it combines algebraic simplicity with the characteristic Gauss quadrature factor of 2 bonus in formal order of accuracy. This fast rate of convergence is summarized in Figure 9.1. For nonperiodic functions on $[-1, 1]$, one can have either the algebraic simplicity (Clenshaw–Curtis) or the extra order of accuracy (Gauss), but not both. And now we turn to another context in which the factor of 2 is lost.

10. Unevenly spaced quadrature points. The periodic trapezoidal rule is defined by equispaced points and equal quadrature weights. However, its exponential convergence for analytic integrands does not depend on equal spacing. The convergence is still exponential for certain periodic non-equispaced point sets, provided the quadrature weights are adjusted according to the design principle of §8: interpolate the data by a trigonometric polynomial, then integrate the interpolant. Let us call the resulting periodic quadrature formula for nonuniform points *quadrature by trigonometric interpolation.* Note that this is different from the more elementary trapezoidal rule for unevenly spaced points, based on connecting the data by trapezoids, whose accuracy is just $O(h^2)$.

We begin with a numerical illustration. The upper image of Figure 10.1 shows the periodic integrand $v(\theta) = e^{\cos(\theta)}$ of §5 sampled on a 12-point trapezoidal grid, the same grid sketched in Figure 8.1. The error listed for the trapezoidal rule reveals 11-digit accuracy for this problem. The lower image of the figure shows the same result for a perturbed grid (the 3rd, 8th, and 9th points have moved moved left by distances 0.2, 0.4, and 0.2, respectively). Now the accuracy cuts in half, to 6 digits. In other words, the quadrature now approximately matches the accuracy of the trigonometric interpolant, as described in the last section. Perturbing the grid has eliminated the "Gauss quadrature factor of 2," but it has not eliminated the exponential convergence.

It is interesting to note the analogous situation with a Faraday cage made from a wire mesh, as discussed in §3. The effectiveness of the cage does not depend on the mesh having equispaced wires, for any nonuniformity in the spacing is compensated by nonuniformity in the electric charge distribution.

To make these observations precise, ideally, we would state generalizations of equations (8.19)–(8.20) for quadrature by trigonometric interpolation. We are not

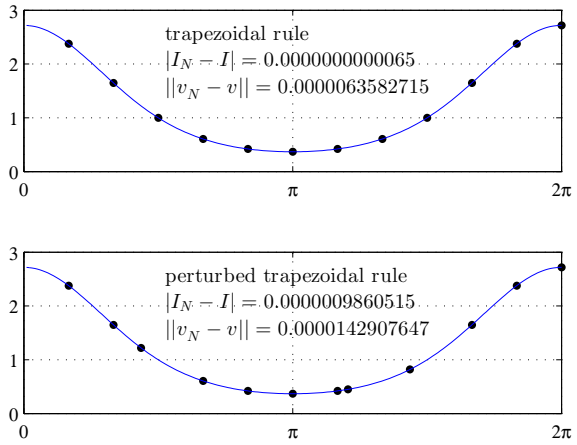


FIG. 10.1. *If the nodes of the periodic trapezoidal rule are perturbed, one still gets exponential convergence for analytic integrands, but at a rate cut in half. This observation holds in theory for all perturbations that leave the asymptotic distribution of quadrature points uniform (Theorem* 10.1*). However, if the quadrature rule is to be useful in practice and converge for less smooth integrands, or in the presence of rounding errors, additional conditions are needed to keep the points apart.*

aware of such generalizations, but let us summarize some of what is known that is relevant to this problem. It is not clear that anything has been published that explicitly addresses quadrature by trigonometric interpolation, but there are two bodies of literature that have direct implications for this problem.

The first concerns interpolation of analytic functions and features theorems that might be regarded as mainly of theoretical interest, because they are unstable with respect to perturbations. Like so much in the field of interpolation, the idea goes back to Runge [140] and his use of the Hermite integral formula for estimating accuracy of polynomial interpolants.

Let $\{z_k^{(N)}\}$ be a family of N-point grids on the unit circle for each $N \geq 1$, and, following a line of work going back to Weyl [183], suppose that as $N \to \infty$, these grids are *uniformly distributed* in the sense that the fraction of points on any arc of the circle of length L converges as $N \to \infty$ to $L/2\pi$. There is no assumption that the points are well separated or even distinct. Let $u(z)$ be analytic in $|z| \leq r$ for some $r > 1$, and let p_{N-1} be the degree $N-1$ polynomial interpolant to u in the grid $\{z_k^{(N)}\}$. (If two nodes are equal, one interpolates both f and f' there, and similarly for confluences of higher order.) Then the interpolants satisfy $\|u - p_{N-1}\| = O(r^{-N})$ as $N \to \infty$. This result is generally attributed to Fejér in 1918 [43], though Fejér stated it without proof, commenting that he was building on ideas of Sierpiński and Weyl and would publish a proof on another occasion. (Fejér's main concern was generalizations by conformal mapping to regions other than the disk.) Perhaps a proof was first spelled out by Kalmár in 1926 together with a converse to the effect that $O(r^{-N})$ convergence for all functions analytic in $|z| \leq r$ implies uniform distribution of the grids [87]. So the result is sometimes known as the Fejér–Kalmár theorem; see [34, 53] for discussions.

The same conclusion carries over from a disk to an annulus or a periodic interval: analyticity and uniform distribution of the interpolation points are enough to guarantee convergence at the expected geometric rate. For the case of an annulus, see [76, Theorem 3]. The following statement for a periodic interval can be derived from [90,

Theorem 5].

THEOREM 10.1. *Let v be a 2π-periodic function on the real line that is analytic in the strip $-a < \mathrm{Im}\,(\theta) < a$ for some $a > 0$, let $\{\theta_k^{(N)}\}$ be a uniformly distributed system of grids in $[0, 2\pi]$ as defined above, and let $\{I_N\}$ be the results of quadrature by trigonometric interpolation applied in these points. Then*

$$\limsup_{N\to\infty} |I_N - I|^{1/N} \le e^{-a/2}. \tag{10.1}$$

A theorem is a theorem, but it is clear that the convergence promised by this result must be unstable. The condition of uniform distribution does not prevent grid points coalescing, and if this happens, a vanishing denominator will appear in any interpolation formula. In the standard analysis of approximation theory, this effect shows up as Lebesgue constants for the interpolation process that satisfy no finite bound, even for finite N [169].[8] Mathematically, this will not affect the accuracy of the interpolant to an analytic function, but interpolation of non-analytic functions will be another matter, or interpolation of analytic functions in the presence of rounding errors. The situation is analogous to polynomial interpolation of entire functions in equispaced grids on $[-1, 1]$, where one has exponential convergence in theory but exponential divergence in practice.

Thus uniform distribution of the interpolation points is not enough for robust interpolation. And yet it is clear from Figure 10.1 that interpolation in perturbed points will sometimes be highly accurate. This leads us to the question, what degree of perturbation of $\{\theta_k^{(N)}\}$ from the equispaced configuration can be permitted? There does not seem to be a literature on this question from the perspective of approximation theory or quadrature, but there is a related literature on *sampling theory* going back many years. A basic question in this area is, what properties of an interpolatory process suffice to guarantee that it will reproduce a sampled function exactly? Our Corollaries 2.3, 4.3, and 6.3 concerning band-limited functions are of this form, and such results have been generalized to irregular grids in work beginning with Paley and Wiener, Levinson, and Kadec, among others. These and other references are given in [5]. When stability is brought into the picture, key concepts that arise are the notions of *Riesz bases* and *frames,* which are related to boundedness of the reconstruction process in L^2 and other norms. Closely related is the literature of algorithms associated with the so-called *nonuniform Fast Fourier Transform* [39].

The emphasis on band-limited functions, exact reconstruction, and integral norms gives the literature of sampling theory a different style from our interest here in the rate of convergence of the trapezoidal rule for smooth but nonband-limited functions. We do not know what analogues there are of Theorem 10.1 for quadrature by trigonometric interpolation with a robustness condition.

11. The Euler–Maclaurin formula. When people set out to explain the fast convergence of the trapezoidal rule for periodic integrands, an argument commonly used is based on the Euler–Maclaurin formula. Indeed, this was Poisson's approach in the 1820s, for the formula was already well established in his day, having originated independently with Euler and Maclaurin in the 1730s and 1740s. By such reasoning one can see that the periodic trapezoidal rule has accuracy $O(N^{-m})$ for any m

[8] By contrast, the interpolation processes underlying the theorems on fast convergence for analytic functions in the other sections of this paper, such as Theorems 2.1–2.2, 4.1–4.2, and 6.1–6.2, are robust, with Lebesgue constants that are finite for each N and grow only slowly as $N \to \infty$.

when applied to a C^∞ function, though it is not easy to see that the accuracy is $O(\exp(-CN))$ if the function is analytic.

Returning to the setting of §4, let v be a continuous function on the real line, which initially we do not require to be periodic, and define

$$I = \int_0^{2\pi} v(\theta)\,d\theta, \qquad I_N = h{\sum_{k=0}^{N}}' v(kh),$$

where $h = 2\pi/N$ and the prime indicates that the terms $k = 0$ and N are multiplied by $1/2$. For a nonperiodic integrand, the accuracy of the formula will normally be $I_N - I = O(h^2)$ as $N \to \infty$, i.e., as $h \to 0$. The observation underlying the Euler–Maclaurin formula is that if v is sufficiently smooth, this figure can be improved to $O(h^4)$, $O(h^6)$, and so on by subtracting appropriate multiples of $v'(2\pi) - v'(0)$, $v'''(2\pi) - v'''(0)$, and so on. Specifically, one statement of the formula in terms of an asymptotic series is as follows:

THEOREM 11.1. *For the trapezoidal rule applied to a function $v \in C^\infty[0, 2\pi]$, $I_N - I$ has the asymptotic series*

$$I_N - I \;\sim\; h^2\frac{B_2}{2!}[v'(2\pi) - v'(0)] + h^4\frac{B_4}{4!}[v'''(2\pi) - v'''(0)] + \dots \tag{11.1}$$

as $N \to \infty$, where $\{B_k\}$ are the Bernoulli numbers ($B_2 = 1/6$, $B_4 = -1/30$, $B_6 = 1/42, \dots$).

Alternative formulations give an explicit expression for the error if the series is truncated after a certain number of steps, in which case v need only have the corresponding number of derivatives.

The case of interest to us is what occurs if v and its derivatives are 2π-periodic. Then all the coefficients in the series are zero, and we have

$$I_N - I \;\sim\; 0h^2 + 0h^4 + \dots. \tag{11.2}$$

This is an interesting asymptotic expansion indeed! One might make the mistake of thinking that it implies that for 2π-periodic $v \in C^\infty(\mathbb{R})$, I_N must be exactly equal to I. But this is not true. The correct implication is that in this case $I_N - I$ decreases faster than any finite power of h as $N \to \infty$. This conclusion holds if v is infinitely differentiable, and can be modified appropriately if it has just a finite number of derivatives.

Poisson understood the limitations of asymptotic series, and wrote [136]

> Euler's series, after having been convergent in the initial terms, sometimes ends by becoming divergent and consequently inexact.[9]

As mentioned above, the Euler–Maclaurin formula provides a popular explanation of the special accuracy of the trapezoidal rule for periodic functions. It is a rather odd explanation, however, in that it imagines that v has a discontinuity at the endpoints and then takes the discontinuity away again, Cheshire cat-style. Perhaps such an argument is most natural in a context in which v is less smooth at the endpoints than in the interior, as occurs, for example, in the analysis of the IMT and tan rules for numerical integration (§15).

[9] "...la série d'Euler, après avoir été convergente dans les premiers termes, finit quelquefois par devenir divergente, et par conséquent inexacte."

Closer to the spirit of this article, with its emphasis on analytic functions, would be a formulation of the Euler–Maclaurin idea that made use of analyticity. Specifically, in §4 we gave a proof of Theorem 4.1 based on a contour integral over the contour depicted in Figure 4.2. Similarly, Theorem 4.2 can be proved with the use of a rectangular contour symmetric with respect to the real axis. Now suppose v is analytic as in Theorem 4.2 but not necessarily periodic. Could the same contour integral be used to derive a more general bound that incorporated both the exponential decay of Theorem 4.2 and also the finite order error terms of (11.1)? Could such a bound yield Theorem 11.1 in one limit and also Theorem 4.2 if v happens to be periodic?

This project has been carried out in forthcoming work by the first author and Javed [86]. For any $m \ge 0$ and sufficiently smooth v, define

$$Q_m = \sum_{\substack{k=1 \\ k \text{ odd}}}^{m} h^{k+1} \frac{v^{(k)}(2\pi) - v^{(k)}(0)}{(k+1)!} B_{k+1}. \tag{11.3}$$

The theorem from [86] is as follows.

THEOREM 11.2. *Given $a > 0$ and $m \ge 0$, let v be bounded by $|v(z)| \le M$ and have a bounded $(m+1)$st derivative in the region defined by $0 \le \mathrm{Re}\, z \le 2\pi$, $-a < \mathrm{Im}\, z < a$, and be analytic in the interior of this region. Then*

$$I_n - I - Q_m = O(h^{m+2}\Delta^{(m+1)}) + O(e^{-2\pi a/h}\Delta) + O(e^{-2\pi a/h}M) \tag{11.4}$$

as $n \to \infty$, where

$$\Delta = \max_{\substack{1 \le k \le m \\ k \text{ odd}}} \left| v^{(k)}(2\pi) - v^{(k)}(0) \right| \tag{11.5}$$

and

$$\Delta^{(m+1)} = \sup_{-a<y<a} \left| v^{(m+1)}(2\pi + iy) - v^{(m+1)}(iy) \right|. \tag{11.6}$$

The constants implicit in the big O symbols are absolute, i.e., independent of m, a, and v.

If v is 2π-periodic, the first two terms on the right in (11.4) are zero, and we recover the $O(e^{-2\pi a/h}M)$ exponential dependence of Theorem 4.2. For more discussion of the implications of Theorem 11.2, see [86].

12. History. At this stage in the paper, we would like to be able to summarize who were the earliest authors to investigate the fast convergence of the trapezoidal rule in the three special cases of the periodic interval, the unit circle, and the real line. We find this surprisingly difficult. The roots of the fast trapezoidal rule go back to Euler and Maclaurin, if not Archimedes, but it is hard to find explicit discussions. One source of information on early works on quadrature in general is the 1961 bibliography by Stroud [158].

For the periodic interval, as told in the introduction, the first such discussion may be that of Poisson in the 1820s [136, 137]. Many mathematicians of the 19th century must have been aware of Poisson's work, but we do not know what use was made of it, and there does not seem to be any author who explicitly stated the geometric convergence for analytic functions until Davis in 1959 [35]. Early papers in the computer era include those of Fettis in 1955 [44] and Moran in 1958 [125], but their discussions, based on the Poisson summation formula, do not include general

convergence statements. This 150-year gap seems odd, and it is hardly surprising that Davis (who does not cite Poisson) calls the result "folklore." Davis's paper considers both intervals and circles, and as in §2, he makes use of convergent series. Our Theorem 4.2 is essentially his Theorem 1, which is restated in §4.6 of the book by Davis and Rabinowitz [36, p. 243]. One can also find essentially the same result as Theorem 9.28 of [98] or Exercise 12.6 of [167].

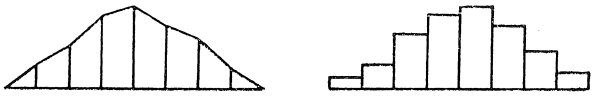
For the unit circle, after Davis, a key early figure was James Lyness in the 1960s [37, 111, 112, 114, 115]. In a succession of papers, some of them written jointly with Delves, Moler, and Sande, Lyness advocated the use of the trapezoidal rule on the unit circle to evaluate contour integrals numerically for applications including evaluation of functions, calculation of derivatives, and zerofinding. Such techniques are the subject of the next section. These methods are closely linked to the Fast Fourier Transform, which was an established tool by the 1970s. The field of digital signal processing developed rapidly in that era, and many of its algorithms are based on computation with data at roots of unity [11].

In the case of the real line, interest in the equispaced sampling of functions originated about a century ago. One line of research arose from the introduction of cardinal interpolants of equispaced data by E. T. Whittaker in 1915 [184]. Whittaker's starting point was the construction of functions by applying finite difference formulas on wider and wider equispaced grids, and his work was closely related to operational calculus and divergent series, subjects of interest to British mathematicians including Heaviside, Hardy, Littlewood, Milne-Thompson, and Titchmarsh. We wish Hardy were here to ask about the trapezoidal rule! Whittaker's son J. M. Whittaker wrote a tract some years later on what he called *interpolatory function theory* [185], with a bibliography listing work by many authors including Copson, Ferrar, Hurwitz, and Takenaka. Meanwhile, across the Atlantic, the subject of *sampling theory* was being invented beginning in the 1920s by Hartley, Nyquist, and Shannon (not to mention less-often cited authors Kotelnikov, Raabe, Gabor, and Someya [173]), wherein Whittaker's cardinal function is interpreted as a band-limited interpolant, and Paley and Wiener's *Fourier Transforms in the Complex Domain* appeared in 1934 [134].

The people who first used the trapezoidal rule on the real line for numerical calculation of integrals may have been the early British statisticians, who were concerned with the estimation of moments of a distribution such as mean, variance, skewness, and kurtosis. The first publication we have found that discusses the matter is Aitken's 1939 book *Statistical Mathematics* [3]. Though Aitken does not speak of geometric convergence or state general theorems, he touches upon the same issues we have discussed in §§6 and 7, and we hope the reader looking at Figure 12.1 will share some of the pleasure we felt in encountering this material in Aitken's book.

A more focussed analysis of the fast convergence of the trapezoidal rule on the real line appears in a paper by Turing written in 1939 and published in 1943 [172]. Turing's aim was a numerical exploration of the Riemann Hypothesis, and he introduced a method for evaluating the zeta function based on the trapezoidal rule applied to a smooth function on the real line. He writes, "We approximate to the integral by the obvious sum $\sum_{k=-K}^{K} \frac{1}{K} h\left(\frac{k}{K}\right)$ and we find that... this gives a remarkably accurate result; when the number of terms taken is $T = 2K + 1$ the error is of the order of magnitude of $e^{-\frac{1}{2}\pi T}$." He derives a bound with the flavor of Theorem 6.1, with a proof using contour integrals, and alludes to the use of the trapezoidal rule on more general contours (§16).

Several times in the course of writing this article we have encountered leads to



Frequency Polygon. Histogram.

Remark. The relation between an *integral* and a *sum* of equidistant ordinates of the kind here considered enters into pure mathematics in the *Euler-Maclaurin summation formula*, by which a sum of ordinates is expressed as an integral over the range plus correction terms involving the derivatives of odd order taken at the boundaries. In many cases, where the derivatives $\phi'(x)$, $\phi'''(x)$, ... are not absolutely zero but converge to zero as a limit, the representation of the integral on the left of (1) by the sum on the right needs very careful investigation. It is found, however, that for the statistical functions to which Sheppard's corrections are usually applied the difference between the integral and the sum can be made negligibly small by taking values of the class-interval w of a size quite customary in practice. Usually it is enough for w to be less than the standard deviation. The following derivation of the formulæ must be regarded as approximate only.

Ex. 1. The following two comparisons of integral with sum over an infinite range are interesting in this respect:

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^2} = \pi = 3{\cdot}14159 \text{ nearly,}$$

whereas

$$\sum_{-\infty}^{\infty} \frac{1}{1+x^2} = 3{\cdot}15336 \text{ nearly,}$$

x taking the values $0, \pm 1, \pm 2, \ldots$. Again,

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx = \sqrt{2\pi} = 2{\cdot}506628275 \text{ nearly,}$$

whereas

$$\sum_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} = 2{\cdot}506628288 \text{ nearly,}$$

x taking the same values as before. The first sum in these examples is only moderately close to the corresponding integral; the second is very close, and still closer results are obtained if a summation with a finer subdivision of x is used.

FIG. 12.1. *Above, a sketch from p.* 28 *of Aitken's* 1939 *book* [3]. *Below, an extract from pp.* 44–45. *Note his comment that in practice, it is good enough to take the step size to be smaller than the standard deviation.*

the mathematical literature of the Soviet Union, including the names of Bakhvalov, Fok, Korobov, and Malozemov. It is quite possible that Fok, for example (known in the West through Hartree–Fock methods), worked with the geometrically convergent trapezoidal rule before Turing. However, our links to the Soviet literature are not as strong as we would like, and we have failed to turn up a definitive publication.

