SIAM REVIEW
Vol. 64, No. 1, pp. 132–150

Exactness of Quadrature Formulas*

Lloyd N. Trefethen†

Abstract. The standard design principle for quadrature formulas is that they should be exact for integrands of a given class, such as polynomials of a fixed degree. We review the subject from this point of view and show that this principle fails to predict the actual behavior in four of the best-known cases: Newton–Cotes, Clenshaw–Curtis, Gauss–Legendre, and Gauss–Hermite quadrature. New results include (i) the observation that x^k is integrated accurately by the Newton–Cotes formula even though the Chebyshev polynomial $T_k(x)$ is not; (ii) the introduction of a parameter-free variant of band-limited quadrature for arbitrary integrands, which is demonstrated to have a factor $\pi/2$ advantage over Gauss quadrature in integrating complex exponentials; (iii) a theorem establishing that chopping the real line to a finite interval achieves $O(\exp(-Cn^{2/3}))$ convergence for n-point quadrature of Gauss–Hermite integrands, whereas for the Gauss–Hermite formula it is just $O(\exp(-Cn^{1/2}))$; and (iv) an explanation of how this result is consistent with the "optimality" of the Gauss–Hermite formula.

Key words. Gauss quadrature, Gauss–Hermite, Newton–Cotes, Clenshaw–Curtis, cubature

AMS subject classifications. 41A55, 65D32

DOI. 10.1137/20M1389522

1. Introduction.

Let f be a real or complex function defined on a domain D, such as an interval in one dimension or a hypercube in s dimensions. A quadrature formula is an approximation

$$I_n(f) = \sum_{j=1}^{n} w_j f(x_j) \tag{1.1}$$

to the definite integral

$$I(f) = \int_D f(x)dx. \tag{1.2}$$

Here the points $\{x_j\}$ are distinct *nodes* in D, and the numbers $\{w_j\}$ are *weights.* Sometimes a further weight function $w(x)$ is introduced in (1.2), as we shall see in section 5. Quadrature formulas generally come in families defined by a rule that specifies how the nodes and weights are determined for each choice of n, and we shall use the word "formula" to refer to both the fixed n case and the family.

The aim with any quadrature formula is that the error

$$E_n(f) = I_n(f) - I(f) \tag{1.3}$$

*Received by the editors January 4, 2021; accepted for publication (in revised form) June 7, 2021; published electronically February 3, 2022.
https://doi.org/10.1137/20M1389522

†Mathematical Institute, University of Oxford, Oxford, OX2 6GG, UK (trefethen@maths.ox.ac.uk).

should be small and, in particular, one would like the convergence to be rapid in the sense that $E_n(f)$ decreases rapidly to 0 as $n \to \infty$ when f is smooth. In applications, n may take values of all sizes. For adaptive quadrature software based on composite application of a rule (1.1) on subintervals, it may be less than 10; in the most common uses of Gauss–Legendre and Gauss–Hermite formulas it is perhaps in the dozens; in the numerical solution of integral equations and in the Chebfun and ApproxFun software systems [12, 34], it is often in the hundreds or thousands.

There is a standard design principle used in deriving quadrature formulas: the formula should be exact when applied to a certain class of integrands f. In the case of quadrature on an interval, this is usually the set P_{n-1} of polynomials of degree at most $n-1$, in which case the result returned by the quadrature formula is equal to the integral of the unique degree $n-1$ polynomial interpolant through the data $\{f(x_j)\}$ at the points $\{x_j\}$. This *exactness principle* has proved effective for a wide range of problems, and it is the starting point of most discussions of quadrature. It is the organizing theme of this paper, which is both a survey of some of the best-known quadrature formulas and also a presentation of new results.

We shall find that the exactness principle is not a reliable guide to the actual accuracy of four well-known quadrature rules: the Newton–Cotes, Clenshaw–Curtis, and Gauss formulas on $[-1,1]$, and the Gauss–Hermite formula on $(-\infty,\infty)$. The failure of the exactness principle is particularly extreme in the cases of Newton–Cotes quadrature (as is well known) and Gauss–Hermite quadrature (not so well known). We shall also mention related effects for additional problems including Gauss–Laguerre quadrature and cubature in a hypercube.

None of our observations depends on rounding error; they concern fundamental mathematical properties that hold in both floating-point and exact arithmetic. If our figures had been computed in exact arithmetic, they would all look the same apart from the elimination of the rounding error plateaus around 10^{-15} in Figures 4.2, 5.1, and 6.1.

There are a number of excellent books on quadrature, including the classic by Davis and Rabinowitz [8] and the more recent work by Brass and Petras [4].

2. Newton–Cotes Quadrature. Here and in the next two sections, our domain is the interval $D = [-1,1]$; other intervals $[a,b]$ are readily scaled to this standard setting. The *Newton–Cotes formula,* going back to Isaac Newton in 1676 and Roger Cotes in 1722, is the formula that results from taking $\{x_j\}$ as equally spaced points from -1 to 1, with $\{w_j\}$ determined so that $I_n(f) = I(f)$ for all $f \in P_{n-1}$. Most numerical analysis textbooks have a chapter on numerical integration in which they discuss two quadrature formulas. First, the Newton–Cotes formula is introduced and it is observed that it has polynomial exactness degree $n-1$. Then Gauss quadrature is presented, based on optimal points $\{x_j\}$ as defined by exactness degree (section 4), and in most textbooks, it is implied that the main advantage of Gauss over Newton–Cotes is that it has exactness degree $2n-1$.

This is highly misleading. Gauss quadrature is indeed better than Newton–Cotes, but the main reason is not its doubled degree of exactness. Newton–Cotes doesn't merely converge less quickly as $n \to \infty$, it diverges at an exponential rate for most integrands f, even when they are analytic. Meanwhile the Gauss formula is guaranteed to converge so long as f is continuous, and at an exponential rate if it is analytic.

The failure of Newton–Cotes quadrature for larger values of n is well known. This became apparent to experts after the appearance in 1901 of Runge's paper on polynomial interpolation in equispaced points [40], which shows that these interpolants

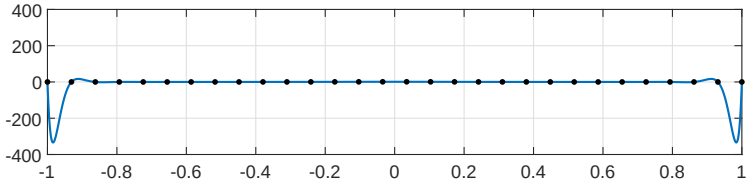

Fig. 2.1 *Degree 29 polynomial interpolant to the integrand $f(x) = 1/(1+25x^2)$ in 30 equispaced points of $[-1,1]$. The Runge phenomenon of oscillations near the boundary leads to exponential divergence of the Newton–Cotes formula as $n \to \infty$.*

Table 2.1 *Errors in Newton–Cotes integration of polynomials x^k and $T_k(x)$ with $n = 30$. (Only even values of k are shown since the integrals and errors are zero when k is odd.) The errors for $T_k(x)$ are huge for $k \geq n$, a reflection of exponentially large quadrature weights of alternating signs.*

k	$\lvert E_n(x^k)\rvert$	$\lvert E_n(T_k(x))\rvert$
⋮	⋮	⋮
26	0	0
28	0	0
30	.0000007	399.5
32	.000005	2711.1
34	.00002	8923.1
36	.00004	18765.9
38	.00009	27812.9
⋮	⋮	⋮

experience oscillations whose amplitude grows exponentially as $n \to \infty$, even for analytic functions f (see Figure 2.1). The conclusion was made rigorous by Pólya in 1933 [39]. Pólya also showed that convergence as $n \to \infty$ for all $f \in C([-1,1])$ occurs if and only if the sum $\sum_{j=1}^n |w_j|$ is bounded as $n \to \infty$ [8, sect. 2.7]. Since $\sum_{j=1}^n w_j = 2$ for any quadrature formula on $[-1,1]$ that integrates constants correctly, this condition certainly holds if the weights are positive, as is the case for the Gauss and Clenshaw–Curtis formulas. The Newton–Cotes weights, however, have alternating signs and grow in amplitude at a rate of order 2^n as $n \to \infty$.[1] Figure 2.1 illustrates the divergence of the Newton–Cotes formula with a plot of the degree 29 polynomial interpolant to $f(x) = 1/(1+25x^2)$ (the celebrated Runge function). The corresponding quadrature estimate is $I_{30}(f) \approx -21.8$. Even the sign is wrong, and the amplitude grows exponentially, with $I_{50}(f) \approx -24{,}965$, for example.

What is not well known is how this failure relates to the exactness principle, which we now consider in Table 2.1. With $n = 30$, the Newton–Cotes formula integrates x^k exactly for $k < 30$, with $E_n(x^k) = 0$. A check of $E_n(x^k)$ for $k \geq 30$ looks unexpectedly promising, with the errors coming out as very small, smaller than one might have dreamed for a quadrature formula of exactness degree $n-1$. However, this apparent good behavior is an illusion associated with the exponential ill-conditioning of monomial bases on $[-1,1]$. Although, algebraically, x^k is of course a polynomial of degree

[1] See the final formula of [37], from 1925, after which Ouspensky writes "One sees that the coefficients $A_2, A_3, \ldots, A_{n-2}$ tend to infinity, making it evident that the Cotes formula loses all practical value as the number of ordinates grows considerable."

k, its numerical degree is only $O(k^{1/2})$ in the sense that for any $\varepsilon > 0$, there is a constant $\delta > 0$ such that x^k can be approximated to accuracy ε on $[-1, 1]$ by polynomials of degree $< \delta k^{1/2}$ for all sufficiently large k [32]. That is why the numbers in the x^k column of Table 2.1 are so small. When we switch from x^k to the Chebyshev polynomial $T_k(x)$ in the final column of the table, huge errors set in immediately at degree $k = 30$.[2] Evidently for the Newton–Cotes formula, exact integration of degree $n - 1$ polynomials has told us next to nothing about accuracy in integrating other functions.

3. Clenshaw–Curtis Quadrature. The *Clenshaw–Curtis formula*, originating in 1960 [6], consists of integrating the degree $n - 1$ polynomial interpolant through n *Chebyshev points*

$$x_j = \cos(j\pi/(n-1)), \quad 0 \le j \le n-1. \tag{3.1}$$

This is the natural formula to apply in the context of Chebyshev spectral collocation methods for differential equations, and it is essentially the method by which Chebfun integrates a function, after first reducing it to a polynomial of sufficiently high degree [12]. Alternatively, nodes and weights can be computed explicitly in $O(n \log n)$ operations [50] and are available in Chebfun with the command `[x,w] = chebpts(n)`.