Our overall assessment of the situation in the first half of the 20th century is that the fast convergence of the trapezoidal rule would have been readily understood by many mathematicians, but that, before the introduction of computers, it did not stand out as a topic worthy of much special attention. This kind of mathematics was used more often to approximate sums by integrals than integrals by sums, and convergence as $h \to 0$ was not a prominent issue. Indeed, the use of the actual term "trapezoidal

TABLE 12.1

Twelve approaches to theorems about the fast convergence of the trapezoidal rule, with pointers to publications we know of in each category before 1970*. Page numbers indicate locations in this article where proofs of each kind are spelled out. As the blank third row suggests, we do not know of published works before* 1970 *that derive the fast convergence of the trapezoidal rule from analysis of interpolants, or after* 1970 *either, even though, as discussed in* §10*, this is the kind of argument that is most robust for generalizations such as perturbed quadrature nodes. We also do not know of literature on the Euler–Maclaurin formula on the unit circle before or after* 1970*, apart from a* 2006 *paper of Berrut* [18]*. The reason two page numbers are cited in some cases is that in fact, each box of the table could be doubled along a third dimension corresponding to the alternatives of nonsymmetric domains of analyticity, such as a disk or a half-plane, or symmetric ones, such as an annulus or a strip.*

	circle	interval	real line
series or transform	Lyness 1967 Davis 1959 (pp. 6,8)	Fettis 1955 Moran 1958 Davis 1959 (pp. 12,15)	Fettis 1955 Moran 1958 (p. 19)
contour integral	 (p. 7)	 (pp. 14,15)	Turing 1943 Goodwin 1949 Luke 1956 Hunter 1964 McNamee 1964 Martensen 1968 Schwartz 1969 (p. 20)
interpolation	(pp. 25,26)	(pp. 27,27)	
Euler–Maclaurin		Poisson 1826 Milne-Thompson 1933 Sag & Szekeres 1964 Isaacson & Keller 1966 (p. 34)	Aitken 1939 Faddeeva 1954

rule" in this context was not so common until later, perhaps in the 1950s. Davis uses this term, as do Isaacson and Keller in their 1966 textbook [84], which relates the periodic trapezoidal rule to the Euler–Maclaurin formula.

After Turing, a paper often cited as an early contribution to the trapezoidal rule on the real line is that of Goodwin in 1949 [65], who begins, "It is well known to computers that the approximate formula [the trapezoidal rule] yields a surprising degree of accuracy. . . ." This use of the word "computers" is in the charming older sense of the term. Goodwin and Turing were colleagues at the National Physical Laboratory during 1945–48, working on the ACE computer project, and Goodwin states that his purpose is to give the details of a generalization of Turing's method. In his formulation of the matter, it is assumed that the integrand decays at the rate $\exp(-x^2)$, a condition that will usually hold in statistical applications. The first appearance of a more general theorem along the lines of Theorem 6.1, again based on a contour integral, seems to be in the 1968 paper of Martensen [118], who cites both Turing and Davis. Schwartz in 1969 [146] also shows exponential convergence using a contour integral, without citing any previous publications.

By 1970, Mori and Takahasi were developing transformed methods of integration

based on the trapezoidal rule (§15), spectral methods were being invented for ODE and PDE [95, 133, 167], Stenger was developing his theory of numerical methods based on sinc functions [155], FFT-fueled digital signal processing was taking off with the discrete z-transform as its central tool [173], and Martensen's student Kress was beginning a career that would apply related ideas to many problems of quadrature and integral equations [96, 97]. At this point the literature related to the fast convergence of the trapezoidal rule begins to get voluminous, and with this in mind, we summarize in Table 12.1 the contributions we are aware of before 1970.

Part II. Applications

13. Contour integrals. We now begin the second half of this article, devoted to applications.

In the preceding pages, we have applied contour integrals to the analysis of the trapezoidal rule. In this section, we turn around to consider the trapezoidal rule as a method for evaluating contour integrals. It seems ideally suited to this task, since contour integrals are often posed over circles in the complex plane, and by the nature of the problem, the integrand is analytic. The ideas for this go back to Poisson and Cauchy in Paris in the 1820s, though we doubt these two men made the connection. Things changed after the introduction of the Fast Fourier Transform in 1965, providing a powerful tool for all kinds of computations with periodic discrete data.

We begin with equations (2.4) and (2.5). If u is analytic and bounded by M in the disk $|z| < r$ for some $r > 1$, then it has a Taylor series

$$u(z) = \sum_{j=0}^{\infty} c_j z^j \tag{13.1}$$

with coefficients

$$c_j = \frac{1}{2\pi i} \int_{|z|=1} z^{-j-1} u(z)\, dz. \tag{13.2}$$

Following (2.2), we define the N-point trapezoidal rule approximation to c_j by

$$c_j^{[N]} = \frac{1}{N} \sum_{k=1}^{N} z_k^{-j} u(z_k), \tag{13.3}$$

where $\{z_k\}$ are the Nth roots of unity, $z_k = e^{2\pi i k/N}$. For $j = 0$, this is the same setting as in Theorem 2.1, apart from a factor of 2π, and that theorem implies $|c_0 - c_0^{[N]}| \le M/(r^N - 1)$. For $j > 0$, however, we find ourselves in the situation corresponding to an integrand u in (2.1) with a pole at $z = 0$, analytic in an annulus but not a disk. Thus we must turn to arguments related to Theorem 2.2 to get a statement of geometric convergence, and in so doing, we might as well relax the assumption on u to analyticity in an annulus rather than a disk. The definitions (13.2) and (13.3) still apply, now for all integers j, and the Laurent series representation of u is

$$u(z) = \sum_{j=-\infty}^{\infty} c_j z^j. \tag{13.4}$$

Here is the estimate that results, which can be found as equation (3.5) of [73].

THEOREM 13.1. *Suppose u is analytic and satisfies $|u(z)| \le M$ in the annulus $r^{-1} < |z| < r$ for some $r > 1$. Then for any $N \ge 1$ and any j,*

$$|c_j^{[N]} - c_j| \le \frac{M(r^j + r^{-j})}{r^N - 1}. \tag{13.5}$$

Proof. In analogy to (2.17), the error is given by the aliasing formula,

$$c_j^{[N]} - c_j = \sum_{k=1}^{\infty} (c_{j+kN} + c_{j-kN}). \tag{13.6}$$

By (2.18), the Laurent coefficients satisfy $|c_k| \le Mr^{-|k|}$. Combining these observations gives (13.5). □

Theorem 13.1 implies that Taylor and Laurent coefficients can be computed with geometric accuracy by applying the trapezoidal rule on the unit circle. In the early FFT era, this idea was analyzed and applied to a variety of problems of complex analysis by Henrici [73, 74]. An algorithm with Fortran software was proposed by Lyness and Sande in 1971 [115]. For much more discussion and context, see [11].

Computing the jth Taylor series coefficient c_j of a function u is equivalent to evaluating its jth derivative at $z = 0$, $u^{(j)}(0) = j!c_j$, and thus Theorem 13.1 also provides the error bound $Mj!(r^j + r^{-j})/(r^N - 1)$ for the evaluation of the jth derivative by the trapezoidal rule. A related method of evaluating derivatives, without mention of the FFT, was proposed by Lyness and Moler [114], who pointed out the great contrast with the numerical instability suffered by finite difference methods based on just real arguments:

> Once complex arguments are allowed, the principal difficulties encountered in numerical differentiation simply disappear.

Lyness and Moler gave the example of the evaluation of the fifth derivative at $z = 0$ of the function

$$u(z) = \frac{e^z}{\sin^3(z) + \cos^3(z)},$$

with exact answer $u^{(5)}(0) = -164$. Using a seven-point centered finite difference with step size h, the best result one can get is about -164.02 with $h \approx 0.0024$. Taking the fifth derivative of a polynomial interpolant to u in 37 Chebyshev points in $[-0.5, 0.5]$ (to avoid the pole at $z \approx -0.785$) improves this result to about -1.63999999985, i.e., about 10 digits of accuracy. Convergence at the rate $\approx (0.5/0.785)^n$ to 15 digits of accuracy is achieved, however, by the trapezoidal rule over a circle of radius 0.5:

N	Estimate of $u^{(5)}(0)$
20	$-\mathbf{164.0}13$
40	$-\mathbf{164.00000}16$
60	$-\mathbf{164.00000000}19$
80	$-\mathbf{164.000000000000}22$

In this example, the trapezoidal rule was applied over a circle of radius $\rho = 0.5$ rather than 1 to avoid a pole. Even when there are no singularities to avoid, the choice

of radius can make a big difference to both the accuracy and the numerical stability of these methods, as has been recognized from the start. An adaptive algorithm for choosing a good radius ρ was published by Fornberg in 1981 [49, 50]. Recently this matter has been investigated comprehensively by Bornemann [22], who has shown that for a wide class of functions u, an optimal choice of ρ enables computation of derivatives of all orders with condition number essentially $O(1)$.[10] As the order of the derivative increases, so does ρ.

In fact, as mentioned in §2, the use of the trapezoidal rule for a Cauchy integral may be numerically advantageous even for the evaluation of the zeroth derivative, i.e., the function itself. Suppose, for example, we wish to evaluate the "phi function"

$$u(z) = \frac{e^z - 1 - z}{z^2} \tag{13.7}$$

for $z = 10^{-8}$ [77, 88, 144] . There is nothing wrong with this function mathematically: it has a removable singularity at $z = 0$. On a computer in standard 16-digit arithmetic, however, evaluation of $u(z)$ directly from the formula (13.7) fails completely due to cancellation error, giving $u(z) = -0.6077\ldots$. The 16-point trapezoidal rule on the unit circle, by contrast, gives the correct answer $u(z) \approx 0.500000001666667$ to full machine precision [144]. This method of stabilized evaluation of difficult functions is applied to matrices and operators in [88, 144], and generalizations to rational functions are exploited in [51] and (implicitly) in [64].

Related to Taylor coefficients are *generating functions.* For example, the jth Taylor coefficient c_j of the function $z/(e^z - 1)$, mentioned in (2.14), is $1/j!$ times the jth Bernoulli number B_j. Thus the trapezoidal rule can be used to compute Bernoulli numbers. With $N = 128$ points on the circle of radius 4, $B_0, \ldots, B_{15}$ are computed to almost full precision in 16-digit arithmetic.

Computing a Taylor coefficient from (13.2) amounts to the evaluation of a contour integral over a circle for which the integrand has a pole right at the center of the circle. However, the more general notion of a Cauchy integral involves a pole at an arbitrary location inside the integration contour. If u is an analytic function in the closed unit disk, then its value for any a with $|a| < 1$ is given by the Cauchy integral formula,

$$u(a) = \frac{1}{2\pi i} \int_{|z|=1} \frac{u(z)}{z - a} dz. \tag{13.8}$$

(The generalization of this formula to the case where a becomes a matrix A will be a central theme of §19.) Regardless of the choice of a, this integrand is analytic in an annulus around the unit circle, and since 1965, mathematicians have been attracted to the idea of exploiting the trapezoidal rule for evaluating the integral. Geometric convergence is guaranteed, with a rate determined by whichever is closer to the unit circle: the circle of analyticity of u, or the point a.

[10] An extreme choice of radius is utilized in the method of *complex step differentiation,* which goes back at least to Squire and Trapp in 1998 [153]. Here, the first derivative of a real analytic function f at a real argument x is computed to machine accuracy by the formula $(f(x + i\varepsilon) - f(x - i\varepsilon))/(2i\varepsilon)$, where ε is a number smaller than machine epsilon, such as 10^{-100} in standard 16-digit IEEE arithmetic. From the point of view of machine arithmetic, the reason the method works is that the real and imaginary parts of a function evaluation are decoupled, so the subtraction is performed with negligible cancellation error. From the point of view of the mathematics of the trapezoidal rule, one may interpret this as the 2-point trapezoidal rule (more precisely the 2-point midpoint rule) evaluated over a circle of radius $\rho = 10^{-100}$, with high accuracy guaranteed by (2.15).

THEOREM 13.2. *Suppose u is analytic and bounded in the disk $|z| < r$ for some $r > 1$, and $u^{[N]}(a)$ is the approximation to $u(a)$ for some a with $|a| < 1$ obtained by applying the N-point trapezoidal rule on the unit circle to* (13.8). *Then as $N \to \infty$,*

$$u(a) - u^{[N]}(a) = O(\max\{r^{-N}, |a|^N\}). \tag{13.9}$$

Proof. This can be derived from the arguments used to prove Theorem 2.2. □

This method of combining the trapezoidal rule with Cauchy integrals is powerful and flexible. Nevertheless, when a is close to the circle of integration, it may be far from optimal, when the second term of the maximum in (13.9) is bigger than the first. Suppose, for example, we wish to evaluate $u(z) = \exp(z - 0.9)$ at $a = 0.9$ from values at $N = 32$ roots of unity; the exact result is 1. The Cauchy integral method gets just 1.5 digits of accuracy:

Trapezoidal rule for Cauchy integral: $u(0.9) \approx 1.035557779939565$.

By contrast, suppose we use exactly the same 32 data values to define a degree 31 polynomial interpolant, and then evaluate the interpolant at a. We now get full machine precision:

Polynomial interpolation: $u(0.9) \approx 1.000000000000000$.

To understand the striking difference between these two quantities, it is interesting to compare two lines of Matlab that may be used to calculate them. If `z` is a Matlab vector of roots of unity, we may write

```
cauchy = mean(z.*u(z)./(z-z0))
interp = mean(z.*u(z)./(z-z0))./mean(z./(z-z0))
```

The first line implements the trapezoidal rule for the Cauchy integral based on data at the roots of unity, whereas the second computes the polynomial interpolant from the same data by the barycentric interpolation formula [11, 19]. When a is equal to 0, the two formulas are the same, and this observation was the basis of §8. For general a, note that the second formula can be interpreted as a Cauchy integral evaluation of u at a divided by a Cauchy integral evaluation of the function $g(z) = 1$ at a. The same two formulas are compared from other points of view in [11, 71, 82]. The main point is that one must be cautious in evaluating Cauchy integrals near the contour of integration. This fact is well known to practitioners of the numerical solution of integral equations (§20), where various methods have been proposed to cope with it [71, 91].

The idea of applying the trapezoidal rule to Cauchy integrals can be applied to many other problems besides evaluation of functions and derivatives. For example, if u is analytic in the closed unit disk, what is the number ν of zeros of u in the disk? Let us assume that u is nonzero on the unit circle. Then by the principle of the argument, ν is equal to the winding number of $u(z)$ about the origin as z traverses the unit circle once in the counterclockwise direction. This is in turn equal to $(2\pi i)^{-1}$ times the change in $\log(u(z))$ as z traverses the circle. Since the derivative of $\log(u(z))$ is $u'(z)/u(z)$, we accordingly have

$$\nu = \frac{1}{2\pi i} \int_{|z|=1} \frac{u'(z)}{u(z)}\, dz. \tag{13.10}$$

Another way to interpret this formula is to note that $u'(z)/u(z)$ is a function with simple poles at the zeros of $u(z)$, each with residue equal to 1, and the contour integral will give the sum of the residues, namely ν. From (2.1) and (2.2), we conclude that ν can be calculated from the trapezoidal rule as

$$\nu \approx \frac{1}{N} \sum_{j=1}^{N} \frac{z_j u'(z_j)}{u(z_j)}. \tag{13.11}$$

The convergence as $N \to \infty$ will be geometric, and since the limit is an integer, it will usually not be hard to spot its correct value. For example, the 40-point trapezoidal rule estimates that $u(z) = \sin^3(2z) + \cos^3(2z)$ has $2.99863\ldots$ zeros in the unit disk, and increasing N to 100 gives $2.9999999256\ldots$.

Equation (13.10) can be generalized to enable one not just to count zeros but to find them. If u is an analytic function with a single simple zero ζ in the unit disk, then ζ is given by

$$\zeta = \frac{1}{2\pi i} \int_{|z|=1} \frac{z u'(z)}{u(z)}\, dz. \tag{13.12}$$

A numerical utilization of this formula was perhaps first recommended by McCune [120]. If u has N zeros $\{\zeta_j\}$ in the disk, the formula generalizes to

$$\sum_{j=1}^{N} \zeta_j^k = \frac{1}{2\pi i} \int_{|z|=1} \frac{z^k u'(z)}{u(z)}\, dz \tag{13.13}$$

for the sum of the kth powers of the zeros [85]. (Equation (13.10) corresponds to $k = 0$: we can count the zeros by adding their zeroth powers.) In other words, the sums on the left in (13.13) are the Laurent coefficients of u'/u. From here it is not far to the algorithm for finding zeros of analytic functions proposed by Lyness and Delves in 1967 [37, 112] and improved by Kravanja and Van Barel [94, §1.2]. The algorithm proposed later by Luck and Stevens [107] is a variant of this based on computing $\int (z/f(z))dz \,/ \int (1/f(z))dz$ rather than $\int z(f'(z)/f(z))dz$, and this too can be generalized to multiple roots [94, §1.6]. Changing $1/f(z)$ to $g(z)$ gives an algorithm for finding a pole of a meromorphic function, related to the FEAST algorithm for computing matrix eigenvalues to be mentioned at the end of §19. Besides works by Lyness and Kravanja and their coauthors, there are also a number of other contributions in this area, such as a method proposed by Henrici [73, §3.2]. Connections to the so-called Burniston–Siewert method for zerofinding are analyzed in [9].

The issues summarized in this section are investigated more fully in [11], where among other things, it is shown that Kravanja's generalization of the Luck–Stevens method to multiple zeros is mathematically equivalent to the determination of the poles of a linearized rational interpolant to $1/f(z)$.

14. Rational approximation. Every quadrature formula can be related to rational or meromorphic functions. In this section we spell out some of these connections for the trapezoidal rule and show how it has been used to derive approximate solutions to two famous problems of rational approximation theory.

To begin the discussion, recall one of our proofs involving contour integrals, say, the proof of Theorem 2.1 for the integral over the unit circle of a function $u(z)$ analytic in $|z| < r$ for some $r > 1$ (p. 7). The key idea was to multiply $u(z)$ by the

rational function $m(z)$ of (2.11), which is listed in the upper-left position of Table 14.1. This function has simple poles at the quadrature points (the Nth roots of unity) with residues equal to $(2\pi i)^{-1}$ times the corresponding quadrature weights $2\pi/N$. It follows that the contour integral of $m(z)u(z)$ around the circle $|z| = r$ is equal to I_N, the result of the trapezoidal rule. The other entries of Table 14.1 list analogous functions for the other five fundamental cases of this paper.

TABLE 14.1

Characteristic functions m for trapezoidal quadrature in the six fundamental cases reviewed in this paper. In each case, the result of the trapezoidal rule applied to an integrand f is equal to the contour integral of f times m over an appropriate contour Γ. In the top row, the contours Γ are the circle $|z| = r$, the line segment $[0, 2\pi] - ia$, and the line $\mathbb{R} - ia$. In the bottom row they are the boundary of the annulus $r^{-1} < |z| < r$, the periodic rectangle $[0, 2\pi] \pm ia$, and the strip boundary $\mathbb{R} \pm ia$, all traversed in the usual positive sense.

	circle	interval	real line
nonsymmetric	$\dfrac{-iz^{-1}}{1-z^{-N}}$	$\dfrac{1}{1-e^{-iN\theta}}$	$\dfrac{1}{1-e^{-2\pi ix/h}}$
symmetric	$-\frac{i}{2}z^{-1}\,\dfrac{1+z^N}{1-z^N}$	$-\frac{i}{2}\cot(N\theta/2)$	$-\frac{i}{2}\cot(\pi x/h)$

Now the true integral I can also be represented by a contour integral of u times a characteristic function, which we shall call μ. For the case of the unit circle just discussed, by (2.6), we have $\mu(z) = -iz^{-1}$. This function is listed in the upper-left position of Table 14.2, and the remaining entries show the other five cases.

TABLE 14.2

Characteristic functions μ for exact integration in the six fundamental cases. In each case, the true integral I is equal to the contour integral of f times μ over the same contours Γ as in Table 14.1*. Each function in Table* 14.1 *is an approximation to the corresponding function in Table* 14.2*, whose accuracy improves with distance to Γ.*

	circle	interval	real line
nonsymmetric	$-iz^{-1}$	1	1
symmetric	$-\frac{i}{2}z^{-1},\ \|z\|>1$ $\frac{i}{2}z^{-1},\ \|z\|<1$	$-\frac{1}{2},\ \mathrm{Im}(\theta)>0$ $\frac{1}{2},\ \mathrm{Im}(\theta)<0$	$-\frac{1}{2},\ \mathrm{Im}(x)>0$ $\frac{1}{2},\ \mathrm{Im}(x)<0$

Contour integral estimates for the accuracy of the trapezoidal rule for analytic integrands are based on comparing a function in Table 14.1 with the corresponding function in Table 14.2. For example, consider the middle of the bottom row in the two tables, corresponding to integration of a 2π-periodic function $v(\theta)$ analytic in a strip of half-width a around the real axis in the complex θ-plane (see [86]). By the definitions of $m(\theta)$ and $\mu(\theta)$, the error of the trapezoidal rule can be expressed as the

integral

$$I_N - I = \int_\Gamma v(\theta)[m(\theta) - \mu(\theta)]d\theta, \tag{14.1}$$

where Γ is the rectangle $[0, 2\pi] \pm ia$ traversed in the positve direction. (We can ignore the ends of the rectangle, as their contributions cancel by periodicity.) As $\mathrm{Im}\,(\theta) \to \infty$, $m(\theta)$ approaches $-1/2$ exponentially, and as $\mathrm{Im}\,(\theta) \to -\infty$, it approaches $1/2$. Thus $m(\theta) - \mu(\theta)$ is exponentially small for large $|\mathrm{Im}\,(\theta)|$. It follows that if v is analytic in a wide region around $[0, 2\pi]$ and does not itself grow too fast with $|\mathrm{Im}\,(\theta)|$, (14.1) delivers a powerful accuracy bound.

In other words, contour integrals connect quadrature to the subject of rational or more generally meromorphic approximation. The function to be approximated is μ, with a jump discontinuity along the integration interval. (The discontinuity is explicit in the second row of Table 14.2 and implicit in the first row.) The function m approximates that jump discontinuity by a sequence of poles.[11] In the upper-left position of the tables, for example, after dividing by $-iz^{-1}$, we note that $(1-z^{-N})^{-1}$ is a rational function that is close to 1 outside the unit disk and close to 0 inside.[12]

Connections between quadrature formulas and rational functions have been a part of quadrature theory for 200 years. Most classically, for integration of a nonperiodic function on the interval $[-1, 1]$, the function $\mu(\xi)$ that arises is $\log((\xi - 1)/(\xi + 1))$, with a branch cut along $[-1, 1]$ and analytic elsewhere in the extended complex plane. When Gauss invented Gauss quadrature in 1814, he did it by approximating μ by what we would now call its type $(N-1, N)$ Padé approximant at $\xi = \infty$ (he used the language of continued fractions). Comparisons of $m(\xi)$ and $\mu(\xi)$ for other quadrature formulas on $[-1, 1]$ were investigated by Takahasi and Mori [161]. For the special case of Gauss vs. Clenshaw–Curtis quadrature, see Figure 5 of [168].

We now turn to two particular problems in the subject of rational approximation where the trapezoidal rule can be used to powerful effect, following the discussion in the chapter on "Two famous problems" in [169].

The first problem is approximation of $y(\xi) = |\xi|$ on $[-1, 1]$ by type (n, n) rational functions, a prototypical rational approximation problem for non-smooth functions. Let $E_{nn}(|\xi|)$ denote the minimal value of $\| |\xi| - r(\xi)\|$ over all rational functions r of type (n, n), where $\|\cdot\|$ is the ∞-norm on $[-1, 1]$. It was a striking surprise when Donald Newman proved in 1964 that the accuracy of these approximations can improve root-exponentially with n [130]. Since polynomials can never do better than $O(n^{-1})$, this established that rational functions can be far more powerful than polynomials in certain applications, a theme implicit in a number of parts of this article. Later Stahl found the precise asymptotic behavior [154]:

$$E_{nn}(|\xi|) \sim 8e^{-\pi\sqrt{n}}. \tag{14.2}$$

[11] These formulations can be regarded as uses of the theory of *hyperfunctions*, which are generalized functions defined by analytic functions in the lower and upper half-plane that may have singularities or discontinuities along the real axis [66]. Table 14.1 is thus a list of hyperfunction approximations to the hyperfunctions of Table 14.2.

[12] Why, one might wonder, would one wish to approximate the simple rational function $-iz^{-1}$ by the more complicated rational function $-iz^{-1}/(1-z^{-N})$? This question is not idle, but reflects the fact that by the mean value theorem, if $u(z)$ is analytic in $|z| \le 1$, its integral over $|z| = 1$ can be evaluated by a single point evaluation at $z = 0$. So indeed, sometimes one would *not* want to use the more complicated function.

Following [169] and arguments by Stenger [156], it is now apparent that a simpler argument than Newman's could have sufficed to estabslish the root-exponential behavior, as follows. We start from the identity

$$|\xi| = \frac{2\xi^2}{\pi} \int_{-\infty}^{\infty} \frac{e^x dx}{e^{2x} + \xi^2},$$

which can be established by an elementary substitution that converts the integral on the right to an arctangent. This is an attractive integral to work with because the integrand decays exponentially as $|x| \to \infty$. We now get a rational approximation of $|\xi|$ by approximating this integral by the trapezoidal rule with node spacing $h > 0$:

$$r(\xi) = \frac{2h\xi^2}{\pi} \sum_{k=-(n-2)/4}^{(n-2)/4} \frac{e^{kh}}{e^{2kh} + \xi^2}. \tag{14.3}$$

Here n is a positive even number, and there are $n/2$ terms in the sum, so $r(\xi)$ is a rational function of ξ of type (n, n). There are two sources of error that make $r(\xi)$ differ from $|\xi|$. The fact that the sum has been terminated at a limit $n < \infty$ introduces an error on the order of $e^{-nh/4}$, and the finite step size $h > 0$ introduces an error on the order of $e^{-\pi^2/h}$. (The integrand is analytic in the strip around the real x-axis of half-width $a = \pi/2$, corresponding to a convergence rate $e^{-2\pi a/h}$.) Balancing these sources of error as in §7 suggests the condition $e^{-nh/4} \sim e^{-\pi^2/h}$, that is,

$$h \sim \frac{2\pi}{\sqrt{n}}, \quad n \to \infty,$$

with error of order

$$\|f - r\| \approx e^{-(\pi/2)\sqrt{n}}.$$

This achieves the root-exponential behavior of (14.2), falling short of the actual optimal convergence rate by a factor of 4 in the sense that it requires n to be 4 times larger than in (14.2) to achieve the same error.

The second famous problem is approximation of e^t on $(-\infty, 0]$. This problem was introduced in a 1969 paper of Cody, Meinardus and Varga [29], which drew attention to the connection of such approximations with the numerical solution of partial differential equations and showed that the convergence was at least geometric. What is the optimal rate? The idea that it might be $O(9^{-n})$ became known in the 1970s as the $1/9$ *conjecture*. In fact, the optimal convergence rate turned out to be $O(H^n)$ with $H \approx 1/9.28903$, a number now known as *Halphen's constant*, as was conjectured numerically by Trefethen and Gutknecht [170] and proved by Gonchar and Rakhmanov [63].

Again we can derive the exponential behavior from the trapezoidal rule, employing methods to be detailed in §16. We begin with the Laplace transform identity

$$e^t = \frac{1}{2\pi i} \int_\Gamma \frac{e^s ds}{s - t}, \quad t < 0,$$

where the integral is over a Hankel contour Γ of the kind to be described in §16. Choosing the contour to be a parabola, we convert this to an integral over the real x-axis by the change of variables

$$s = (ix + a)^2, \quad ds = 2i(ix + a)\, dx$$

for a constant $a > 0$, which gives

$$e^t = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{e^{(ix+a)^2}(ix+a)}{(ix+a)^2 - t}\, dx.$$

We now approximate this integral by the trapezoidal rule with node spacing $h > 0$:

$$r(t) = \frac{h}{\pi} \sum_{k=-(n-1)/2}^{(n-1)/2} \frac{e^{(ikh+a)^2}(ikh+a)}{(ikh+a)^2 - t}. \tag{14.4}$$

Here n is a positive odd number, and since t rather than t^2 appears in each term, we now take n terms in the sum rather than $n/2$ as in (14.3) to make $r(t)$ a rational function of type $(n-1, n)$, and hence also of type (n, n).

This time, the integrand has Gaussian decay rather than just exponential as $|x| \to \infty$, so choosing $h = O(1/\sqrt{n})$ is enough to make the errors from endpoint truncation exponentially small. We also have the parameter a to play with (the details are spelled out in §16). By taking $a = O(\sqrt{n})$, we can make the errors due to grid spacing exponentially small too, and in this fashion we can achieve geometric convergence. More precisely, the choices

$$a = \sqrt{\frac{\pi n}{24}}, \quad h = \sqrt{\frac{3\pi}{2n}} \tag{14.5}$$

lead to the convergence rate

$$\|f - r\| = O(e^{-\pi n/3}) \approx O(2.849^{-n}). \tag{14.6}$$

This has the optimal geometric behavior up to a constant factor of about 2 in the sense that n has to be about twice the value of n in the best approximation case to reach the same error.

Alternative applications of the trapezoidal rule to approximation of e^t are discussed in §16, and the underlying rational approximations are investigated in [171]. In particular, contour plots in the complex plane are presented in [171] comparing best rational approximations to e^x on $(-\infty, 0]$ with approximations derived from the trapezoidal rule on hyperbolic, parabolic, and Talbot Hankel contours as discussed in §16.

15. Exponential and double exponential quadrature rules. When a function $y(\xi)$ defined on $[-1, 1]$ is smooth, there are all kinds of ways to integrate it efficiently, such as Gauss–Legendre or Clenshaw–Curtis quadrature [169, Chapter 19]. If there are singularities at the endpoints, however, such methods converge very slowly. One of the striking applications of the trapezoidal rule is to the integration of functions with endpoint singularities, by first transplanting $y(\xi)$ to a function $w(x)$ on the real line as in §6 or sometimes to a periodic function $v(\theta)$ on an interval as in §4. Methods of this kind were introduced in the 1960s [141, 146] and developed much further in the 1970s by Mori, Takahasi, and other Japanese researchers [83, 126, 129, 160, 161, 162].[13] They can be amazingly effective at treating nearly arbitrary endpoint singularities. For a very readable personal account, see Mori's recollections [128].

[13] Of course, the idea of changing variables to improve an integral originated long before the 20th century. An example close to our interests can be found in [125, p. 35].

We shall mainly discuss methods of the first type, based on a mapping $\xi = \phi(x)$ from $\mathbb{R}$ to $[-1,1]$, which alters the integral according to the formula

$$\int_{-1}^{1} y(\xi)\,d\xi = \int_{-\infty}^{\infty} y(\phi(x))\phi'(x)\,dx. \tag{15.1}$$

A possible mapping, proposed in [146] and [161], is the

$$\textit{tanh rule:}\quad \xi = \tanh(x). \tag{15.2}$$

One can see the power of this change of variables by considering the example

$$\int_{-1}^{1} \frac{d\xi}{\sqrt{1-\xi^2}} = \int_{-\infty}^{\infty} \frac{dx}{\cosh(x)}. \tag{15.3}$$

On the left, there are inverse-square-root singularities at both ends—integrable, but not even bounded. On the right, the integrand decays like $O(e^{-|x|})$ as $|x| \to \infty$, and it is analytic in a strip of half-width $a = \frac{1}{2}\pi$. By Theorem 6.1, the trapezoidal rule will converge geometrically at a rate $O(\exp(-\pi^2/h))$, and if it is truncated to $N = 2n+1$ points as in (6.4), the choice $h = \pi/\sqrt{n}$ as in the first row of Table 7.1 gives $\exp(-\pi\sqrt{n})$ convergence.

Square root singularities could be integrated in many ways, such as Gauss–Jacobi quadrature. The point about transplantation methods is that they are very general, requiring no particular structure in the integrand. For the tanh rule, the following theorem shows that the root-exponential convergence just described is typical. In practice, almost any integrand of interest will satisfy the conditions of this theorem, since it places no restriction on the nature of any branch cuts or other singular behavior along $(-\infty,-1]$ and $[1,\infty)$.

THEOREM 15.1. *Let $y(\xi)$ be analytic in a neighborhood of $(-1,1)$ that includes the intersection of a neighborhood of $\xi = -1$ with the wedge $-a < \arg(\xi+1) < a$ for some $a > 0$, where it satisfies $y(\xi) = O(|\xi+1|^\tau)$ for some $\tau > -1$, and similarly at $\xi = 1$. If $h = cn^{-1/2}$ for some $c > 0$, then as $N = 2n+1 \to \infty$, for some $C > 0$,*

$$I_N - I = O(e^{-C\sqrt{N}}). \tag{15.4}$$

Proof. For any a with $0 < a < \pi$, the function $\xi = \tanh(x)$ maps the infinite strip $-a < \mathrm{Im}(x) < a$ in the x-plane to the lens-shaped region about $(-1,1)$ in the ξ-plane bounded by circular arcs intersecting at ± 1 with half-angle a. The result follows from this observation together with Theorem 6.1. Also see citeLuBo92. □

The tanh transformation has been applied to good effect by many authors and is at the heart of the sinc function techniques used by Stenger and others for the solution of ODEs and PDEs and related problems [155, 157]. Its weakness is the root-exponential convergence, which one would like to improve upon. With this in mind, Takahasi and Mori also proposed two other transformations in [161] and [162], respectively:

$$\textit{erf rule:}\quad \xi = \mathrm{erf}(x), \tag{15.5}$$

$$\textit{tanh-sinh}\ \text{or}\ \textit{DE rule:}\quad \xi = \tanh(\tfrac{1}{2}\pi\sinh(x)) \tag{15.6}$$

(the constant $\frac{1}{2}\pi$ is justified below). The first of these maps has Gaussian behavior, and the second has exponential-of-exponential or "double exponential" behavior. We can see these effects by noting that the maps transform (15.3) to

$$\frac{2}{\sqrt{\pi}}\int_{-\infty}^{\infty}\frac{\exp(-x^2)}{(1-\operatorname{erf}^2(x))^{1/2}}\,dx, \qquad \frac{\pi}{2}\int_{-\infty}^{\infty}\frac{\cosh(x)}{\cosh(\frac{1}{2}\pi\sinh(x))}\,dx,$$

respectively, with integrands of size $O(\exp(-\frac{1}{2}x^2))$ and $O(\exp(-\frac{1}{4}\pi e^{|x|}))$ as $|x|\to\infty$. If truncation and discretization errors are balanced as in §7, one obtains the excellent convergence rates summarized in Table 15.1.[14]

TABLE 15.1
The three most commonly used transplantations from $[-1,1]$ *to* $\mathbb{R}$. *The convergence rates apply to many integrands, though precise theorems require careful hypotheses* [139, 174].

Name	Formula	References	Convergence
tanh	$\xi=\tanh(x)$	[146, 155, 161]	$\exp(-cN^{1/2})$
erf	$\xi=\operatorname{erf}(x)$	[161]	$\exp(-cN^{2/3})$
tanh-sinh or DE	$\xi=\tanh(\frac{1}{2}\pi\sinh(x))$	[162]	$\exp(-cN/\log(N))$

Of the methods we have mentioned, tanh-sinh quadrature with its nearly-geometric convergence has attracted the most attention. Bailey, Borwein, and their collaborators have applied it with remarkable effect to problems of high-precision experimental mathematics [12, 13, 14], and in [14] they write,

> It is the nearest we have to a truly all-purpose quadrature scheme at the present time.

For most of the remainder of this section, we explore properties of the tanh-sinh method, highlighting both its power and its limitations in floating point arithmetic, a problem that to varying degrees afflicts most quadrature methods based on variable transformations. At the end, we comment briefly on some of the other methods.

Consider the quadrature rule

$$\int_{-1}^{1} y(\xi)\,d\xi=\int_{-\infty}^{\infty} y(\phi(x))\phi'(x)\,dx\approx h\sum_{k=-n}^{n} y(\phi(kh))\phi'(kh), \tag{15.7}$$

where for some $\mu>0$,

$$\phi(x)=\tanh(\mu\sinh(x)), \qquad \phi'(kh)=\mu\cosh(x)\operatorname{sech}^2(\mu\sinh(x)),$$

and $N=2n+1$. We now explain why $\mu=\frac{1}{2}\pi$ is optimal for the class of entire functions y. The larger μ is, the faster the integrand decays. However, the mapping introduces poles where $\cosh(\mu\sinh(t))=0$. The two poles nearest the real axis are located at $x=\arcsin(\frac{1}{2}\pi/\mu)i$ and its conjugate. The maximum distance these poles can be from the real axis is $\frac{1}{2}\pi$, which is the case for all $0<\mu\le\frac{1}{2}\pi$; as μ increases beyond the value $\frac{1}{2}\pi$, the poles approach the real axis. The choice $\mu=\frac{1}{2}\pi$ therefore

[14]The reader is invited to confirm that for the example (15.3), the optimal step sizes for the three transformations are approximately $h=\pi n^{-1/2}$, $2.92n^{-2/3}$, and $n^{-1}\log(4\pi n)$, leading to convergence rates of $\exp(-\pi n^{1/2})$, $\exp(-4.28n^{2/3})$, and $\exp(-\pi^2 n/\log(4\pi n))$, respectively. Here $N=2n+1$. For the erf transformation, it will be necessary to know that the complex roots of $\operatorname{erf}(x)+1=0$ closest to the real axis are located approximately at $1.3548\pm 1.9915i$; cf. [161].

maximizes the width of the strip of analyticity and at the same time guarantees rapid decay of the integrand.

To estimate the optimal step size for (15.7), we match discretization and truncation errors as usual. We have seen that the half-width of the strip of analyticity for entire functions is $\frac{1}{2}\pi$, but singularities of $y(\phi(x))$ in the x-plane may limit it further. Hence we estimate the discretization error as usual by $O(\exp(-2\pi a/h))$, where now $a \le \frac{1}{2}\pi$. The truncation error can be estimated as $\phi'(nh) = O(\exp(-\frac{1}{2}\pi e^{nh}))$. By balancing the two we obtain the formula

$$2\pi\frac{a}{h} = \frac{1}{2}\pi e^{nh}, \tag{15.8}$$

which leads to

$$h = \frac{W(4an)}{n}, \tag{15.9}$$

where W is the Lambert W-function.[15] This choice of step size perhaps does not appear in the literature except in [139]; instead, in the original paper [162] and later, (15.8) is solved by asymptotic iteration after taking the logarithm. With a starting value $h = 1/n$, one improvement leads to

$$h = \frac{\log(4an)}{n}, \tag{15.10}$$

which agrees with (15.9) in view of the fact that $W(x) \sim \log(x)$ as $x \to \infty$. Despite this equivalence for large n, it turns out that (15.9) often leads to markedly superior results for intermediate n, as we shall see below. In either case, the convergence is asymptotically of the form $O(\exp(-2\pi an/\log(n)))$ as $n \to \infty$.

To appreciate the difference between (15.9) and (15.10), it is interesting to consider the integration of a constant function,[16]

$$\int_{-1}^{1} 1\,d\xi = \frac{\pi}{2}\int_{-\infty}^{\infty} \frac{\cosh(x)}{\cosh^2(\frac{1}{2}\pi\sinh(x))}\,dx. \tag{15.11}$$

Here $a = \frac{1}{2}\pi$, as mentioned above. With $n = 12$, for example, the formulas (15.9) and (15.10) specify that the interval should be truncated at the quite different values $|x| = 3.56$ and 4.32, respectively. In the latter case, the integrand is as small as 8.5×10^{-50} at the truncation point, which is quite incommensurate with the expected discretization error of $\exp(-\pi^2/h) = 1.3\times10^{-12}$. Using (15.9) produces the more balanced estimates 4.7×10^{-15} and 5.9×10^{-17} for the truncation and discretization errors, respectively.

In deriving (15.9) and (15.10), we assumed that the function y itself makes no appreciable contribution to the decay rate. If $|y(\xi)|$ decreases to 0 as $\xi \to \pm1$, the transformation can be modified; see [162] for a suggestion. Consider, for example,

$$\int_{-1}^{1} \sqrt{1-\xi^2}\,d\xi = \frac{\pi}{2}\int_{-\infty}^{\infty} \frac{\cosh(x)}{\cosh^3(\frac{1}{2}\pi\sinh(x))}\,dx. \tag{15.12}$$

[15] $W(x)$ is defined implicitly by $we^w = x$. For its properties and applications, see [32].