Clenshaw–Curtis quadrature, like Newton–Cotes, has polynomial exactness degree $n - 1$, but with none of the misbehavior as $n \to \infty$ since the weights are always positive. Pólya's theory guarantees convergence for all $f \in C([-1, 1])$ at a rate that follows the smoothness of f; if f is analytic, the convergence is exponential [48, Thms. 19.3 and 19.4]. Thus the obvious expectation for Clenshaw–Curtis quadrature is that it should converge like Gauss quadrature, but at approximately half the rate, since Gauss has polynomial exactness degree $2n - 1$.

This is not what happens. Gauss quadrature behaves as expected; the surprise is that Clenshaw–Curtis often converges at the Gauss rate too [48, Thm. 19.5]. Figure 3.1 shows a typical example. This effect was noted experimentally by Clenshaw and Curtis themselves, who wrote: "We see that the Chebyshev formula, which is much more convenient than the Gauss, may sometimes nevertheless be of comparable accuracy" [6]. A paper on the subject was published in 1968 by O'Hara and Smith, who wrote, "The Clenshaw–Curtis method gives results nearly as accurate as Gaussian quadratures for the same number of abscissae" [33]. Subsequently, the effect was mentioned in the books of Evans [13] and Kythe and Schäferkotter [24] and then became more widely known through a paper of mine in 2008 [45].

O'Hara and Smith's explanation of the unexpected accuracy of the Clenshaw–Curtis formula is that although its errors in integrating degree k polynomials are nonzero for $k \ge n$, they are still very small for $n \le k \ll 2n$. Note that this is precisely a failure of exactness as a guide to accuracy. The small errors are shown numerically in Table 3.1, a repetition of Table 2.1 (without the distracting x^k column) for Clenshaw–Curtis. Figure 3.2 gives a visual picture of what is going on. The reason the errors oscillate up and down has to do with aliasing, much as wagon wheels in a movie may seem to turn forward, backward, then forward again as the wagon speeds up. Specifically, on the n-point grid, the Chebyshev polynomial T_{N+j} takes the same values as $T_{|j|}$ whenever N is an even multiple of $n - 1$ and $|j| \le n - 1$ [48, Thm. 4.1].

[2] The phenomenon that x^k has numerical degree $O(k^{1/2})$ is hidden in a discussion of quadrature many years ago by Gautschi [17]. Gautschi writes of the "futility of moment-related tests" of accuracy, showing that a formula can integrate x^k very accurately and yet be inaccurate for other integrands. The trouble goes away if x^k is replaced by $T_k(x)$.

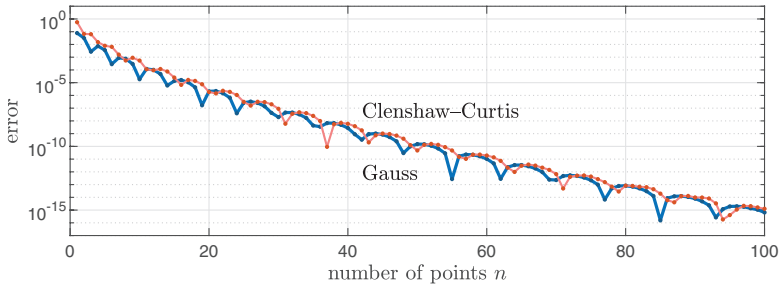

Fig. 3.1 *Convergence of the Clenshaw–Curtis and Gauss quadrature formulas for the integrand* $f(x) = \exp(-1/x^2)$. *The Gauss convergence is at the expected rate; the surprise is that Clenshaw–Curtis converges at this rate too.*

Table 3.1 *Errors in Clenshaw–Curtis integration of polynomials* $T_k(x)$ *with* $n = 30$. *Though nonzero, the errors for* $n \le k \ll 2n$ *are small, of order* $O(n^{-2})$. *This is the explanation given by O'Hara and Smith* [33] *of the unexpected accuracy of Clenshaw–Curtis.*

k	$\lvert E_n(T_k(x))\rvert$
⋮	⋮
26	0
28	0
30	0.0003
32	0.001
34	0.002
⋮	⋮
54	0.1
56	0.7
58	2.0
60	0.7
⋮	⋮

When k is close to an even multiple of $n-1$, T_k therefore looks like a low-degree polynomial on the grid, which Clenshaw–Curtis integrates to a number $O(1)$, whereas the true answer should be $O(k^{-2})$, giving an error of $O(1)$. These are the peaks in the plot. When k is close to an odd multiple of $n-1$, on the other hand, T_k looks more like T_{n-1}, giving a smaller integral and error of order $O(n^{-2})$.