[16] LNT first learned of some of these quadrature formulas as a graduate student visiting Peter Henrici at the ETH in Zurich in 1979. Henrici introduced him to Jörg Waldvogel and laughed, "Dr. Waldvogel has a new quadrature formula that doesn't even integrate constants right!"

Comparison with (15.11) shows that $a = \frac{1}{2}\pi$ as before, but the decay as $|x| \to \infty$ has now improved from $O(\exp(-\frac{1}{2}\pi e^{nh}))$ to $O(\exp(-\frac{3}{4}\pi e^{nh}))$. This example shows that, curiously, a singular function such as $\sqrt{1-\xi^2}$ can in fact be more suitable for this type of integration than an entire function such as a constant, even if the difference is only slight. Pursuing arguments such as these, it is possible to show that if $y(\xi)$ is of the form $(1-\xi^2)^{\alpha-1}g(\xi)$ for some $\alpha > 0$, where $g(\xi)$ is bounded away from 0 at $\xi = \pm 1$, the estimates (15.8)–(15.9) can be improved to $2\pi a/h = (\pi/2)\alpha e^{nh}$ and $h = W(4an/\alpha)/n$, with a small gain in accuracy. If $y(\xi)$ has a zero at only one of the endpoints, one can exploit this by truncating the sum (15.7) at different values of n left and right.

The discussion above was rather informal and intended for practical situations. A rigorous error bound of the form $C\exp(-c_1 n/\log(c_2 n))$ for (15.7) can be established, but for functions $y(\xi)$ in a certain class. This is the class of functions defined on $(-1,1)$ that can be extended analytically into a domain defined by the image of the strip $|\mathrm{Im}(x)| < a$ under the transformation $= \tanh(\frac{1}{2}\pi\sinh(x))$, which is an infinitely sheeted Riemann surface that wraps around $\xi = \pm 1$. A proof can be found in [164], where related theorems for other single and double exponential transformations are also given. Further variations are considered in [1, 139].

As mentioned above, the tanh–sinh transformation has been a particular success in the field of experimental mathematics, where integrals are required to accuracies much higher than standard double precision [12]. It owes its success to its fast convergence, robustness against endpoint singularities, and simplicity, since the weights and nodes are explicitly given, unlike those of Gauss–Legendre quadrature, for example. For a comparison of the performance of tanh–sinh transformation with erf transformation and Gauss quadrature in the high-accuracy setting, we refer to [13].

Motivated by these high-accuracy computations, let us consider as an example an integral from [12],

$$\int_0^1 (\xi+1)^{-1}\log^6(\xi)\arctan\Big(\frac{\sqrt{3}\xi}{2-\xi}\Big)\,d\xi = 4.742841654850862\ldots. \tag{15.13}$$

Convergence curves for the tanh–sinh rule for different choices of h are shown in Figure 15.1. To quantify the convergence rates seen, we note that the half-width of the strip of analyticity is determined by the point where the argument of the arctan term in (15.13) assumes the values $\pm i$; the value is $a = \arcsin(2/3) \approx 0.7297$. Two of the curves in the figure correspond to the choices (15.9) and (15.10), reaffirming the superiority of the one based on the Lambert function. In the third and best error curve, we have taken into account the fact that unequal truncation points in (15.7) are advantageous for this integrand; see the discussion above and below (15.12). In particular, the sixfold zero at $\xi = 1$ in (15.13) allows one to truncate (15.7) at $n = -n_-$ and $n = n_+$ with $n_+ \ll n_-$. By balancing discretization and truncation errors, it is possible to find the optimal relationship between n_- and n_+, which was used to generate the error curve (c) in Figure 15.1 (with $N = 1 + n_- + n_+$).

The great drawback of exponential and double exponential quadrature rules is their difficulties with underflow, overflow, and cancellation errors in floating point arithmetic. Using the erf rule (15.5), for example, the integrand of (15.3) transforms to $g(x) = (2/\sqrt{\pi})e^{-x^2}(1-\mathrm{erf}^2(x))^{-1/2}$. In standard double precision arithmetic, the domain is limited to $|x| \le 6$, approximately, because for larger $|x|$ the value of $\mathrm{erf}(x)$ rounds precisely to 1 and the integrand becomes effectively singular. Yet the true value is $g(6) \approx 4 \times 10^{-8}$, which means the truncation error can be made no smaller

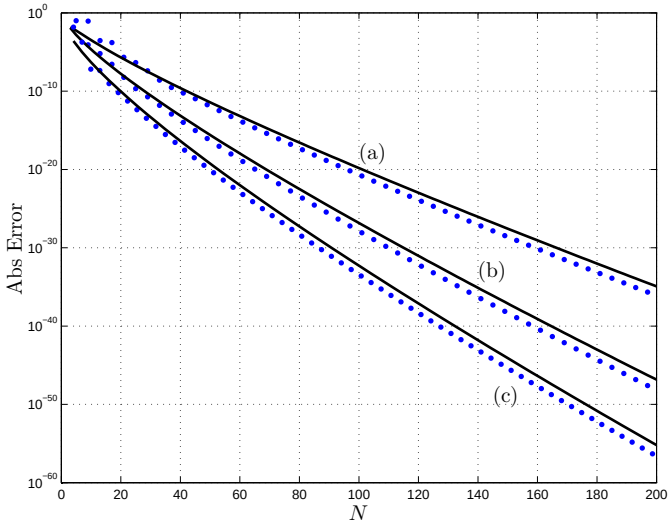


FIG. 15.1. *Convergence of* tanh–sinh *quadrature applied to the integral* (15.13) *after linear scaling to* $[-1,1]$. *Cases* (a) *and* (b) *correspond to symmetric truncations of* (15.7) *with* h *given by* (15.10) *and* (15.9), *respectively. For case* (c), h, n_+, *and* n_- *have been chosen corresponding to a nonsymmetric truncation of* (15.7) *to exploit the different decay rates of the integrand as* $x \to \pm\infty$. *In each case the curves show theoretical approximations based on balancing various error estimates and the dots are the actual computed errors.*

than this number. In this case the remedy is easy, which is to exploit the identity $1 - \mathrm{erf}(x) = \mathrm{erfc}(x)$. However, not all integrands can be adjusted in such a way, and certainly not without user intervention. This difficulty has prevented variable transformation methods from becoming general tools in floating point arithmetic.

The mappings of Table 15.1 all map $[-1,1]$ to $\mathbb{R}$. The so-called periodizing transformations, on the other hand, map $\mathbb{R}$ to $[-\pi,\pi]$. One example is the

$$\textit{tan rule:} \quad x = \tan(\theta/2), \tag{15.14}$$

intended for slowly decaying integrands on $\mathbb{R}$. The following is a test example from [162]:

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^4} = \frac{1}{2}\int_{-\pi}^{\pi} \frac{1+\tan^2(\theta/2)}{1+\tan^4(\theta/2)}\,d\theta. \tag{15.15}$$

The transplanted integrand is 2π-periodic, and the singularities at $\theta = \pm\pi$ are removable, so Theorem 4.2 predicts a convergence rate of $O(e^{-aN}) \approx O(e^{-0.88N})$, where $a = \mathrm{arctanh}(2^{-1/2})$. This agrees perfectly with the numerical results reported in [162]. The singularities at $\theta = \pm\pi$ in (15.15) are not always removable, however, as is shown by the example

$$\int_{-\infty}^{\infty} \frac{e^{-x^2}}{1+x^2}\,dx = \frac{1}{2}\int_{-\pi}^{\pi} e^{-\tan^2(\theta/2)}\,d\theta. \tag{15.16}$$

Here, the essential singularities at $x = \pm\infty$ have been transplanted to $\theta = \pm\pi$, and Theorem 4.2 is not applicable as there is no analyticity in any strip around $[-\pi,\pi]$. One still obtains algebraic converence of all orders since the derivatives at $\theta = \pm\pi$ vanish. This particular integral has been analyzed in [178], where it was shown that the convergence rate is $O(e^{-\frac{3}{2}N^{2/3}})$, similar to the second example in Table 7.1.

There are also periodizing transformations that map finite intervals to finite intervals. The first of these, one of the earliest examples of a transformation method, is by Sag and Szekeres [141]. Perhaps the most famous is the IMT rule by Iri, Moriguti, and Takasawa, given not in explicit form but as an indefinite integral [83].

Polynomial-based periodizing transformations have been proposed by Korobov and Laurie [92, 101], and a $\sin^m$-transformation by Sidi [149, 150]. Because these are not exponential transformations, they do not achieve exponential accuracy but rather high algebraic order $O(h^p)$. At the same time they do not suffer to the same extent from floating point difficulties because they do not space the points as densely as the exponential rules do. These transformations have been used in the integration of multivariate functions [92].

Finally, we mention the mapping $X = \sinh(x)$, which maps $\mathbb{R}$ to $\mathbb{R}$ with the purpose of improving the decay of slowly converging integrals such as (15.15). We shall apply this transformation for solving a PDE in §17.

16. Laplace transforms and Hankel contours. The Laplace transform and its inverse are defined by

$$\widehat{f}(s) = \int_0^\infty e^{-st} f(t)\,dt, \quad f(t) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} e^{st}\widehat{f}(s)\,ds, \quad \sigma > \sigma_0, \tag{16.1}$$

assuming convergence of the integrals, where σ_0 is a number known as the convergence abscissa. The second formula, known as the Bromwich integral, is the one that has to be evaluated numerically when tables of Laplace transforms fall short. Two of its key properties are these:

(i) The factor $\widehat{f}(s)$ is analytic not only in the half-plane $\mathrm{Re}(s) > \sigma_0$, but also, in many cases, in much of the rest of the complex plane.
(ii) The factor e^{st} is likewise analytic and decays rapidly as $\mathrm{Re}(s) \to -\infty$.

These properties allow one to deform the Bromwich line $\mathrm{Re}(s) = \sigma$ into a contour that is better suited for computation, exploiting the rapid decay to turn the integral into the kind treated by the trapezoidal rule in §§4–7 and §15. For all kinds of applications, this idea has proved effective. Many of the examples of §17 and §18 are of this type.

The idea of deforming the contour and applying the trapezoidal rule seems to have originated independently in two different hemispheres in the mid-1950s. In England, Alan Talbot and his doctoral student J. S. Green considered integrating the Bromwich integral on a steepest descent contour [67]. Around the same time, at a conference in Australia, John Butcher suggested a Hankel contour in the form of a parabola

> ... which bends towards the negative half-plane [thus introducing] an exponentially decreasing factor which may speed up the overall convergence of the integral [27].

Butcher never returned to this work, but Talbot did, and twenty years later he published a much cited paper [163] in which he generalized and improved the earlier work of Green.

Consider a Hankel contour parameterized by a real variable x,

$$\Gamma: \ s = s(x), \quad x \in \mathbb{R}, \tag{16.2}$$

with the property that $s(\pm\infty) = -\infty \pm i\gamma$ for some positive γ (possibly infinite). Examples can be seen in Figures 1.2, 16.2, and 16.3. The line $\mathrm{Re}(s) = \sigma$ in (16.1)

can be deformed to the contour Γ, provided it remains in the domain of analyticity of $\widehat{f}(s)$. There is also a restriction on the decay of $\widehat{f}(s)$ in the left half-plane, namely that $\widehat{f}(s) \to 0$ uniformly in $\mathrm{Re}(s) \le \sigma_0$ as $|s| \to \infty$ [163, 171]. This rules out transforms such as $\widehat{f}(s) = e^{-s}$ or $e^{-\sqrt{s}}$, which correspond to functions $f(t)$ with singularities at $t = 0$.

The contour deformation (16.2) leads to

$$f(t) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} e^{s(x)t} \widehat{f}(s(x)) s'(x)\, dx, \tag{16.3}$$

and approximation by the truncated trapezoidal rule yields

$$f(t) \approx \frac{h}{2\pi i} \sum_{k=-n}^{n} e^{s_k t} s'_k \widehat{f}_k. \tag{16.4}$$

Here h is the uniform node spacing, and

$$x_k = kh, \quad s_k = s(x_k), \quad s'_k = s'(x_k), \quad \widehat{f}_k = \widehat{f}(s_k). \tag{16.5}$$

If it is more convenient to use the midpoint rule, then $x_k = (k + \frac{1}{2})h$ and the upper limit of the sum (16.4) is $n - 1$ instead of n.

Observe that if $\widehat{f}(s)$ is real-valued when s is real, and if the contour Γ is symmetric with respect to the real s-axis, then the terms in the sum (16.4) appear as conjugate pairs, which cuts the cost in half. This is a significant savings in PDE problems, for example, where each evaluation of $\widehat{f}_k$ requires the solution of a large linear system.

Just as there are many quadrature rules that can be applied to (16.3), so there are many choices for the Hankel contour (16.2). Most of these can achieve geometric convergence as $n \to \infty$ if certain parameters are well chosen. Our plan of the remainder of this section is, first, to outline five particular choices of contours, summarized in Table 16.1: parabolas, hyperbolas, circles, and so-called Talbot contours and modified Talbot contours. The first two are based on transplantations to the real line as in (16.3), and the other three on transplantations to $[-\pi, \pi]$. Then, focussing on the simplest case of parabolas, we will show the kind of analysis that can be used to derive the optimal parameters listed in the table.

To show in the most concrete way how effective these formulas can be, we consider as an example the plot of $|\Gamma(z)|$ shown in Figure 16.1, the absolute value of the gamma function in the complex plane. This example is based on the Laplace transform formula

$$\frac{t^{z-1}}{\Gamma(z)} = \frac{1}{2\pi i} \int_{\Gamma} e^{st} s^{-z}\, ds, \tag{16.6}$$

which can be found in most tables of integrals. Following [144, 167], Figure 16.1 has been computed by evaluating this integral, with $t = 1$, on a modified Talbot contour discretization with 32 points. The figure shows $|\Gamma(z)|$ on a grid with spacing 0.1 in each direction of the complex plane. (The symmetry with respect to the real axis in the figure is not exploited.) The complete Matlab code (except for plotting commands) is as follows:

TABLE 16.1

Choices of contours for the numerical inversion of the Laplace transform by the approximation (16.4)*, showing the best asymptotic convergence rates in the case where* $\widehat{f}(s)$ *has singularities on the negative real axis and* t *is fixed. Formulas for the parabola and hyperbola are given after scaling the trapezoidal formula to* $\theta \in [-\pi, \pi]$ *via* $x = (nh/\pi)\theta$*. If* $\widehat{f}$ *has real symmetry, only nodes in* $[0, \pi]$ *are used; if not, nodes in* $\theta \in [-\pi, \pi]$ *are needed and the listed convergence rate cuts effectively in half. For comparison, the final line lists the asymptotic convergence rate of algorithms based on the use of actual best* ∞*-norm rational approximations of* e^s *on* $(-\infty, 0]$*, an idea that shows promise but has received relatively little study.*

Name, eq.	Convergence	Refs.	Formula: $s = (n/t)\times$
parabola (16.8)	$O(e^{-2.09n})$	[181]	$0.2618 - 0.2387\theta^2 + 0.5i\theta$
hyperbola (16.9)	$O(e^{-2.32n})$	[181]	$4.4921(1 - \sin(1.1721 - 0.3443i\theta))$
circle	$O(e^{-0.91n})$	–	$-0.3533 + 0.5569\exp(i\theta)$
Talbot (16.7)	$O(e^{-1.90n})$	[179]	$-0.4814 + 0.6443\theta\cot(\theta) + 0.3642i\theta$
mod. Talbot (16.7)	$O(e^{-2.72n})$	[38]	$-1.2244 + 1.0034\theta\cot(0.6407\theta) + 0.5290i\theta$
best approx.	$O(e^{-4.46n})$	[29, 171]	

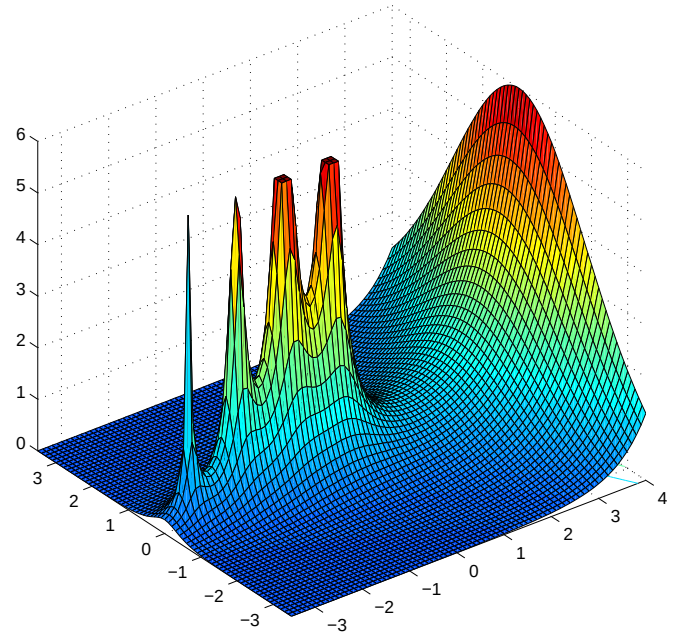


FIG. 16.1. *Modulus of the gamma function in the complex plane computed by a short Matlab code segment based on a* 32*-point modified Talbot contour discretization of* (16.6)*, with* $t = 1$*, following* [144, 167]*. The computation gives* 10 *to* 15*-digit accuracy over most of the domain and takes less than a half a second on an ordinary laptop.*

```
n = 16;
th = [-n+0.5:n-0.5]*pi/n;                              % Nodes in [-pi,pi]
sg = -1.2244; mu = 1.0034; nu = 0.5290i; bt = 0.6407; % Modified T. parameters
s =  n*(sg+mu*th.*cot(bt*th)+nu*th);                   % Talbot pts in s-plane
sp = n*(mu*(cot(bt*th)-bt*th./sin(bt*th).^2)+nu);      % Derivative of s
x = linspace(-3.5,4,76); y = linspace(-2.5,2.5,71);
[xx,yy] = meshgrid(x,y);  zz = xx+1i*yy;               % Plotting grid
```

```
gi = zeros(size(zz));
for k = 1:2*n;                                          % Loop over contour pts
     gi = gi+exp(s(k))*sp(k)*s(k).^(-zz);
end
gi = 1/(2i*n)*gi;                                       % 1/Gamma(z) on grid
gamma = 1./gi;                                          % Gamma(z) on grid
```

If the function is costly to evaluate at each node, as would be the case when solving PDEs (§17), the **for** loop can be replaced by a parallel **for** loop.

This code segment is based on the contour

$$s = \sigma + \mu\theta\cot(\beta\theta) + \nu i\theta, \quad \theta \in [-\pi, \pi]. \tag{16.7}$$

Note that we have implemented the midpoint rule version, as it avoids the need to compute the value of s at the removable singularity at $\theta = 0$.

The contour (16.7) with $\beta = 1$ is the original contour of Talbot [163]. The version with $0 < \beta < 1$ is known as the modified Talbot contour [38]. The optimal choices of parameters in the modified contour appear in the code and in Table 16.1, where the expected rate of convergence is also listed. At a fixed value of z, say $1 + i$, the absolute errors in $1/\Gamma(z)$ for $n = 5$ and $n = 10$ are about 1.6×10^{-5} and 2.2×10^{-11}, respectively. This represents an empirical convergence rate of $O(e^{-2.70n})$, which is in excellent agreement with the theoretical $O(e^{-2.72n})$ listed in the table.

Another idea listed in Table 16.1 is a parabolic contour,

$$s = (ix + a)^2, \quad x \in \mathbb{R}, \tag{16.8}$$

with $a > 0$, already used in §14. This idea goes back to Butcher [27], and optimal parameters were derived in [181].

A third idea is a hyperbola, and the contour

$$s = \mu(1 + \sin(ix - \alpha)), \quad x \in \mathbb{R}, \tag{16.9}$$

was considered in [104, 105]. (This formula does indeed describe a hyperbola, as becomes clear when the sine term is expanded in its real and imaginary parts.) The parameter $\alpha \in (0, \frac{1}{2}\pi)$ can be adjusted to control the asymptotic angle of the hyperbola, which makes this a popular choice for transforms $\widehat{f}$ with singularities in a wedge-like region around the negative real axis. These arise when solving differential or integral equations involving sectorial operators §17. (However, in practice it does not matter much if one uses a hyperbolic, parabolic, or other contour for inverse Laplace transforms of sectorial operators, since the $e^{st} < 10^{-16}$ for $\mathrm{Re}\, s < -37$. In machine arithmetic, the complex plane is a half-plane.)