The effect shown in Table 3.1 and Figure 3.2 leads readily to an understanding of the surprising convergence rate of Clenshaw–Curtis quadrature as seen in Figure 3.1. Any Lipschitz continuous integrand f will have an absolutely and uniformly convergent Chebyshev series,

$$f(x) = \sum_{j=0}^{\infty} a_j T_j(x), \tag{3.2}$$

from which it follows that the error (1.3) in Clenshaw–Curtis quadrature is

$$E_n(f) = \sum_{\substack{j=n \\ j \text{ even}}}^{\infty} a_j E_n(T_j). \tag{3.3}$$

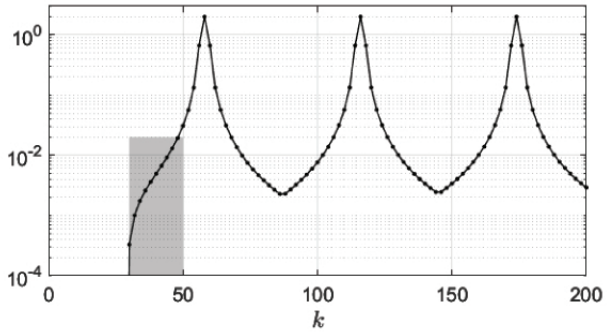

Fig. 3.2 *Errors in Clenshaw–Curtis integration with $n = 30$ of $T_k(x)$ as a function of even indices k. The shading highlights the small errors for $n \le k \ll 2n$.*

If all the errors $E_n(T_j)$ were of the same size $O(1)$, then $E_n(f)$ would depend just on the Chebyshev coefficients $\{a_j\}$, and this is approximately what happens in the regime $2n \le j < \infty$. For the "shaded coefficients" a_j with $n \le j \ll 2n$, however, the errors $E_n(T_j)$ are of size $O(n^{-2})$, and whether or not they contribute significantly to $E_n(f)$ depends on the rate of decrease of $\{a_j\}$ as $j \to \infty$, and hence on the regularity of f. If f is not analytic, as in the example of Figure 3.1, then $\{a_j\}$ decreases slowly enough that $a_j E_n(T_j)$ is much smaller for $n \le j \ll 2n$ than for $j \approx 2n$, making the contributions of the shaded coefficients to $E_n(f)$ negligible and producing the doubled-degree effect in its cleanest form. (There is still a complication, however, in that both Clenshaw–Curtis and Gauss converge at a rate faster than expected by one power of n [53].) If f is analytic, then $\{a_j\}$ decreases exponentially and the terms $a_j E_n(T_j)$ for $n \le j \ll 2n$ are no longer negligible in comparison to the later ones, making the asymptotic convergence rate of Clenshaw–Curtis indeed half that of Gauss. For details, including the "kink phenomenon" observed for Clenshaw–Curtis quadrature of analytic integrands as the initial Gauss convergence rate cuts in half after a certain value of n, see [45, 52].

4. Gauss Quadrature. Gauss quadrature, introduced by Gauss in 1814 [15], is defined by having the maximum possible polynomial degree of exactness, $2n - 1$. This is achieved by taking the nodes $\{x_j\}$ as the roots of the degree n Legendre polynomial P_n, and the name *Gauss–Legendre* is also used to distinguish this case from that of integrals in which a nonconstant weight function is introduced in (1.2) (see the next section). The weights for Gauss quadrature $\{w_j\}$ are all positive and the formula is extremely effective in practice. Thanks to new algorithms introduced in the past 15 years and available with the Chebfun command `[x,w] = legpts(n)`, the nodes and weights can be computed in a fraction of a second even when n is in the millions [3, 12, 20].

Gauss quadrature is often described as optimal, but this is only precisely true in senses that are tied to polynomials. Specifically, it can be shown to be optimal by certain measures for integrating functions that are analytic in a *Bernstein ellipse,* an ellipse in the complex plane with foci ± 1; see [38] and sections 4.9 and 6.9 of [4]. Analyticity in an ellipse, however, is a skewed form of smoothness, requiring a function's Taylor coefficients for expansions in the middle of the interval to grow more slowly than those for expansions near the endpoints. Polynomials can resolve much

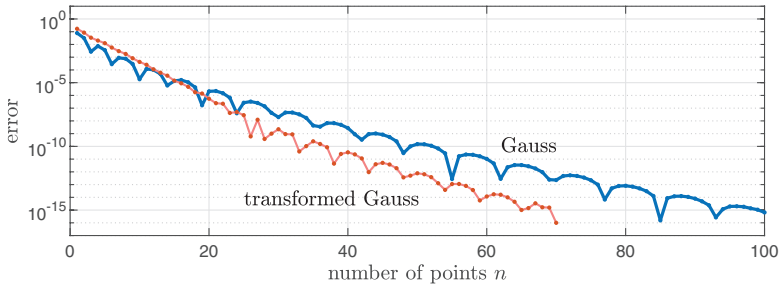

Fig. 4.1 *Convergence of the Gauss formula and a conformally transformed Gauss formula* (4.2) *for the same integrand* $f(x) = \exp(-1/x^2)$ *as in Figure* 3.1. *While such transformed formulas have a limited role in practice, they illustrate that Gauss quadrature is not optimal.*

faster wiggles near endpoints than in the interior [11], and Gauss quadrature (likewise Clenshaw–Curtis) inherits this property—as one sees intuitively from its strong clustering of sample points near the ends. For example, Gauss quadrature converges faster for the integrand $\sqrt{1.01 - x}$ than for $\sqrt{0.1i - x}$, though the singularity of the first function is ten times closer to $[-1, 1]$ than that of the second.[3]

From a user's point of view, it would seem more natural to consider integrands with uniform smoothness across $[-1, 1]$. In the analytic case, one might require analyticity in an ε-neighborhood of this interval, and quadrature formulas based on this assumption can be derived by transplanting Gauss quadrature via a conformal map g with $g([-1, 1]) = [-1, 1]$ of a Bernstein ellipse onto a neighborhood of $[-1, 1]$ with approximately straight sides (a "cigar," one might say), or more simply, an infinite strip. Here the integral (1.2) becomes

$$I(f) = \int_{-1}^{1} g'(s) f(g(s)) ds, \tag{4.1}$$

and applying Gauss quadrature in the s variable gives the transformed quadrature formula

$$I_n(f) = \sum_{j=1}^{n} w_j g'(s_j) f(g(s_j)). \tag{4.2}$$

An example is shown in Figure 4.1, with g taken as the conformal map of the Bernstein ellipse with parameter $\rho = 1.4$ (the sum of the semiminor and semimajor axes) onto an infinite strip. Such transplanted formulas were introduced in [21], and the conformal mapping idea goes back in the theoretical literature to Bakhvalov in 1967 [2] and was applied for spectral methods by Kosloff and Tal-Ezer [23]. As expected, the transformed quadrature nodes are much more uniformly distributed, with density $\approx \pi/2$ times greater in the middle of the interval than for Gauss quadrature (not shown).

A different approach to developing quadrature formulas with more uniform behavior, *band-limited quadrature,* is based on time-frequency analysis. Suppose one seeks a formula (1.1) that will integrate the functions $\exp(ikx)$ to high accuracy for all the wave numbers $k \in [-c, c]$ for some $c > 0$. Note that this is a continuum of

[3] In Chebfun, try `cheb.x, plotcoeffs([sqrt(1.01-x) sqrt(0.1i-x)])`.

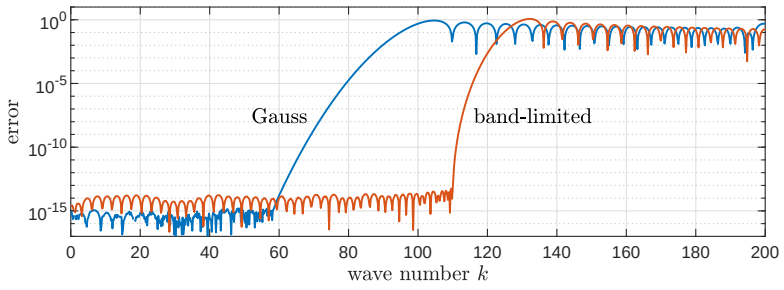

Fig. 4.2 *Errors of 50-point Gauss and band-limited quadrature formulas integrating* $\exp(ikx)$ *for* $k \in [-c, c]$*, with* c *given by* (4.3)*. By this measure of integrating complex exponentials, the band-limited formula is notably more efficient.*

wave numbers, not just the integer multiples of π that would make the integrand 2-periodic. No choice of nodes and weights can integrate the continuum exactly, because its dimension is infinite. However, it has been known since the work of Slepian, Landau, and Pollak at Bell Labs in the 1950s and 1960s that the numerical dimension of this space is finite, just a bit larger than $2c/\pi$ [25, 26, 41]. Specifically, the singular values σ_j of the bivariate kernel function $\exp(ikx)$ for $x \in [-1, 1]$, $k \in [-c, c]$ decrease exponentially for $j > 2c/\pi$.[4] By applying the method known as generalized Gauss quadrature [5, 36] to an appropriate set of prolate spheroidal wave functions (PSWFs), one can develop quadrature rules that integrate these functions to high accuracy, as shown by Xiao, Rokhlin, and Yarvin (see [36, 54]).[5] Later work has led to fast methods for calculating the nodes and weights of these formulas numerically [35, 36].

In the literature of band-limited quadrature, the usual point of view is that these formulas are targeted at the special case of integrands that are band-limited to a known range $k \in [-c, c]$ or nearly so. Although the formulas are indeed excellent in such cases, this is a bit like saying that Gauss quadrature is targeted at the integration of polynomials. In fact, band-limited quadrature can be used for general integrands just like Gauss, provided one eliminates the free parameter c by making it depend systematically on n. To this end, for quadrature to 15-digit accuracy, we propose the formula

$$c = \pi n - 12 \log n \qquad (n \geq 8). \tag{4.3}$$

The constant 12 has been chosen empirically, but there is a good deal of relevant theory; see [27] and [36, Thm. 2.4 and Prop. 17]. Figure 4.2 shows the accuracy of band-limited quadrature based on (4.3) with $n = 50$ for approximating $\exp(ikx)$ for various k. Figure 4.3 repeats the convergence comparison of Figures 3.1 and 4.1.[6]

Both of the alternatives to Gauss quadrature we have outlined in this section lead to the conclusion that the potential gain is a factor of $\pi/2$ in convergence rate. This is hardly big enough to be of great importance in practice (at least in one space dimension [46, Fig. 11]), and I share the view, which is discussed with particular

[4] In Chebfun, try `K = chebfun2(@(x,k) exp(1i*k.*x),[-1 1 -c c]), semilogy(svd(K),'.')`.