One way to compare these various contours is to view their connection with the double exponential quadrature formulas of §15.[17] To see the connection, note that on the parabolic contour (16.8) and the hyperbolic contour (16.9), the integral (16.3) has Gaussian and double exponential decay, respectively, as can be seen from

$$|e^{st}| = O(e^{-tx^2}), \quad |e^{st}| = O(e^{-\frac{1}{2}\mu t(\sin\alpha)e^{|x|}}), \quad |x| \to \infty. \tag{16.10}$$

[17] It is curious to note that the people working on the inverse Laplace transform (Butcher, Green, Talbot) and quadrature by variable transformation (Mori, Schwartz, Takahasi) were doing effectively the same thing around the same time, but they seem to have been unaware of each other.

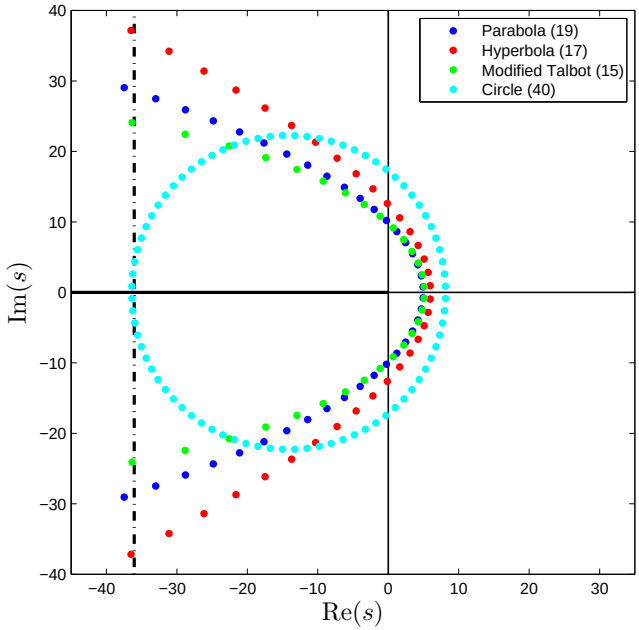


FIG. 16.2. *Quadrature nodes for four of the contours with parameters optimized to achieve the convergence rates shown in Table* 16.1. *In each case the value of n (shown in the legend) is chosen so that the accuracy is approximately* 10^{-16}, *as marked by the vertical dashed line. To avoid clutter the actual contours (connecting the dots) are not drawn.*

This is assuming that the transform $\widehat{f}(s)$ does not contribute to the decay rate. The original Talbot contour, eq. (16.7) with $\beta = 1$, corresponds to faster than double exponential decay, namely

$$|e^{zt}| = O(e^{-2\pi^2\mu t/(\pi^2-\theta^2)}), \quad |\theta| \to \pi. \tag{16.11}$$

The decay here is like that of a bump function with essential singularities at $\theta = \pm\pi$, which makes it similar to the IMT rule (§15). This decay is too fast, and the reason for considering $0 < \beta < 1$ in (16.7) is precisely to control it better [38, 179].

Figure 16.2 shows most of the contours we have discussed. The figure illustrates the conclusion, also evident in Table 16.1, that while the modified Talbot method achieves a given accuracy with the least number of nodes, there is not a great difference between the various Hankel contours.

The circle is simple and has been used to good effect on the $\Gamma(z)$ example; see [167]. As it is not a true Hankel contour, it commits a branch cut "crime" in crossing the negative real axis. By moving the crossing point further into the left half-plane, however, the magnitude of the crime can be made as small as desired, and to derive the optimal center and radius listed in Table 16.1, we balanced it against the discretization and truncation errors. The circle is inferior to the Hankel contours not because of the branch crime, but because it wastes so many of its nodes far in the left half-plane. Using an optimal half-circle improves the convergence rate marginally, to $O(e^{-1.11n})$.

As mentioned earlier, we will now conclude the section by outlining an example of the kind of analysis by which good parameters for these contours can be found. Even though the parabola does not represent the fastest convergence in the present group

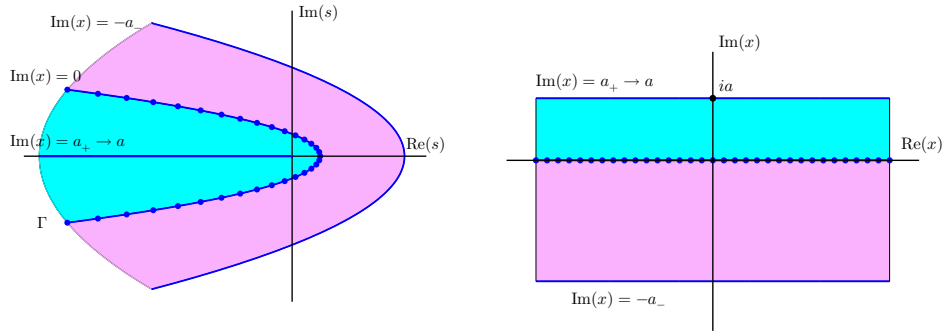


FIG. 16.3. *Conformal map of the parabola* (16.8). *The interior (resp. exterior) of the contour of integration* Γ *in the* s*-plane maps to the upper (resp. lower) half of the* w*-plane. Theorem* 6.1 *is applied to each half-plane to estimate errors and hence compute optimal parameters.*

of methods, we shall restrict our analysis to this contour because of its simplicity. Also, it leads to explicit formulas, which is not the case for the other contours.

We follow the argument of [179] and consider the conformal mapping (16.8), with x now extended into the complex plane. See Figure 16.3 for a schematic representation of this map. A strip of analyticity in the x-plane is mapped to the cut s-plane, with the cut along the negative real axis. Note that the strip on the right is bounded from above by the singularity at $\mathrm{Im}(x) = a$, but there is no restriction in the lower half-plane. The latter property is inherited from the fact that the integrand in the s-plane is analytic in the exterior of the parabola.

We now estimate discretization errors by applying Theorem 6.1; specifically, the generalization of that theorem to the case of an unsymmetric strip of analyticity, as discussed below (6.15). Hence, consider in Figure 16.3 (right diagram) the half-strips $\mathrm{Im}(x) \in [0, a_+]$ and $\mathrm{Im}(x) \in [-a_-, 0]$, where $0 < a_+ < a$, $0 < a_-$. The error associated with $[0, a_+]$ is $O(e^{-2\pi a_+/h})$, and if we assume the only singularities of $\widehat{f}(s)$ are on the negative real axis we can let $a_+ \to a$ to obtain $O(e^{-2\pi a/h})$. The error associated with $[-a_-, 0]$ is $O(e^{(a+a_-)^2 t} e^{-2\pi a_-/h})$, where the first and second exponential terms come from the factor M and the denominator of (6.6), respectively. Minimizing with respect to a_-, as was done below (5.2), gives the optimal value as $a_- = -a + \pi/(th)$, assuming h sufficiently small for this quantity to be positive. The corresponding error contribution is $O(e^{(\pi/th^2)(2ath-\pi)})$. Finally, there is also a truncation error of magnitude $O(e^{t(a^2-(hn)^2)})$. Balancing these three error terms gives

$$-\frac{2\pi a}{h} = \frac{\pi}{th^2}(2ath - \pi) = t(a^2 - (nh)^2), \tag{16.12}$$

from which it follows that

$$a = \frac{1}{2}\sqrt{\frac{\pi n}{3t}}, \quad h = \frac{1}{2}\sqrt{\frac{3\pi}{tn}}. \tag{16.13}$$

The corresponding convergence rate is $O(e^{-2\pi n/3})$. These formulas for a and h correspond to the formulas used in the rational approximation problem discussed at the end of §14, where $t = 1$ and n is replaced by $n/2$.

In Table 16.1 we chose to scale all these contours to the interval $\theta \in [-\pi, \pi]$, which is done by setting $x = (nh/\pi)\theta$. Substituting this as well as (16.13) into (16.8) gives $s = (n/t)(\pi/12 - 3/(4\pi)\theta^2 + i\theta/2)$, as recorded in the first line of the table.

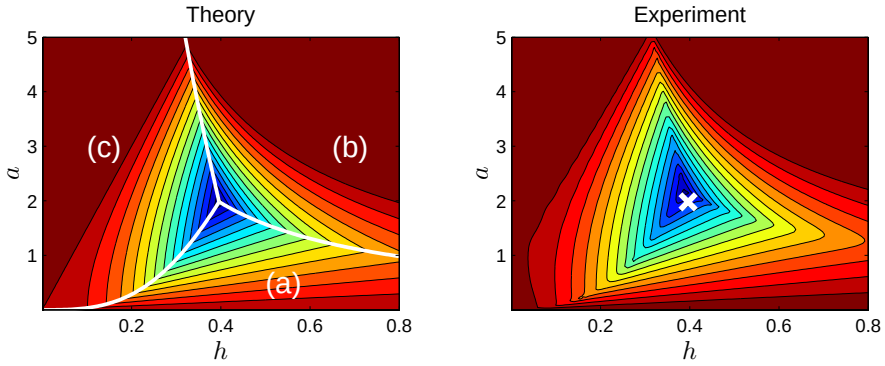


FIG. 16.4. *On the left, a graphical representation of the derivation of the optimal parameters* a *and* h *of* (16.13) *for the parabolic contour* (16.8). *In region* (a) *of the parameter plane, the discretization error associated with the singularities on* $(-\infty, 0]$ *dominates, of size* $O(e^{-2\pi a/h})$. *In* (b), *the discretization error associated with the magnitude of* e^{st} *in the right half* s*-plane dominates, of size* $O(e^{(\pi/th^2)(2ath-\pi)})$. *In* (c), *the truncation error from replacing the infinite sum by a finite one dominates, of size* $O(e^{t(a^2-(hn)^2)})$. *The white curves are the boundaries of the regions as defined by* (16.12). *The color bands represent absolute errors of magnitude* $10^{-\ell}$, $\ell = 0, 1, \ldots, 15$, *counting from the outside in. This is for the special case* $t = 1$ *and* $n = 15$, *with all implied constants set to* 1. *On the right, actual errors for the* $1/\Gamma(z)$ *example at* $z = 1 + i$, *using the parabolic contour with* $n = 15$. *The cross marks the location of the theoretical optimal parameters* (16.13).

In Figure 16.4 the theoretical analysis above is summarized and validated against an actual example, namely the computation of $1/\Gamma(z)$ discussed above. The agreement between theory and experiment is striking. This would be a good time, however, to point out the limitations of the analysis. It was assumed that the transform $\widehat{f}(s)$ made only $O(1)$ contributions to the various discretization and truncation error estimates; the only information of $\widehat{f}(s)$ that was actually used is the fact that its singularities are on the negative real axis. This is a valid assumption for a transform like $\widehat{f}(s) = (s+\lambda)^{-1}$, with $\lambda > 0$, which is the appropriate model problem for parabolic PDEs; see §17 and [181]. In the case of the gamma function, however, $\widehat{f}(s) = s^{-z}$ can have large or small magnitude when z is far in the left or right half-planes and the contour parameters (16.13) can become less reliable [143]. The same thing was pointed out in the case of the generalized Mittag-Leffler function in [54].

If $\widehat{f}(s)$ has singularities off the real axis, the contour parameters (16.13) are likewise not applicable. If the singularities lie in a parabolic region enclosing the negative real axis, however, the analysis can be adapted by limiting the value of a_+ by a value strictly less than a. This situation arises for example in solving linear advection-diffusion PDEs by the method described in §15. Formulas for optimal parabolas were derived in in [81, 180], where it was shown that the convergence is still geometric but the decay constant gets smaller the wider the region. In the extreme case when the singularities extend to infinity along the imaginary axis (as is the case in wave-type PDEs) all accuracy is lost, signalling the failure of these contour deformation methods.

An issue we have not discussed in this section is the fact that the optimal contours all scale with n; see Table 16.1. A contour that extends well into the right half-plane will therefore make the integrand of (16.1) large. The result can be destructive cancellation errors on a computer even when a formula should mathematically be

accurate. Roundoff control in the case of parabolas, hyperbolas, and Talbot contours is discussed in [180], [105], and [38], respectively.

Another issue worthy of mention is that the contour deformation technique we have described is not the only approach to the numerical inversion of the Laplace transform that makes use of the trapezoidal rule. A popular alternative is an expansion in Laguerre polynomials

$$f(t) = e^{\sigma t} \sum_{j=0}^{\infty} c_j e^{-bt} L_j(2bt), \quad t > 0, \tag{16.14}$$

known as the Weeks method [176]. Here σ is defined as in (16.1), b is a positive constant, and the expansion coefficients c_j are defined by the power series

$$\frac{2b}{1-z} \widehat{f}(s) = \sum_{j=0}^{\infty} c_j z^j, \quad s = \sigma + b\frac{1+z}{1-z}, \tag{16.15}$$

with radius of convergence determined by the singularities of $\widehat{f}$. For derivations of (16.15) we refer to [176, 177], both of which also provide guidelines for choosing the parameters σ and b. The coefficients c_j are computed from (16.15) by trapezoidal rule approximation and the FFT as described in §13; see also [113].

A closely related method is one for computing the Fourier integral (6.7) when $w(x)$ is slowly decaying on $\mathbb{R}$. (When it is rapidly decaying, (6.7) can simply be truncated.) This method, which leads to a Laguerre expansion for $\widehat{w}(\xi)$ similar to (16.14), is due to Weber [175].

Both the Weeks method for the Laplace transform and the Weber method for the Fourier transform are related to the tan rule (15.14). A clue to this is the fact that the argument of $\widehat{f}$ in (16.15) reduces to $s = \sigma + ib\cot(\theta/2)$ when $z = e^{i\theta}$. The use of the trapezoidal rule, however, is shifted away from the direct computation of the integrals, (6.7) or (16.1), to the computation of expansion coefficients as in (16.15).

17. Partial differential equations. The geometrically converging trapezoidal rule has been used in many ways in algorithms for the numerical solution of PDEs. In this section we concentrate on methods exploiting Hankel contour integrals as described in the last section.

Consider the system of ODEs

$$\frac{d\boldsymbol{u}}{dt} = A\boldsymbol{u}, \quad \boldsymbol{u}(0) = \boldsymbol{u}_0, \tag{17.1}$$

$\boldsymbol{u} \in \mathbb{C}^m$, $A \in \mathbb{C}^{m\times m}$, which arises upon semi-discretization of a PDE such as $u_t = \mathcal{A}u$. This could be the heat equation, with $\mathcal{A}$ the Laplacian operator in some coordinate system and A its matrix approximation based on finite differences, finite elements, or spectral methods. The solution of (17.1) is $\boldsymbol{u}(t) = \exp(At)\boldsymbol{u}_0$, so what we are about to describe can be seen as a method for computing the action of the matrix exponential on a vector without computing the exponential itself. More general functions $f(A)$ besides the exponential are the subject of §19.

Taking the Laplace transform of (17.1) with respect to t leads to the integral

$$\boldsymbol{u}(t) = \frac{1}{2\pi i} \int_{\sigma-i\infty}^{\sigma+i\infty} e^{st} \widehat{\boldsymbol{u}}(s)\, ds, \quad \widehat{\boldsymbol{u}}(s) = (sI - A)^{-1}\boldsymbol{u}_0, \tag{17.2}$$

where the contour passes to the right of all eigenvalues of A.[18] Deforming the contour to the type of Hankel curve described in §16 leads to

$$\boldsymbol{u}(t) \approx \frac{h}{\pi} \mathrm{Im} \left\{ {\sum_{k=0}^{n}}' e^{s_k t} s_k' \widehat{\boldsymbol{u}}_k \right\}, \tag{17.3}$$

where we have used the symmetry considerations alluded to below (16.5) to reduce the cost. The numbers s_k and s_k' are defined by (16.5), and the vectors $\widehat{\boldsymbol{u}}_k$ are the solutions of the equations

$$(s_k I - A)\widehat{\boldsymbol{u}}_k = \boldsymbol{u}_0, \quad k = 0, 1, \ldots, n. \tag{17.4}$$

The major expense in a practical implementation of (17.3) is the solution of the shifted systems (17.4), which are typically large and require complex arithmetic even when A is real. When the structure of A is such that these systems can be solved efficiently by a direct method, the summation (17.3) can be implemented in a parallel **for** loop as the summands may be computed independently.[19] When direct methods are impractical, Krylov subspace iteration can be considered. A key point is that because of shift-invariance, only one Krylov basis needs to be constructed for all the systems (17.4). This reduces the problem to solving, instead of (17.4), a sequence of upper-Hessenberg systems of much smaller dimension; see [123, 151]. For a comparison of Krylov methods with parallel direct methods we refer to [81], where the important issue of preconditioning is also addressed.

Another way to reduce the cost is to pick a good contour so that the number of linear systems (17.4) to be solved is small. If A arises from discretization of the heat equation, its spectrum is typically on the negative real axis, and an optimal contour can be constructed by the type of analysis that led to (16.13).

The formula (17.3) computes the solution directly at a given value of t. In certain applications this can give the method an advantage over the method-of-lines approach, in which (17.1) is integrated on all of $[0, t]$ by some Runge–Kutta or multistep formula. While (17.3) can be applied at two or more values of t for the same set of $\widehat{\boldsymbol{u}}_k$, one should keep in mind that the optimal contours are t-dependent; cf. (16.13) for example. If the contour is picked optimally for a particular value of t, the accuracy may drop significantly for other values of t; see [105, 181].

As far as we are aware, the first mention of the use of (17.2) as a method for solving the heat equation goes back to the original paper of Talbot [163], but no numerical results were reported there. Since then, numerous variants of the method have been proposed and numerical results presented for the heat equation and more general parabolic PDEs. Relevant papers include [56, 81, 180, 181] (parabolic contours), [105, 121, 148, 181] (hyperbolic contours), and [38, 171] (Talbot contours). A recent monograph devoted to the topic is [57].

The greatest impact of the contour integral method has been in the area of parabolic problems. However, there also exists a contour integral formulation for elliptic PDEs defined on cylindrical domains [57]. To illustrate this, consider the second-order analogue of (17.1),

$$\frac{d^2\boldsymbol{u}}{d\xi^2} + A\boldsymbol{u} = 0, \quad \boldsymbol{u}(0) = \boldsymbol{0}, \quad \boldsymbol{u}(1) = \boldsymbol{u}_1, \tag{17.5}$$

[18] More generally, the same formulas apply if A is a closed linear operator generating a C_0 semigroup [135].

[19] In [69] the term "backslash matrix" is proposed for such matrices.

with $\xi \in (0,1)$. If A is an approximation to an elliptic operator in d dimensions with suitable boundary conditions, the system of ODEs (17.5) is an approximation to an elliptic PDE in $d+1$ dimensions. Again one should think of A as a suitable discretization of the Laplacian. The assumption that the boundary conditions are independent of ξ restricts the problem to a cylindrical domain.

Taking the Laplace transform of (17.5), one obtains, in analogy to (17.2),

$$\boldsymbol{u}(\xi) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} E(\xi;s)\widehat{\boldsymbol{u}}(s)\,ds, \quad \widehat{\boldsymbol{u}}(s) = (sI-A)^{-1}\boldsymbol{u}_1, \tag{17.6}$$

where $E(\xi;s)$ is the normalized sine function

$$E(\xi;s) = \frac{\sin(\xi\sqrt{s})}{\sin(\sqrt{s})}. \tag{17.7}$$

The contour in (17.6) passes to the right of the spectrum of A, as before. What is new is that $E(\xi;s)$ now has singularities, which lie at $s=\pi^2$ and its positive multiples, putting a further restriction on the location of the contour. In practice one would therefore select a contour that intersects the real axis in $(0,\pi^2)$.

In [57] a method based on contour deformation of (17.6) is proposed and analyzed. Because the function $E(\xi;s)$ in (17.6) does not decay as rapidly in the left half-plane as the function e^{st} in (17.2), however, the contour deformation is perhaps less crucial here. Instead, one can solve the problem on a vertical line in combination with a sinh transformation as in §15. Based on the type of parameter tuning arguments seen elsewhere in the paper, we propose

$$s(x) = \frac{1}{2}\pi^2(1+i\sinh(x)), \quad x\in\mathbb{R}.$$

The approximation is then

$$\boldsymbol{u}(\xi) \approx \frac{h}{\pi}\,\mathrm{Im}\left\{ {\sum_{k=0}^{n}}' E(\xi;s_k)s_k'\widehat{\boldsymbol{u}}_k \right\}, \tag{17.8}$$

where $s_k = s(kh)$ and $s_k' = s'(kh) = \frac{1}{2}\pi^2\cosh(kh)i$, and $\widehat{\boldsymbol{u}}_k = (s_kI-A)^{-1}\boldsymbol{u}_1$. Note that in this method the dimension of the problems to be solved has been reduced by 1. A 2D elliptic problem can be solved by solving a sequence of 1D problems, namely one resolvent system for each quadrature node. Likewise a 3D problem can be reduced to solving 2D problems, etc. The method gives one the option of computing the solution at a specified value of ξ, without having to compute it through the entire domain as is the case with many other methods. Accuracy is lost as $\xi\to 1$, however. As with (17.2), the resolvent systems can be solved in parallel.

As an example, consider the last problem of the SIAM 100-Digit Challenge [23]. This involved a particle in a 10×1 rectangle undergoing a two-dimensional Brownian motion. Assuming the motion starts from the center, the problem was to compute the probability that the particle will reach one of the ends before it touches the sides. To make the problem more challenging, we consider here the 3D version, i.e., a rectangular box of dimensions $10\times 1\times 1$. See Figure 17.1.

The problem can be posed as a Laplace problem for the determination of a harmonic measure; see [23, Chapter 10]. We consider a box of dimensions $2b\times 2a\times 2a$, and making use of the symmetries, we reduce the problem to solving (17.5) on the

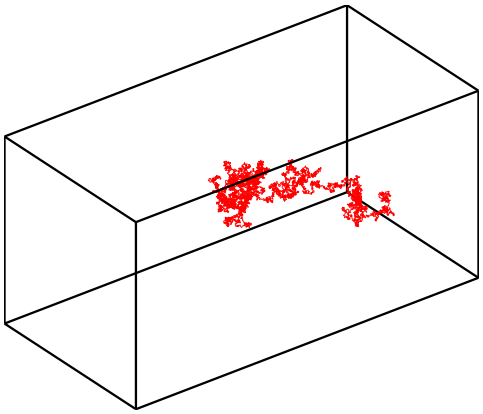


FIG. 17.1. *A particle undergoing Brownian motion in a rectangular box. A trajectory has been selected that hits an end before the sides; in this* $2 \times 1 \times 1$ *geometry the fraction of trajectories with this property is* ≈ 0.038. *For a* $10 \times 1 \times 1$ *box, the case calculated by a contour integral method in the text, this figure shrinks to* 7.30×10^{-10}.

domain $[-b, b] \times [0, a] \times [0, a]$ with boundary conditions 0 on one end and 1 on the other. Here A is an approximation to the Laplacian on the square $[0, a] \times [0, a]$ with Dirichlet conditions on two adjacent sides and Neumann conditions on the others. For A, we therefore use

$$A = I \otimes D + D \otimes I, \tag{17.9}$$

where D is the Chebyshev spectral differentiation matrix on $[0, a]$ that incorporates homogeneous Neumann and Dirichlet conditions at the left and right endpoints, respectively. The matrix I is the identity of the same dimension as D, and $\otimes$ denotes the Kronecker product [167].

This method yields a probability 7.298818×10^{-10}, which is correct to all digits shown (as verified by a series solution). This accuracy is achieved by using an 8×8 differentiation matrix D and 30 terms in the quadrature sum (17.8), with a step size of $h = 0.2773$ (as prescribed by the parameter tuning). Note that the total cost involved is the solution of 30 resolvent systems, each of size 64×64. If the same method is applied to the original 2D example, it yields the required ten digits, namely $3.837587979 \times 10^{-7}$, at the cost of solving 30 resolvent systems of dimension 8×8.

As a final example, consider a semi-discretization of a fractional-time diffusion equation

$$\frac{d^\alpha \boldsymbol{u}}{dt^\alpha} = A\boldsymbol{u}, \quad \boldsymbol{u}(0) = \boldsymbol{u}_0, \tag{17.10}$$

where the so-called Caputo fractional derivative on the left is defined for $0 < \alpha < 1$ by

$$\frac{d^\alpha \boldsymbol{u}}{dt^\alpha} = \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{\boldsymbol{u}'(y)}{(t-y)^\alpha} dy.$$

The analytic solution of (17.10) is given by

$$\boldsymbol{u}(t) = E_\alpha(t^\alpha A)\boldsymbol{u}_0, \quad \text{where} \quad E_\alpha(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k + 1)}.$$

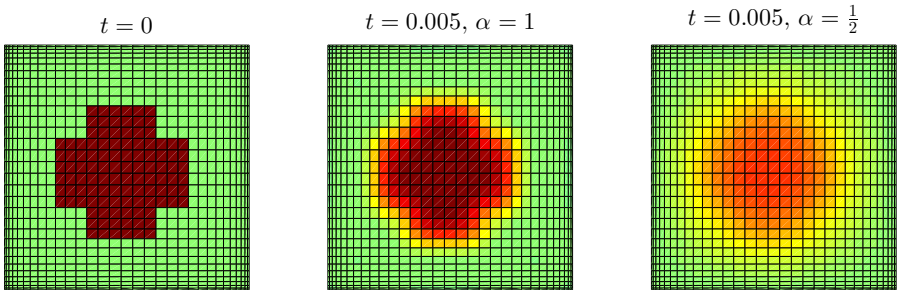


FIG. 17.2. *Diffusion under the equation* (17.10) *computed by the inverse Laplace transform discretized by the trapezoidal rule on a Hankel contour. The left image shows the initial condition. The middle, with* $\alpha = 1$*, shows ordinary diffusion. The right, with* $\alpha = 1/2$*, corresponds to diffusion governed by a fractional time derivative.*

The function $E_\alpha(z)$ is the Mittag-Leffler function, which reduces to e^z and $e^{z^2}\mathrm{erfc}(-z)$ in the special cases $\alpha = 1$ and $\alpha = \frac{1}{2}$, respectively; see [54] and [116].

By taking a Laplace transform of (17.10), one can derive a contour integral representation and its approximation identical to (17.2) and (17.3), but the transform is now [116]

$$\widehat{\boldsymbol{u}}(s) = (sI - s^{1-\alpha}A)^{-1}\boldsymbol{u}_0. \tag{17.11}$$

Note that if $\alpha = 1$, this reduces to the standard heat equation (17.2). On the other hand, when $0 < \alpha < 1$, the transform involves a branch cut. If this cut is defined to be on the negative real axis, the contours displayed in Table 16.1 should still be good choices.

As an experiment, we used the same Talbot contour displayed in the code segment of (16) to invert (17.11). For A we chose a spectral collocation approximation to the Laplacian on a tensor-product grid of 32×32 Chebyshev points; cf. (17.9). With $n = 16$ nodes on a midpoint discretization of the Talbot contour the results are as shown in Figure 17.2. Note that because of symmetry, only 8 resolvent systems were solved for each α, each of which involved a matrix of size $32^2 \times 32^2$.

The numerical examples presented in this section are primarily of academic interest. One area of more realistic applications is mathematical finance, where convection-diffusion equations such as the Black–Scholes and Heston PDEs can be solved by these techniques [81, 102].

A related integral transform method, in which the trapezoidal rule is applied on a hyperbolic contour, has been applied by Fokas and colleagues to solve a variety of evolutionary and stationary PDEs [46, 47, 48]. Another variation on the theme is the use of the method known as convolution quadrature [106]. This approach is not restricted to parabolic operators to the same extent as the method (17.3), and applications to wave equations are described in [15, 16, 33]. Parallelization in time is effected by clever use of the z-transform, which is inverted numerically by applying the trapezoidal rule on a circle.

18. Special functions. Integrals evaluated by variants of the trapezoidal rule are one of the mainstays of special function evaluation. Such applications have been illustrated already by Poisson's example in §1 (elliptic functions), the Bessel function example of §5, and the gamma function of §16.

TABLE 18.1

Examples of special functions computable by the trapezoidal rule. The symbols z and x denote formulas that are valid for complex and real arguments, respectively. An asterisk in the second column indicates that some contour deformation/transformation in the classical integral formula is applied to map to the real line. The special functions of [109] *were also discussed earlier in* [108].

Function	Type	References
Airy, $\mathrm{Ai}(z)$	real line*	[58]
Bessel, $J_n(z)$, $I_n(z)$	periodic interval	[44, 109, 147, 159]
Bessel, $J_\nu(z), Y_\nu(z), K_\nu(z)$	real line*	[44, 78, 109, 147]
Complete elliptic integrals, $E(k)$, $F(k)$	periodic interval	[109]
Error and related, $w(z)$, $\mathrm{erf}(z)$, $\mathrm{erfc}(z)$	real line	[42, 44, 109, 119, 127]
Gamma, $\Gamma(z)$	real line*	[143]
Incomplete gamma, $\Gamma(a, z)$	real line*	[6]
Mittag-Leffler, $E_\alpha(x)$	real line*	[54]
Parabolic cylinder, $U(a, z)$	real line*	[59]
Psi, $\Psi(a, b; x)$	real line*	[7]
Zeta, $\zeta(z)$	real line*	[8, 147, 172]

Some of the early attempts in the computer era at table generation were directly based on this strategy. An example is the complex error function, also known as the plasma dispersion or Faddeeva function, $w(z)$ [42, 166]. For z in the upper half-plane, the function can be expressed as the Hilbert transform of the Gaussian,

$$w(z) = \frac{i}{\pi} \int_{-\infty}^{\infty} \frac{e^{-t^2}}{z - t}\, dt, \tag{18.1}$$

with the integral to be interpreted in the principal value sense when z is real. It can also be written in terms of the complementary error function,

$$w(z) = e^{-z^2}\mathrm{erfc}(-iz) = e^{-z^2}\Big(1 + \frac{2i}{\sqrt{\pi}} \int_0^z e^{t^2}\, dt\Big).$$

Several other functions can be related to it, such as the Fresnel and Voigt functions, as well as Dawson's integral. In 1954, Faddeeva and Terent'ev compiled a set of tables for $w(z)$ in the rectangle $z = x + iy$, $0 \le x \le 5$, $1 \le y \le 5$ [42]. (For $0 \le y \le 1$ they used Taylor series, which are more effective, and in the other quadrants symmetry formulas can be applied.) By bounding the error with the Euler–Maclaurin formula, they were able to generate values to an accuracy of about 10^{-8}, an impressive feat at the time. The success of the trapezoidal rule when applied to (18.1) is partially due to the Gaussian decay of the integrand (just as in the example of §5). The fact that the integrand in (18.1) has a pole at z, however, limits the width of the strip of analyticity and hence the convergence rate. This is why, in [42], the trapezoidal rule was used only for $y \ge 1$. Later authors explicitly removed the pole by a residue subtraction [4, 28, 79, 119, 127]. One must be mindful, however, of the possibility of floating point cancellation errors in the resulting formula [165].

When the decay of the integrand is not as rapid as Gaussian, a useful strategy to force a more rapid decay is contour deformation. Methods of this kind based on general Hankel contours were discussed at length in §16. Another contour deformation strategy is to integrate along a path of steepest descent or stationary phase in order to eliminate oscillations.

A summary of special functions for which the trapezoidal rule has proved effective is presented in Table 18.1.

19. Functions and eigenvalues of matrices and operators. One of the most powerful applications of the trapezoidal rule in scientific computing involves the computation of functions $f(A)$, where A is a matrix or linear operator and f is a function such as $\exp(z)$, $\log(z)$, or $z^{1/2}$. For a comprehensive presentation of this problem, including precise definitions of $f(A)$, see Higham's book *Functions of Matrices* [75]. A related application concerns the computation of eigenvalues.

The power of these techniques stems from a conjunction of two circumstances (already relevant in §17, though not discussed explicitly there). On one hand is the mathematical fact that a special kind of Cauchy integral, sometimes known as the Dunford–Taylor integral, provides a representation for general functions. If A is a matrix or a bounded operator in a Banach space and f is analytic in a region of the complex plane that encloses the spectrum of A, then

$$f(A) = \frac{1}{2\pi i} \int_\Gamma (zI - A)^{-1} f(z) dz, \tag{19.1}$$

where Γ is any contour in the complex plane that encloses the spectrum while lying within the region of analyticity [89]. This equation is a natural candidate for quadrature by the trapezoidal rule, which approximates (19.1) by matrices or operators of the form

$$f_N(A) = \sum_{j=1}^{N} c_j (z_j I - A)^{-1} \tag{19.2}$$

for appropriate weights c_j (depending on f). In simple cases Γ may be a circle, and in other situations it may be a Hankel contour as considered in §§16 and 17, which can be regarded as devoted to the important special case $f(z) = e^z$. (Here one may employ a generalization of (19.1) to unbounded operators.) Note that whereas the integrand of (13.8) has a single pole inside the contour of integration, that of (19.1) will have as many as n poles if A is an $n \times n$ matrix, namely the eigenvalues of A.

The other circumstance that makes (19.2) so powerful is a fact of numerical linear algebra. Suppose A is a matrix and we wish to apply $f_N(A)$ to a particular vector $\boldsymbol{v}$ of the appropriate dimension,

$$f_N(A)\boldsymbol{v} = \sum_{j=1}^{N} c_j (z_j I - A)^{-1} \boldsymbol{v}. \tag{19.3}$$

Then the jth summand of (19.3) is the vector $\boldsymbol{x}_j$ that is the solution of a linear system of equations:

$$f_N(A)\boldsymbol{v} = \sum_{j=1}^{N} c_j \boldsymbol{x}_j, \quad (z_j I - A)\boldsymbol{x}_j = \boldsymbol{v}. \tag{19.4}$$

Now the solution of linear systems of equations is the most highly developed task in all of computational science. Methods exist that can solve $(z_j I - A)\boldsymbol{x}_j = \boldsymbol{v}$ for $\boldsymbol{x}_j$ even when A has dimension in the millions, as will often be the case for matrices arising in the discretization of partial differential equations. Thus the magic of (19.4) is that it reduces the matrix function problem for the general function f to the case of the matrix function we know almost everything about, z^{-1}. An additional bonus is that the algorithm is easily parallelizable on a computer, as mentioned in §17,

since each of the summands may be computed independently. A drawback is that complex arithmetic is usually required, which may be more costly than real arithmetic. However, in many cases one may recover a factor of 2 by exploiting the symmetry between two solves involving z_j and $\overline{z}_j$.

Here is a prototypical theorem on the geometrical convergence of these methods, a generalization of Theorem 13.2. We shall not give a proof.

THEOREM 19.1. *Let Γ be the circle $|z - z_0| = r$ in the complex plane for some z_0 and $r > 0$, let A be a matrix or linear operator whose spectrum lies within the disk $|z - z_0| < \rho$ for some $\rho < r$, and let f be a function analytic in the disk $|z - z_0| \le R$ for some $R > r$. Then in any matrix norm,*

$$\|f(A) - f_N(A)\| = O(\max\{(r/R)^N, (\rho/r)^N\}) \tag{19.5}$$

as $N \to \infty$.

For many years, as the subject of numerical linear algebra was developing, numerical methods derived from (19.1) did not get much attention. In fact, the first three editions of the landmark text by Golub and Van Loan described this formula as "fairly useless from the computational point of view." In the past decade, however, contour integrals have come to be recognized as practical methods for computing matrix functions, and the comment was dropped from the 2013 fourth edition [62].

Sometimes, the trapezoidal rule can be applied directly on a circle as described above. For example, this method is used for computing phi functions in [88, 144].

In other applications, the trapezoidal rule becomes competitive after a change of variables corresponding to a conformal map. This idea was put forward in [69]. In the problem considered in that paper, A is a matrix whose spectrum lies in an interval $[m, M]$ of the positive real axis, and f is a function that is analytic in the complex plane minus the negative real axis $(-\infty, 0]$. This configuration covers the important cases of $\log(A)$, $A^{1/2}$, or more generally A^α for a symmetric positive definite matrix A, whose condition number will be at most $\kappa = M/m$. Now suppose one applies the trapezoidal rule directly. To enclose the spectrum while staying away from the line of singularities of f, the circle of integration will have to pass through the gap between $z = 0$ and $z = m$, as in Figure 19.1. Theorem 19.1 indicates that the trapezoidal rule will require $N = O(\kappa)$ quadrature points per digit of accuracy, an intolerably slow rate of convergence for ill-conditioned matrices.

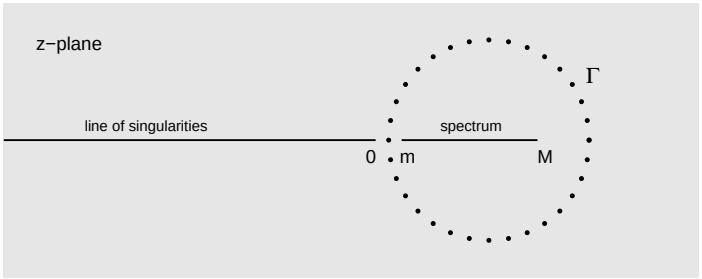


FIG. 19.1. *One way to apply the trapezoidal rule for evaluating $f(A)$, where A is symmetric positive definite with spectrum in $[m, M]$ and f is analytic except on the negative real axis (figure from [69]). According to Theorem 19.1, the convergence rate will be just $O((1 - \kappa^{-1})^N)$, which is intolerably slow if $\kappa \gg 1$.*

A conformal map, however, can improve this figure drastically to $N = O(\log(\kappa))$ quadrature points per digit, as shown in Figure 19.2. Let the whole complex plane,

minus the negative real axis and the interval $[m, M]$, be conformally mapped onto an annulus in the s-plane, with $[m, M]$ mapping to the inner boundary and $(-\infty, 0]$ to the outer one, and let the trapezoidal rule be applied on a circle in the annulus. (The details of the conformal map, which involves a Jacobi elliptic function, are spelled out in [69].) Now the convergence is quite satisfactory. Experiments in [69] show 10-digit accuracy achieved for $A^{1/2}$ or $\log(A)$ for a matrix with condition number 10^4.

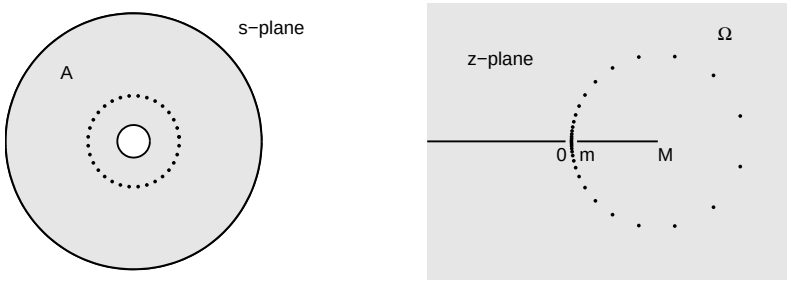


FIG. 19.2. *By a conformal map to an annulus, the rate of convergence can be improved greatly to* $O((1 - (\log(\kappa))^{-1})^N)$ *(figure from* [69]*).*

The technique just outlined is called "Method 1" in [69], and applies for functions f with arbitrary singularities on $(-\infty, 0]$. In the common situation where $(-\infty, 0]$ is just a branch cut of f, with no singularity except at $z = 0$, a modified conformal map can be used that leads to a "Method 2" converging twice as fast. For the particular case $f(z) = z^{1/2}$, a further "Method 3" simplification reduces the whole problem to one of real arithmetic, and an example is given in [69] where it takes 1 second to compute the square root of a matrix of dimension 4096 on a computer for which the Matlab `sqrt` command, employing a Schur decomposition of the matrix, requires 26 minutes. In this particular application, it is pointed out in [69] that the equispaced nodes of the trapezoidal rule, which we have seen in §14 can be regarded as poles of a meromorphic approximation, are related to the equispaced poles of a Jacobi elliptic function utilized by Zolotarev in 1877 to derive best rational approximations [189].

Methods of this kind are applied to the solution of fractional-in-space reaction-diffusion equations in [26], including the fractional Fisher and Allen–Cahn equations.

Other problems of functions of matrices and operators benefit from other conformal maps. An example of interest to physicists concerns the Fermi–Dirac function of electronic structure analysis. Here the function $\tanh(A)$ arises, and Lin, et al. have presented remarkable results based on conformal maps and the trapezoidal rule [103]. Different conformal maps are needed for the analysis of insulators, whose Hamiltonians have a gap in the spectrum, and conductors, where there is no such gap.