[5] Recent experiments by Jim Bremer (unpublished) show that essentially the same results can be obtained without use of PSWFs by applying generalized Gauss quadrature directly to the functions $\exp(ikx)$.

[6] The Chebfun calculations made use of `c = pi*n-12*log(n), [s,w] = pswfpts(n,c,'ggq')`.

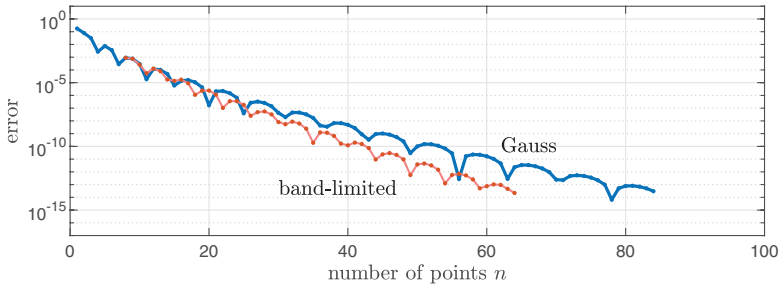

Fig. 4.3 *Comparison of the Gauss and band-limited quadrature formulas for the integrand* $f(x) = \exp(-1/x^2)$*, as in Figures* 3.1 *and* 4.1*, again with* c *given by* (4.3)*. The curves stop at accuracy* 10^{-13} *because our computation of nodes and weights in the band-limited case is a few digits off machine accuracy, as was evident in Figure* 4.2*.*

substance in section 6.9 of [4], that Gauss quadrature is the best choice for general use.

5. Gauss–Hermite Quadrature. Gauss–Hermite quadrature is the standard Gauss quadrature method for integrals over the whole real line. One supposes that a function f is given and one wants to compute the integral

$$I = \int_{-\infty}^{\infty} e^{-x^2} f(x)\,dx, \tag{5.1}$$

which is (1.2) with $D = (-\infty, \infty)$ modified by the introduction of the weight function $\exp(-x^2)$. The approximation will take the form (1.1) as usual, now with nodes $\{x_j\}$ in $(-\infty, \infty)$, and the Gauss–Hermite formula is defined by the nodes and weights taking the unique values such that $I(f) = I_n(f)$ whenever f is a polynomial of degree $2n-1$. According to Gautschi [16], this method was introduced by Gourier in 1883 [19]. As with its unweighted progenitor discussed in the last section, Gauss–Hermite quadrature is intimately associated with orthogonal polynomials [42]. These are the *Hermite polynomials* $H_0(x) = 1$, $H_1(x) = 2x$, $H_2(x) = 4x^2 - 2, \ldots$, which are orthogonal over $(-\infty, \infty)$ with respect to the inner product

$$\langle f, g\rangle = \int_{-\infty}^{\infty} e^{-x^2}\, \overline{f(x)}\, g(x)\,dx. \tag{5.2}$$

The nodes $\{x_j\}$ are the roots of H_n, and as with Gauss–Legendre quadrature, algorithms have been developed to compute the nodes and weights with just $O(n)$ work [18, 44]. In Chebfun, `[x,w] = hermpts(n)`.

In an application, one might start from an integrand $g(x)$ with Gaussian decay, so that g can be written $g(x) = \exp(-x^2)f(x)$, where f is a bounded function on $(-\infty, \infty)$. It has been recognized from the beginning that there may be problems with determining the right decay behavior. What if $g(x)$ looks more like $\exp(-\sigma x^2)$ for some $\sigma > 0$ not equal to 1, or decays in a non-Gaussian manner? We shall avoid these questions and assume that $\exp(-x^2)$ well approximates the shape of g.

Even in this most favorable setting, Gauss–Hermite quadrature is terribly inefficient as $n \to \infty$. For large n, most of the nodes lie far enough out along the real axis that their weights are minuscule; these terms then contribute negligibly to the sum (1.1) and can be thrown away! This curious phenomenon is known to experts,

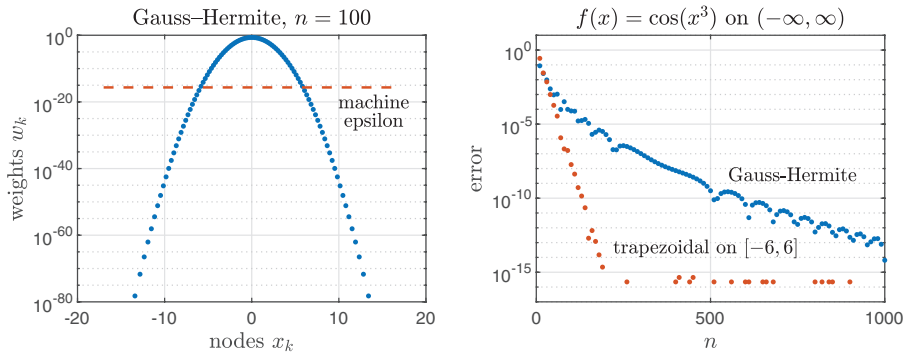

Fig. 5.1 *A practitioner's view of the inefficiency of Gauss–Hermite quadrature for large n. Many of the weights w_j lie below machine precision—a fraction approaching 100% as $n \to \infty$. The convergence rate is just root-exponential, whereas if one truncates to a finite interval and applies a standard quadrature method, it becomes exponential. The missing dots for the trapezoidal rule correspond to errors that are exactly zero in floating-point arithmetic.*

but I am unaware of any discussion in the literature of where the reasoning that led to the Gauss–Hermite formula has broken down. How can a formula that is optimal in a precise mathematical sense be so plainly suboptimal in practice? We shall first illustrate the problem, introducing Theorem 5.1 to make the point solid, and then give an answer to this question.

The right image of Figure 5.1 presents the inefficiency of Gauss–Hermite quadrature as a practitioner might encounter it. For the integrand we take $f(x) = \cos(x^3)$, mixing energy at all wave numbers. The errors for Gauss–Hermite quadrature as a function of n line up along a parabola on a semilog scale, corresponding to slow root-exponential convergence (i.e., errors of size $O(\exp(-C\sqrt{n}))$ for some $C > 0$), and even with $n = 1000$ the error is no smaller than 10^{-13}. Yet $\exp(-x^2)$ is of the order of 10^{-16} or less for $|x| > 6$, so for practical purposes, this integral might as well be posed on the compact interval $[-6, 6]$. Sure enough, the periodic trapezoidal rule applied on that interval shows much faster exponential convergence down to machine precision, and the convergence would be similarly fast for the Clenshaw–Curtis or Gauss formulas on $[-6, 6]$. If $f(x)$ is changed to $\exp(-1/x^2)$ as in Figure 3.1 (not shown), the convergence curves have about the same shapes though with about half the convergence rate.

The left side of Figure 5.1 shows how the Gauss–Hermite formula wastes its effort. For $n = 100$, about half the weights lie below machine precision, 48 of them to be exact, and this fraction will increase with n. For $n = 1000$, the number of weights below machine precision is 836. If f is of order 1, these function samples will not normally contribute to the computed value of the integral (though the weights themselves may be representable in floating point so long as they are bigger than about 10^{-308}). Figure 5.2 illustrates the problem from another angle, plotting the *Hermite functions*

$$\psi_n(x) = H_n(x) e^{-x^2/2} (\sqrt{2\pi}\, n!)^{-1/2}, \quad n = 0, 1, \ldots, \tag{5.3}$$

for three values of n. These functions form a complete orthonormal set in the unweighted space $L^2(-\infty, \infty)$, and Gauss–Hermite quadrature implicitly expands f in

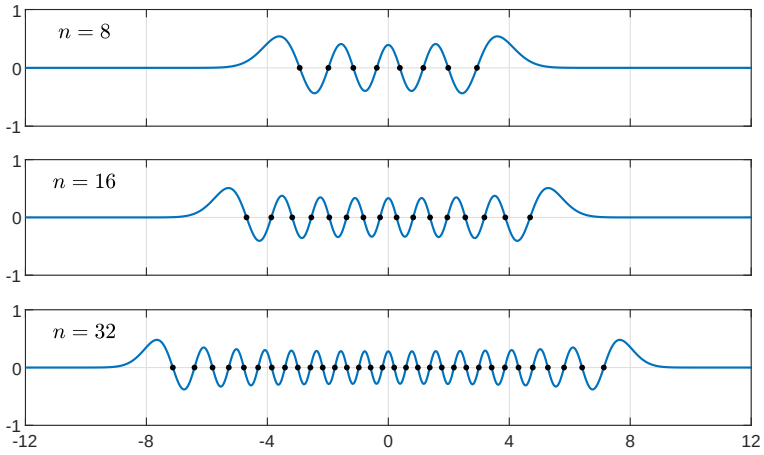

Fig. 5.2 *Hermite functions* (5.3)*, a complete orthonormal set in* $L^2(-\infty,\infty)$ *whose zeros are the Gauss–Hermite quadrature nodes* $\{x_j\}$*. The interval around* $x = 0$ *where the amplitude is* $O(1)$ *broadens in proportion to* $\sqrt{n}$*, even though the integrand of* (5.1) *decays at the rate* $\exp(-x^2)$.

this basis.[7] It is clear from the figure that this is going to be an inefficient way of capturing the part of f that matters to the integral, with ψ_{32}, for example, taking values of size $O(1)$ all across the interval $[-8,8]$ even though $\exp(-8^2) \approx 10^{-28}$.

We have spoken of machine precision for vividness, but the essential point has nothing to do with rounding error. In exact arithmetic, we can get convergence of $I_n(f)$ to $I(f)$ with I_n defined by quadrature over a finite interval, provided the interval grows as $n \to \infty$. The right choice for analytic functions f bounded on $(-\infty,\infty)$ is to truncate to an interval of size $[-O(n^{1/3}), O(n^{1/3})]$, which balances a domain-truncation error of order $\exp(-n^{2/3})$ and a discretization error of the same order since the sample step size will be $h = O(n^{-2/3})$. (Discretization errors for the trapezoidal rule are quantified in [49]; the tool for such estimates is Cauchy integrals.) Note that $[-O(n^{1/3}), O(n^{1/3})]$ is much narrower than the interval $[-O(n^{1/2}), O(n^{1/2})]$ sampled by the Gauss–Hermite formula. Comparing the two enables us to quantify the inefficiency of Gauss–Hermite quadrature. In effect, only a fraction of order $n^{1/3}/n^{1/2} = n^{2/3}$ of the Gauss–Hermite nodes is utilized, meaning that Gauss–Hermite quadrature employs more nodes than necessary by a factor of order $n^{1/3}$. The ratio increases to nearly order $n^{1/2}$ for nonanalytic functions f, where intervals growing just logarithmically rather than algebraically with n are appropriate for balancing domain-truncation and discretization errors.