Up to now, we have spoken of functions of matrices that couple all the eigencomponents, which correspond mathematically to an integral (19.1) in which the contour Γ encloses all the eigenvalues of A. Another door is opened when one considers the same integral with Γ enclosing just some of the eigenvalues, which leads to a projection of $f(A)$ onto the eigenspace associated with those eigenvalues. In the simplest case $f(z) = I$, (19.1) becomes

$$P = \frac{1}{2\pi i} \int_\Gamma (zI - A)^{-1} dz. \tag{19.6}$$

If A is an $n \times n$ matrix regarded as a linear operator on $\mathbb{C}^n$, then P is precisely the projection of $\mathbb{C}^n$ onto the eigenspace just mentioned.

This idea is a powerful one in computational science, where some components of a large matrix or linear operator are often the dominant ones for the application at hand. In the simplest application of this kind of thinking, suppose one wants to find the eigenvalues and eigenvectors enclosed by Γ of a large matrix A, which we suppose are known to be k in number. If P is applied to k randomly chosen vectors in $\mathbb{C}^n$, then with probability 1, the resulting k vectors will exactly span the eigenspace in question, and from that point, eigenvalues and eigenvectors can be determined by standard methods of numerical linear algebra (see [11]). This idea has been proposed as the basis for numerical algorithms by Sakurai and coauthors (see [80, 142] and other publications in the Japanese literature) and Polizzi, whose FEAST algorithm has applied such ideas to large-scale applications in physics [138]. In practical algorithms, P is not applied exactly but is approximated by a quadrature formula, and an outer iteration is introduced to improve the approximation. An obvious choice for some applications is the trapezoidal rule in a circle, though FEAST itself is based on Gauss quadrature. Nonlinear analogues have been considered [10, 20, 187], and spectral projectors are analyzed from other angles in [17]. We expect that algorithms of this kind will be applied to many problems in the years ahead.

20. Integral equations. Together with PDEs, integral equations are among the core techniques of mathematical physics. Discretizations related to the trapezoidal rule have been used by many authors over the years, including [4, 25, 41, 71, 91, 97]. We shall not attempt a thorough review of this large subject, but restrict our discussion to an example and a few comments.

The starting point is the idea of an integral equation of the *first kind,*

$$\int K(x,y)f(y)dy = g(x), \tag{20.1}$$

or the *second kind,*

$$f(x) + \int K(x,y)f(y)dy = g(x). \tag{20.2}$$

Here K is a given kernel function and g is a given forcing function; the problem is to find the unknown function $f(x)$. The integral is over a prescribed domain in the y variable, and when it comes to computation, numerical methods for evaluating the integral are at the heart of the matter. Often the domain of integration is equivalent to a circle or some other smooth closed curve, and consequently the trapezoidal rule makes a regular appearance in the computational integral equations literature.

In a simple quadrature problem, one has an integrand to be integrated approximately by a quadrature formula. For an integral equation, however, the integrand is unknown and one must find it approximately by solving a system of equations. The basic idea of deriving systems of equations from quadrature formulas goes by the name of the *Nyström method* [31, 99, 131].

As an example, consider the solution of Laplace's equation inside an ellipse parameterized by $x = 2\cos(\theta)$, $y = \sin(\theta)$, $0 \le \theta \le 2\pi$, and taking boundary values $\sqrt{2+y}$ on its perimeter. The solution at the origin, say, is given by

$$u(0,0) = \int_0^{2\pi} \frac{\mu(\phi)}{2\cos^2(\phi) + \frac{1}{2}\sin^2(\phi)}\, d\phi, \tag{20.3}$$

where $\mu(\theta)$ is described by the Fredholm integral equation of the second kind,

$$\pi\mu(\theta) + 2\int_0^{2\pi} \frac{\mu(\phi)}{5 - 3\cos(\theta+\phi)}\, d\phi = \sqrt{2+\sin(\theta)}, \quad 0 \le \theta \le 2\pi; \tag{20.4}$$

see [30, Chapter VI].[20]

If one approximates the integral in (20.4) by the trapezoidal rule with spacing $h = 2\pi/N$, a linear system is obtained from which the function values $\mu(kh)$, $k = 1, \ldots, N$, can be determined. The system has coefficient matrix $A = \pi I + 2hB$ with $b_{kj} = 1/(5 - 3\cos((k+j)h))$, and the right-hand side has entries $\sqrt{2+\sin(kh)}$. The matrix A has dimension $N \times N$, but the symmetry about the y-axis can be exploited to reduce it to a matrix of half the size. Once the $\mu(kh)$ are known, the value of $u(0,0)$ can be computed from (20.3), again with the trapezoidal rule. The results show the lightning convergence we have seen over and over in this paper:

h	$u(0,0)$
$2\pi/8$	**1**.41
$2\pi/16$	**1.376**49
$2\pi/32$	**1.37609**30
$2\pi/64$	**1.3760929841944823**

Assuming a convergence rate of $O(e^{-2\pi a/h})$, an empirical fit of the data suggests $a \approx 0.55$. This is consistent with the singularities of the denominator in (20.3), which are located at $\mathrm{Im}\,(\phi) = \pm\mathrm{arcsinh}(1/\sqrt{3}) \approx \pm 0.549$. With a different right-hand side in (20.4), the convergence rate may be determined not by the denominator in (20.3) but rather by the singularities of the numerator, $\mu(\phi)$.

Although this example is a simple one, the same ideas of Nyström's method combined with the trapezoidal rule feature in advanced applications as well. For example, an application in 3D scattering is described in [25]. It is impossible to do justice here to the vast field of integral equations, which encompasses topics including singular integral equations, boundary integral equations, equations of convolution type, and many others. We already noted the concept of convolution quadrature in §17, and for its application to integral equations of Abel–Volterra, Wiener–Hopf, and other types, we refer to the many papers by Lubich and co-workers ([106] reviews many of these developments and contains a large number of references). Likewise, we noted the Hilbert transform at the end of §16, which arises in the the study of airfoils and potential theory and many other applications.

21. Afterword. It would be hard to argue that the fast-converging trapezoidal rule is ever indispensable. It is only a quadrature formula, and there are others. Gauss quadrature is always an option, for example, and now that algorithms with $O(N)$ complexity are available in Chebfun and elsewhere that make it easy to compute Gauss nodes and weights even for millions of points [21, 70], the traditional avoidance of Gauss quadrature for problems with large N has lost its justification. Clenshaw–Curtis quadrature has much the same behavior as Gauss quadrature in many circumstances, and is easy to implement [168, 169]. Indeed, Clenshaw–Curtis

[20] This example fills the second author with no small amount of nostalgia. It was while solving this problem as part of a BSc Honours project in 1979 that he discovered the remarkable speed of convergence of the trapezoidal rule for himself.

quadrature is equivalent to the trapezoidal rule combined with the Joukowski transformation. Finally, for many problems where an algorithm based on quadrature is effective, there may be equally good or better alternatives based on series, best approximations, or other ideas. For example, the trapezoidal rule method used to compute the gamma function for Figure 16.1 is not the best possible algorithm for that problem, and alternatives are discussed in [144].

One reason why the fast-converging trapezoidal rule is nevertheless so valuable is that in some applications it is clearly the most natural tool to use. This is especially apparent for problems involving integration over a periodic interval or a circle. If $v(\theta)$ is periodic on $[0, 2\pi]$, one could ignore the periodicity and apply Gauss or Clenshaw–Curtis quadrature anyway, at the cost of just a constant factor approximating $\pi/2$. But this would seem a very odd approach, and in some applications, a factor of $\pi/2$ may matter a great deal. On the real line or a Hankel contour, on the other hand, the conceptual or practical advantages of the trapezoidal rule are more nuanced.

The most obvious advantage of the trapezoidal rule is its great simplicity, which has been at the heart of its appeal all along. One aspect of its simplicity is its ready accessibility when a user wants to figure out how to get the job done. Another is more mathematical. Since the trapezoidal rule converges at a geometric rate associated with a disk, annulus, half-plane or strip of analyticity, its use almost automatically brings along a known convergence rate for many problems. This easily understood behavior is what makes it so straightforward to determine what step size and truncation point to use in a sum like (6.4), for example, to achieve $O(\exp(-CN^{1/2}))$ convergence. Other methods may be able to achieve the same convergence rate, but rarely with such easy analysis. Relatedly, the mathematical simplicity of the trapezoidal rule means that in theoretical applications it may provide explicit formulas. For example, in §14 we saw how it could be used to derive the explicit formulas (14.3) and (14.4) for nearly-optimal rational approximations to $|\xi|$ on $[-1, 1]$ and e^t on $(-\infty, 0]$.

Finally, an easy tool in widespread use can reveal connections between topics. In this article we have described connections between rational approximations and inverse Laplace transforms, between Cauchy integrals and matrix eigenvalue software, between Faraday cages and the nonuniform Fast Fourier Transform. The trapezoidal rule links different corners of applied mathematics.

Acknowledgments. We first began writing this article in 2005, and many people have given us good advice along the way, including responses to NA Digest postings in 2006 and 2012. We are glad to acknowledge the help of Steven Balbus, Lehel Banjai, Alex Barnett, Sir Michael Berry, Timo Betcke, Folkmar Bornemann, Coralia Cartis, Simon Chandler-Wilde, Maurice Cox, David Elliott, Bengt Fornberg, Walter Gautschi, Anne Gelb, Leslie Greengard, Stefan Güttel, Nick Hale, Dave Hewett, Herbert Homeier, Steven Johnson, Rogelio Luck, James Lyness, Paul Matthews, Yuji Nakatsukasa, Dirk Nuyens, Frank Olver, Daniel Potts, Minvydas Ragulskis, Fernando Reitich, Mark Richardson, Edward Saff, Charlie Schwartz, Ridgway Scott, Avram Sidi, Alastair Spence, Frank Stenger, Gilbert Strang, Jared Tanner, Garry Tee, Nico Temme, Laurette Tuckerman, Gert Van den Eynde, Jörg Waldvogel, and Andrew Zangwill. Georges Klein made substantial contributions to §10. Endre Süli helped us translate [87] from Hungarian and [90] from Russian.

Our greatest debt is to Oxford D.Phil. students Anthony Austin and Mohsin Javed. During our final period of work on this project, in May–June 2013, these two met with us every morning to discuss the paper in detail, section by section from beginning to end. Their suggestions led to innumerable improvements, and we will

look back happily on the memory of these daily trapezoidal meetings.

REFERENCES

[1] B. Adcock and M. Richardson, *New exponential variable transform methods for functions with endpoint singularities,* manuscript, 2013.
[2] L. V. Ahlfors, *Complex Analysis,* 3rd ed., McGraw-Hill, 1978.
[3] A. C. Aitken, *Statistical Mathematics,* Oliver and Boyd, Edinburgh and London, 1939.
[4] M. Alazah, S. N. Chandler-Wilde, and S. La Porte, *Computing Frensel integrals via modified trapezium rules,* Numer. Math., to appear.
[5] A. Aldroubi and K. Gröchenig, *Nonuniform sampling and reconstruction in shift-invariant spaces,* SIAM Rev., 43 (2001), pp. 585–620.
[6] G. Allasia and R. Besenghi, *Numerical calculation of incomplete gamma functions by the trapezoidal rule,* Numer. Math., 50 (1987), pp. 419–428.
[7] G. Allasia and R. Besenghi, *Numerical computation of Tricomi's psi function by the trapezoidal rule,* Computing, 39 (1987), pp. 271–279.
[8] G. Allasia and R. Besenghi, *Numerical calculation of the Riemann zeta function and generalizations by means of the trapezoidal rule,* in Numerical and Applied Mathematics, Part II, Baltzer, Paris, 1989.
[9] E. G. Anastasselou, *A formal comparison of the Delves–Lyness and Burniston–Siewert methods for locating the zeros of analytic functions,* IMA J. Numer. Anal., 6 (1986), pp. 337–341.
[10] J. Asakura, T. Sakurai, H. Tadano, T. Ikegami, and K. Kimura, *A numerical method for nonlinear eigenvalue problems using contour integrals,* JSIAM Letters, 1 (2009), pp. 52–55.
[11] A. P. Austin, P. Kravanja, and L. N. Trefethen, *Numerical algorithms based on analytic function values at roots of unity,* SIAM J. Numer. Anal., submitted, 2013.
[12] D. H. Bailey and J. M. Borwein, *High-precision numerical integration: Progress and challenges,* J. Symbolic Comput., 46 (2011), pp. 741–754.
[13] D. H. Bailey, K. Jeyabalan, and X. S. Li, *A comparison of three high-precision quadrature schemes,* Experiment. Math., 14 (2005), pp. 317–329.
[14] D. H. Bailey and X. S. Li, *A comparison of three high-precision quadrature schemes,* Proceedings of the Real Numbers and Computing Conference, Lyon, Sept., 2003.
[15] L. Banjai and D. Peterseim, *Parallel multistep methods for linear evolution problems,* IMA J. Numer. Anal., 32 (2012), pp. 1217–1240.
[16] L. Banjai and S. Sauter, *Rapid solution of the wave equation in unbounded domains,* SIAM J. Numer. Anal., 47 (2008), pp. 227–249.
[17] M. Benzi, P. Boito, and N. Razouk, *Decay properties of spectral projectors with applications to electronic structure,* SIAM Rev., 55 (2013), pp. 3–64.
[18] J.-P. Berrut, *A circular interpretation of the Euler–MacLaurin formula,* J. Comp. Appl. Math., 189 (2006), pp. 375–386.
[19] J.-P. Berrut and L. N. Trefethen, *Barycentric Lagrange interpolation,* SIAM Rev., 46 (2004), pp. 501–517.
[20] W.-J. Beyn, *An integral method for solving nonlinear eigenvalue problems,* Lin. Alg. Applics., 436 (2012), pp. 3839–3863.
[21] I. Bogaert, B. Michiels, and J. Fostier, *$O(1)$ computation of Legendre polynomials and Gauss–Legendre nodes and weights for parallel computing,* SIAM J. Sci. Comput., 34 (2012), C83–C101.
[22] F. Bornemann, *Accuracy and stability of computing high-order derivatives of analytic functions by Cauchy integrals,* Found. Comput. Math., 11 (2011), pp. 1–63.
[23] F. Bornemann, D. Laurie, S. Wagon, and J. Waldvogel, *The SIAM* 100*-Digit Challenge: A Study in High-Accuracy Numerical Computing,* SIAM, 2004.
[24] H. Brass and K. Petras, *Quadrature Theory: The Theory of Numerical Integration on a Compact Interval,* Amer. Math. Soc, Providence, RI, 2011.
[25] O. P. Bruno, V. Dominguez, and F.-J. Sayas, *Convergence analysis of a high-order Nystrom integral-equation method for surface scattering problems,* Numer. Math., 124 (2013), pp. 603–645.
[26] K. Burrage, N. Hale, and D. Kay, *An efficient implicit FEM scheme for fractional-in-space reaction-diffusion equations,* SIAM J. Sci. Comp., 34 (2012), pp. A2145–A2172.
[27] J. C. Butcher, *On the numerical inversion of Laplace and Mellin transforms,* Conference on Data Processing and Automatic Computing Machines, Salisbury, Australia, 1957.
[28] C. Chiarella and A. Reichel, *On the evaluation of integrals related to the error function,* Math. Comp., 22 (1968), pp. 137–143.

[29] W. J. Cody, G. Meinardus, and R. S. Varga, *Chebyshev rational approximations to e^{-x} in $[0,+\infty)$ and applications to heat-conduction problems,* J. Approx. Th., 2 (1969), pp. 50–65.
[30] L. Collatz, *The Numerical Treatment of Differential Equations,* 3rd ed., Springer-Verlag, Berlin, 1960.
[31] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory,* Springer-Verlag, 1998.
[32] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, *On the Lambert W function,* Adv. Comput. Math., 5 (1996), pp. 329–359.
[33] E. Cuesta, C. Lubich, and C. Palencia, *Convolution quadrature time discretization of fractional diffusion-wave equations,* Math. Comp., 75 (2006), pp. 673–696.
[34] J. H. Curtiss, *A stochastic treatment of some classical interpolation problems,* Proc. 4th Berkeley Sympos. Math. Statist. and Prob., v. II, 1961, pp. 79–93.
[35] P. J. Davis, *On the numerical integration of periodic analytic functions,* in R. E. Langer, ed., On Numerical Integration: Proceedings of a Symposium, Madison, April 21–23, 1958, Math. Res. Ctr., U. of Wisconsin, 1959, pp. 45–59.
[36] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration,* Academic Press, New York, 1975.
[37] L. M. Delves and J. N. Lyness, *A numerical method for locating the zeros of an analytic function,* Math. Comp., 21 (1967), pp. 543–560.
[38] B. Dingfelder and J. A. C. Weideman, *An improved Talbot method for numerical Laplace transform inversion,* arXiv:1304.2505
[39] A. Dutt and V. Rokhlin, *Fast Fourier transforms for nonequispaced data,* SIAM J. Sci. Comput., 14 (1993), pp. 1368–1383.
[40] D. Elliott, *The evaluation and estimation of the coefficients in the Chebyshev series expansion of a function,* Math. Comp., 18 (1964), pp. 274–284.
[41] D. Elliott, *Sigmoidal-trapezoidal quadrature for ordinary and Cauchy principal value integrals,* ANZIAM J., 46 (2004), pp. E1–E69.
[42] V. N. Faddeyeva and N. M. Terentev, *Tables of values of the function $w(z) = e^{-z^2}(1 + 2i\pi^{-1/2}\int_0^z e^{t^2}dt)$ for complex argument,* in V. A. Fok, ed., Mathematical Tables Series, v. 11, Pergamon Press, Oxford, 1961.
[43] L. Fejér, *Interpolation und konforme Abbildung,* Göttinger Nachrichten, 1918, pp. 319–331.
[44] H. E. Fettis, *Numerical calculation of certain definite integrals by Poisson's summation formula,* Math. Comp., 9 (1955), pp. 85–92.
[45] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics,* v. 2, Addison-Wesley, 1964.
[46] N. Flyer and A. S. Fokas, *A hybrid analytical-numerical method for solving evolution partial differential equations. I. The half-line,* Proc. R. Soc. A, 464 (2008), pp. 1823–1849.
[47] A. S. Fokas, *A Unified Approach to Boundary Value Problems,* SIAM, 2008.
[48] A. S. Fokas, N. Flyer, S. A. Smithemana, and E. A. Spence, *A semi-analytical numerical method for solving evolution and elliptic partial differential equations,* J. Comp. Appl. Math., 227 (2009), pp. 59–74.
[49] B. Fornberg, *Numerical differentiation of analytic functions,* ACM Trans. Math. Softw., 7 (1981), pp. 512–526.
[50] B. Fornberg, *ALGORITHM 579: CPSC: Complex power series coefficients,* ACM Trans. Math. Softw., 7 (1981), pp. 542–547.
[51] B. Fornberg and G. Wright, *Stable computation of multiquadric interpolants for all values of the shape parameter,* Computers and Math. w. Applics., 48 (2004), pp. 853–867.
[52] B. G. Gabdulhaev, *Approximate solution of singualr integral equations by the method of mechanical quadratures,* Dokl. Akad. Nauk SSSR, 179 (1968), pp. 260–263.
[53] D. Gaier, *Lectures on Complex Approximation,* Birkhäuser, Boston, 1987.
[54] R. Garrappa and M. Popolizio, *Evaluation of generalized Mittag-Leffler functions on the real line,* Adv. Comp. Math., 2012.
[55] I. P. Gavrilyuk, W. Hackbusch, and B. N. Khoromskij, *$\mathcal{H}$-matrix approximation for the operator exponential with applications,* Numer. Math., 92 (2002), pp. 83–111.
[56] I. P. Gavrilyuk and V. L. Makarov, *Exponentially convergent parallel discretization methods for the first order evolution equations,* Comput. Methods Appl. Math., 1 (2001), pp. 333–355.
[57] I. Gavrilyuk, V. Makarov, and V. Vasylyk, *Exponentianlly Convergent Algorithms for Abstract Differential Equations,* Birkhäuser, 2011.
[58] A. Gil, J. Segura, and N. M. Temme, *Computing complex Airy functions by numerical quadrature,* Numer. Algorithms, 30 (2002), pp. 11–23.
[59] A. Gil, J. Segura, and N. M. Temme, *Computing special functions by using quadrature rules,* Numer. Algorithms, 33 (2003), pp. 265–275.

[60] A. Gil, J. Segura, and N. M. Temme, *Numerical Methods for Special Functions,* SIAM, 2007.
[61] S. Goedecker, *Linear scaling electronic structure methods,* Rev. Modern Phys., 71 (1999), pp. 1085–1123.
[62] G. H. Golub and C. F. Van Loan, *Matrix Computations,* 4th ed., Johns Hopkins University Press, 2013.
[63] A. A. Gonchar and E. A. Rakhmanov, *Equilibrium distributions and degree of rational approximation of analytic functions,* Math. USSR Sb., 62 (1989), 305–348.
[64] P. Gonnet, S. Güttel, and L. N. Trefethen, *Robust Padé approximation via SVD,* SIAM Rev., 55 (2013), pp. 101–117.
[65] E. T. Goodwin, *The evaluation of integrals of the form* $\int_{-\infty}^{\infty} f(x)e^{-x^2}dx$ *,* Proc. Cambr. Philos. Soc., 45 (1949), pp. 241–245.
[66] U. Graf, *Introduction to Hyperfunctions and Their Integral Transforms: An Applied and Computational Approach,* Birkhäuser, 2000.
[67] J. S. Green, *The Calculation of the Time-Responses of Linear Systems,* Ph.D. thesis, U. of London, 1955.
[68] S.-Å. Gustafson, *Quadrature rules derived from linear convergence acceleration schemes,* in Numerical Integration. IV, Birkhäuser, Basel, 1993, pp. 151–165.
[69] N. Hale, N. J. Higham, and L. N. Trefethen, *Computing* A^α*,* $\log(A)$*, and related matrix functions by contour integrals,* SIAM J. Numer. Anal., 46 (2008), pp. 2505–2523.
[70] N. Hale and A. Townsend, *Fast and accurate computation of Gauss–Legendre and Gauss-Jacobi quadrature nodes and weights,* SIAM J. Sci. Comp., 35 (2013), pp. A652–A674.
[71] J. Helsing and R. Ojala, *On the evaluation of layer potentials close to their sources,* J. Comp. Phys., 227 (2008), pp. 2899–2921.
[72] P. Henrici, *Applied and Computational Complex Analysis, v.* 2*: Special Functions, Integrals Transforms, Asymptotics, Continued Fractions,* Wiley, 1977.
[73] P. Henrici, *Fast Fourier methods in computational complex analysis,* SIAM Review, 21 (1979), pp. 481–527.
[74] P. Henrici, *Applied and Computational Complex Analysis, v.* 3*: Discrete Fourier Analysis, Cauchy Integrals, Construction of Conformal Maps, Univalent Functions,* Wiley, 1986.
[75] N. J. Higham, *Functions of Matrices: Theory and Computation,* SIAM, 2008.
[76] E. Hlawka, *Interpolation analytischer Funktionen auf dem Einheitskreis,* in P. Turán, ed., Number Theory and Analysis, Plenum, New York, 1969, pp. 99–118.
[77] M. Hochbruck and A. Ostermann, *Exponential integrators,* Acta Numer., 19 (2010), pp. 209–286.
[78] D. B. Hunter, *The calculation of certain Bessel functions,* Math. Comp., 18 (1964), pp. 123–128.
[79] D. B. Hunter and T. Regan, *A note on the evaluation of the complementary error function,* Math. Comp., 26 (1972), pp. 539–541.
[80] T. Ikegami and T. Sakurai, *Contour integral eigensolver for non-Hermitian systems: a Rayleigh–Ritz-type approach,* Taiwanese J. Math., 14 (2010), pp. 825–837.
[81] K. J. in 't Hout and J. A. C. Weideman, *A contour integral method for the Black–Scholes and Heston equations,* SIAM J. Sci. Comput., 33 (2011), pp. 763–785.
[82] N. I. Ioakimidis, K. E. Papadakis and E. A. Perdios, *Numerical evaluation of analytic functions by Cauchy's theorem,* BIT, 31 (1991), pp. 276–285.
[83] M. Iri, S. Moriguti, and Y. Takasawa, *On a certain quadrature formula,* J. Comput. Appl. Math., 17 (1987), pp. 3–20 (reprinted translation of 1970 paper in Japanese).
[84] E. Isaacson and H. B. Keller, *Analysis of Numerical Methods,* Wiley, 1966.
[85] D. Jackson, *Roots and singular points of analytic functions,* Annals Math., 19 (1917), pp. 142–151.
[86] M. Javed and L. N. Trefethen, *A trapezoid rule error bound unifying the Euler–Maclaurin formula and geometric convergence for periodic functions,* in preparation.
[87] L. Kalmár, *On interpolation,* Mat. Lapok, 33 (1926), pp. 120–149 (Hungarian).
[88] A.-K. Kassam and L. N. Trefethen, *Fourth-order time-stepping for stiff PDEs,* SIAM J. Sci. Comp., 26 (2005), pp. 1214–1233.
[89] T. Kato, *Perturbation Theory for Linear Operators,* Springer-Verlag, Berlin, corrected 2nd ed., 1980.
[90] O. Kis, *On the convergence of the trigonometrical and harmonical interpolation,* Hung. Acta Math., 7 (1956), pp. 173–200 (Russian).
[91] A. Klöckner, A. Barnett, L. Greengard, and M. O'Neil, *Quadrature by expansion: a new method for the evaluation of layer potentials,* arXiv:1207.4461v1, 2012.
[92] N. M. Korobov, *Number-Theoretic Methods in Approximate Analysis,* Fizmatgiz, Moscow, 1963.

[93] J. Korevaar and J. L. H. Meyers, *Spherical Faraday cage for the case of equal point charges and Chebyshev-type quadrature on the sphere,* Integral Trans. and Spec. Func., 1 (1993), pp. 105–117.
[94] P. Kravanja and M. Van Barel, *Computing the Zeros of Analytic Functions,* Springer Lect. Notes Math. v. 1727, 2000.
[95] H.-O. Kreiss and J. Oliger, *Comparison of accurate methods for the integration of hyperbolic equations,* Tellus, 24 (1972), pp. 199–215.
[96] R. Kress, *Interpolation auf einem unendlichen Intervall,* Computing, 6 (1970), pp. 274–288.
[97] R. Kress, *Zur numerischen Integration periodischer Funktionen nach der Rechteckregel,* Numer. Math., 20 (1972), pp. 87–92.
[98] R. Kress, *Numerical Analysis,* Springer, 1998.
[99] R. Kress, *Linear Integral Equations,* Springer, 1989.
[100] R. Kress and E. Martensen, *Anwendung der Rechteckregel auf die reelle Hilberttransformation mit unendlichem Intervall,* Z. Angew. Math. Mech, 50 (1970), pp. 61–64.
[101] D. P. Laurie, *Periodizing transformations for numerical integration,* J. Comput. Appl. Math., 66 (1996), pp. 337–344.
[102] H. Lee and D. Sheen, *Laplace transformation method for the Black–Scholes equation,* Int. J. Numer. Anal. Model., 6 (2009), pp. 642–658.
[103] L. Lin, J. Lu, L. Ying, and W. E, *Pole-based approximation of the Fermi–Dirac function,* Chin. Ann. Math. Ser. B, 30 (2009), pp. 729–742.
[104] M. López-Fernández and C. Palencia, *On the numerical inversion of the Laplace transform in certain holomorphic mappings,* Appl. Numer. Math., 51 (2004), pp. 289–303.
[105] M. López-Fernández, C. Palencia, and A. Schädle, *A spectral order method for inverting sectorial Laplace transforms,* SIAM J. Numer. Anal., 44 (2006), pp. 1332–1350.
[106] C. Lubich, *Convolution quadrature revisited,* BIT Numerical Mathematics, 44 (2004), pp. 503–514.
[107] R. Luck and J. W. Stevens, *Explicit solutions for transcendental equations,* SIAM Rev., 44 (2002), pp. 227–233.
[108] Y. L. Luke, *Simple formulas for the evaluation of some higher transcendental functions,* J. Math. and Phys., 34 (1956), pp. 298–307.
[109] Y. L. Luke, *The Special Functions and Their Approximations,* v. 1–2, Academic Press, New York, 1969.
[110] J. Lund and K. Bowers, *Sinc Methods for Quadrature and Differential Equations,* SIAM, 1992.
[111] J. N. Lyness, *Quadrature methods based on complex function values,* Math. Comp., 23 (1969), pp. 601–619.
[112] J. N. Lyness and L. M. Delves, *On numerical contour integration round a closed contour,* Math. Comp., 21 (1967), pp. 561–577.
[113] J. N. Lyness and G. Giunta, *A modification of the Weeks method for numerical inversion of the Laplace transform,* Math. Comp., 47 (1986), pp. 313–322.
[114] J. N. Lyness and C. B. Moler, *Numerical differentiation of analytic functions,* SIAM J. Numer. Anal., 4 (1967), pp. 202–210.
[115] J. N. Lyness and G. Sande, *Algorithm* 413*: ENTCAF and ENTCRE: evaluation of normalized Taylor coefficients of an analytic function,* Comm. ACM, 14 (1971), pp. 669–675
[116] F. Mainardi and R. Gorenflo, *On Mittag-Leffler-type functions in fractional evolution processes,* J. Comput. Appl. Math., 118 (2000), pp. 283–299.
[117] A. I. Markushevich, *Theory of Functions of a Complex Variable,* 2nd ed., Chelsea, 1965.
[118] E. Martensen, *Zur numerischen Auswertung uneigentlicher Integrale,* Zeit. Angew. Math. Mech., 48 (1968), pp. T83–T85.
[119] F. Matta and A. Reichel, *Uniform computation of the error function and other related functions,* Math. Comp., 25 (1971), pp. 339–344.
[120] J. E. McCune, *Exact inversion of dispersion relations,* Phys. Fluids, 9 (1966), pp. 2082–2084.
[121] W. McLean and V. Thomée, *Time discretization of an evolution equation via Laplace transforms,* IMA J. Numer. Anal., 24 (2004), pp. 439–463.
[122] J. McNamee, *Error-bounds for the evaluation of integrals by the Euler–MacLaurin formula and by Gauss-type formulae,* Math. Comp., 18 (1964), pp. 368–381.
[123] K. Meerbergen, *The solution of parametrized symmetric linear systems,* SIAM J. Matrix Anal. Appl., 24 (2003), pp. 1038–1059.
[124] L. M. Milne-Thompson, *The Calculus of Finite Differences,* Macmillan, 1933.
[125] P. A. P. Moran, *Approximate relations between series and integrals,* Math. Comp., 12 (1958), pp. 34–37.
[126] M. Mori, *An IMT-type double exponential formula for numerical integration,* Publ. Res. Inst.

Math. Sci., 14 (1978), pp. 713–729.
[127] M. Mori, *A method for evaluation of the error function of real and complex variable with high relative accuracy,* Publ. Res. Inst. Math. Sci., 19 (1983), pp. 1081–1094.
[128] M. Mori, *Discovery of the double exponential transformation and its developments,* Publ. Res. Inst. Math. Sci., 41 (2005), pp. 897–935.
[129] M. Mori and M. Sugihara, *The double-exponential transformation in numerical analysis,* J. Comp. Appl. Math., 127 (2001), pp. 287–296.
[130] D. J. Newman, *Rational approximation to $|x|$,* Mich. Math. J., 11 (1964), pp. 11–14.
[131] E. J. Nyström, *Über die praktische Auflösung von Integralgleichungen mit Anwendungen auf Randwertaufgaben,* Acta Math., 54 (1930), 185–204.
[132] H. O'Hara and F. J. Smith, *Error estimation in the Clenshaw–Curtis quadrature formula,* Comput. J., 11 (1968), pp. 213–219.
[133] S. A. Orszag, *Comparison of pseudospectral and spectral approximations,* Stud. Appl. Math., 51 (1972), pp. 253–259.
[134] R. E. A. C. Paley and N. Wiener, *Fourier Transforms in the Complex Domain,* Amer. Math. Soc., New York, 1934.
[135] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations,* Springer-Verlag, New York, 1983.
[136] S.-D. Poisson, *Mémoire sur le calcul numérique des intégrales définies* (abstract of [137]), Nouv. Bull. des sc. de la Societe philomathique Paris, 1826, pp. 161–162.
[137] S.-D. Poisson, *Sur le calcul numérique des intégrales définies,* Mémoires de l'Académie Royale des Sciences de l'Institut de France, 4 (1827), pp. 571–602.
[138] E. Polizzi, *Density-matrix-based algorithm for solving eigenvalue problems,* Phys. Rev. B, 79 (2009), pp. 115112:1–6.
[139] M. Richardson, *Spectral Representation of Functions with Endpoint Singularities,* D. Phil. diss., Math. Inst., U. of Oxford, 2013.
[140] C. Runge, *Theorie und Praxis der Reihen,* G. J. Göschen, Leipzig, 1904.
[141] T. W. Sag and G. Szekeres, *Numerical evaluation of high-dimensional integrals,* Math. Comp., 18 (1964), pp. 245–253.
[142] T. Sakurai and H. Sugiura, *A projection method for generalized eigenvalue problems using numerical integration,* J. Comp. Appl. Math., 159 (2003), pp. 119–128.
[143] T. Schmelzer and L. N. Trefethen, *Computing the gamma function using contour integrals and rational approximations,* SIAM J. Numer. Anal., 45 (2007), pp. 558–571.
[144] T. Schmelzer and L. N. Trefethen, *Evaluating matrix functions for exponential integrators via Carathéodory–Fejér approximation and contour integrals,* Electron. Trans. Numer. Anal., 29 (2007), pp. 1–18.
[145] C. B. Schneider, *Inversion formulas for the discretized Hilbert transform on the unit circle,* SIAM J. Numer. Anal., 35 (1998), pp. 71–77.
[146] C. Schwartz, *Numerical integration of analytic functions,* J. Comp. Phys., 4 (1969), pp. 19–29.
[147] C. Schwartz, *Numerical calculation of Bessel functions,* submitted, 2012.
[148] D. Sheen, I. H. Sloan, and V. Thomée, *A parallel method for time discretization of parabolic equations based on Laplace transformation and quadrature,* IMA J. Numer. Anal., 23 (2003), pp. 269–299.
[149] A. Sidi, *A new variable transformation for numerical integration,* Numerical Integration, IV (Oberwolfach, 1992).
[150] A. Sidi, *Extension of a class of periodizing variable transformations for numerical integration,* Math. Comp., 75 (2006), pp. 327–343.
[151] V. Simoncini and D. B. Szyld, *Recent computational developments in Krylov subspace methods for linear systems,* Numer. Lin. Alg. Appl., 14 (2007), pp. 1–59.
[152] G. Smith, *An Introduction to Classical Electromagnetic Radiation,* Cambridge U. Press, 1997.
[153] W. Squire and G. Trapp, *Using complex variables to estimate derivatives of real functions,* SIAM Rev., 40 (1998), pp. 110–112.
[154] H. Stahl, *Best uniform rational approximation of $|x|$ on $[-1,1]$,* Russian Acad. Sci. Sb. Math., 76 (1993), pp. 461–487.
[155] F. Stenger, *Numerical Methods Based on Sinc and Analytic Functions,* Springer-Verlag, New York, 1993.
[156] F. Stenger, *Explicit nearly optimal linear rational approximation with preassigned poles,* Math. Comp., 47 (1986), pp. 225–252.
[157] F. Stenger, *Sinc Numerical Methods,* CRC Press, 2010.
[158] A. H. Stroud, *A bibliography on approximate integration,* Math. Comput., 15 (1961), pp. 52–80.

[159] A. H. Stroud and J. P. Kohli, *Computation of $J_n(x)$ by numerical integration,* Commun. ACM, 12 (1969), pp. 236–238.
[160] H. Takahasi and M. Mori, *Estimation of errors in the numerical quadrature of analytic functions,* Applicable Anal., 1 (1971), pp. 201–229.
[161] H. Takahasi and M. Mori, *Quadrature formulas obtained by variable transformation,* Numer. Math., 21 (1973), pp. 206–219.
[162] H. Takahasi and M. Mori, *Double exponential formulas for numerical integration,* Publ. Res. Inst. Math. Sci., 9 (1974), pp. 721–741.
[163] A. Talbot, *The accurate numerical inversion of Laplace transforms,* J. Inst. Maths. Applics., 23 (1979), pp. 97–120.
[164] K. Tanaka, M. Sugihara, K. Murota, and M. Mori, *Function classes for double exponential integration formulas,* Numer. Math., 111 (2009), pp. 631–655.
[165] N. M. Temme, *Numerical aspects of special functions,* Acta Numer., 16 (2007), pp. 379–478.
[166] N. M. Temme, *Error functions, Dawson's and Fresnel integrals,* NIST Handbook of Mathematical Functions, U. S. Dept. Commerce, Washington, DC, 2010, pp. 159–171.
[167] L. N. Trefethen, *Spectral Methods in MATLAB,* SIAM, 2000.
[168] L. N. Trefethen, *Is Gauss quadrature better than Clenshaw–Curtis?,* SIAM Rev., 50 (2008), pp. 67–87.
[169] L. N. Trefethen, *Approximation Theory and Approximation Practice,* SIAM, 2013.
[170] L. N. Trefethen and M. H. Gutknecht, *The Carathéodory–Fejér method for real rational approximation,* SIAM J. Numer. Anal., 20 (1983), pp. 420–436.
[171] L. N. Trefethen, J. A. C. Weideman, and T. Schmelzer, *Talbot quadratures and rational approximations,* BIT Numer. Math., 46 (2006), pp. 653–670.
[172] A. M. Turing, *A method for the calculation of the zeta-function,* Proc. Lond. Math. Soc., 48 (1943), pp. 180–197 (written in 1939).
[173] M. Vetterli, J. Kovačević, and V. K. Goyal, *Foundations of Signal Processing,* Cambridge U. Press., 2013.
[174] J. Waldvogel, *Toward a general error theory of the trapezoid rule,* in W. Gautschi et al., eds., Approximation and Computation: in Honor of Gradimir V. Milovanović, Springer, 2012, pp. 267–282.
[175] H. Weber, *Numerical computation of the Fourier transform using Laguerre functions and the fast Fourier transform,* Numer. Math., 36 (1981), pp. 197–209.
[176] W. T. Weeks, *Numerical inversion of Laplace transforms using Laguerre functions,* J. Assoc. Comput. Mach., 13 (1966), pp. 419–429.
[177] J. A. C. Weideman, *Algorithms for parameter selection in the Weeks method for inverting the Laplace transform,* SIAM J. Sci. Comput., 21 (1999), pp. 111–128.
[178] J. A. C. Weideman, *Numerical integration of periodic functions: a few examples,* Amer. Math. Monthly, 109 (2002), pp. 21–36.
[179] J. A. C. Weideman, *Optimizing Talbot's contours for the inversion of the Laplace transform,* SIAM J. Numer. Anal., 44 (2006), pp. 2342–2362.
[180] J. A. C. Weideman, *Improved contour integral methods for parabolic PDEs,* IMA J. Numer. Anal., 30 (2010), pp. 334–350.
[181] J. A. C. Weideman and L. N. Trefethen, *Parabolic and hyperbolic contours for computing the Bromwich integral,* Math. Comp., 76 (2007), pp. 1341–1356.
[182] J. A. C. Weideman and L. N. Trefethen, *The kink phenomenon in Fejér and Clenshaw–Curtis quadrature,* Numer. Math., 107 (2007), pp. 707–727.
[183] H. Weyl, *Über die Gleichverteilung von Zahlen mod. Eins,* Math. Anal., 77 (1916), pp. 313–352.
[184] E. T. Whittaker, *On the functions which are represented by the expansions of the interpolation-theory,* Proc. Roy. Soc. Edinburgh, 35 (1915), pp. 181–194.
[185] J. M. Whittaker, *Interpolatory Function Theory,* Cambridge U. Press, 1935.
[186] S. Xiang and F. Bornemann, *On the convergence rates of Gauss and Clenshaw–Curtis quadrature for functions of limited regularity,* SIAM J. Numer. Anal., 50 (2012), pp. 2581–2587.
[187] S. Yokota and T. Sakurai, *A projection method for nonlinear eigenvalue problems using contour integrals,* JSIAM Letters, 5 (2013), pp. 41–44.
[188] A. Zangwill, *Modern Electrodynamics,* Cambridge U. Press, 2012.
[189] E. I. Zolotarev, *Application of elliptic functions to questions of functions deviating least and most from zero,* Izvestiya Imp. Akad. Nauk, 1877.