A theorem makes the point precise; the proof is given in the appendix. This result supports the view that the best method for evaluating integrals of type (5.1) is not to apply the Gauss–Hermite formula, but to truncate the interval to finite length and apply an unweighted rule like Gauss–Legendre. The constants L and a in the theorem are arbitrary, so long as they are positive. By "equispaced trapezoidal quadrature" we mean that for each n, the trapezoidal rule is applied on any equispaced set of

[7] Hermite functions have been well known to physicists since the 1920s, for they are the eigenfunctions of the Schrödinger equation for the harmonic oscillator. In Chebfun, try `x = chebfun('x',[-3,3]), quantumstates(x^2,25)`.

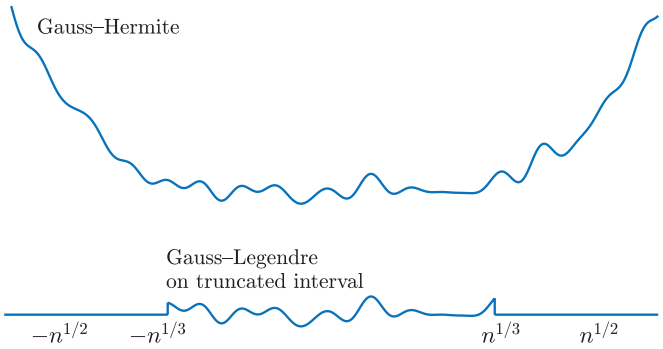

Fig. 5.3 *Schematic illustration of the approximation problems $f \approx v \in V_n$ associated with Gauss–Hermite quadrature for* (5.1) *and the alternative method of applying the Gauss–Legendre formula on a truncated interval $[-O(n^{1/3}), O(n^{1/3})]$, as in Theorem* 5.1. *The Gauss–Legendre method only needs to approximate f on $[-O(n^{1/3}), O(n^{1/3})]$ to achieve an error* (5.5) *of size $O(\exp(-O(n^{2/3})))$. The Gauss–Hermite method is effectively forced to approximate on a wider interval $[-O(n^{1/2}), O(n^{1/2})]$ just to keep the growth of polynomials under control.*

nodes exactly n of which fall in the truncated interval $[-Ln^{1/3}, Ln^{1/3})$. One might choose points such that this amounts to the periodic trapezoidal rule on the truncated interval, but this is unnecessary, and the proof does not rely on periodicity.

THEOREM 5.1. *Let $f(x)$ be analytic and bounded for $x \in (-\infty, \infty)$, and suppose $\exp(-x^2)f(x)$ extends to a bounded analytic function in the infinite strip $-a < \mathrm{Im}\, x < a$ for some $a > 0$. Let $L > 0$ be fixed, and for each $n \geq 1$, let I_n be the estimate of the integral I of* (5.1) *obtained by applying Gauss–Legendre, Clenshaw–Curtis, or equispaced trapezoidal quadrature on the truncated interval $[-Ln^{1/3}, Ln^{1/3}]$. Then for some $C > 0$,*

$$|I - I_n| = O(\exp(-Cn^{2/3})), \qquad n \to \infty. \tag{5.4}$$

To reiterate, the rate $O(\exp(-Cn^{2/3}))$ is much faster than the rate $O(\exp(-Cn^{1/2}))$ achieved by Gauss–Hermite quadrature.

In the literature, several authors have recommended adjustments to Gauss–Hermite quadrature that are related in one manner or another to truncating the domain to a finite interval. Mastroianni and Monegato and their coauthors have published a number of papers in this direction, focusing mainly on the analogous case of Gauss–Laguerre quadrature on $[0, \infty)$ [29, 30] (see the next section). Townsend, Trogdon, and Olver note that for large n, only about $25\sqrt{n}$ of the weights w_j are greater than the smallest machine number $\approx 10^{-308}$ in IEEE double precision, and they recommend "subsampling" to retain only nodes and weights that will matter [44]. In unpublished work presented at a conference in 2018, Weideman showed how $O(\exp(-n^{2/3}))$ convergence with a particularly favorable constant can be obtained by applying the Gauss–Hermite formula with the weight function $\exp(-(n/2)^{1/3}x^2)$ rather than $\exp(-x^2)$ [51].

However, the literature seems not to confront the conceptual question: what has gone wrong with the Gauss–Hermite notion of optimality? Here is an answer, summarized schematically in Figure 5.3. Suppose we have a quadrature formula for (5.1) that is exact for all functions v in an n-dimensional space V_n. Its effectiveness will depend on how well f can be approximated by functions $v \in V_n$ in the weighted

1-norm

$$\|f - v\| = \int_{-\infty}^{\infty} e^{-x^2} |f(x) - v(x)| dx. \tag{5.5}$$

(The study of weighted approximation problems on the real line was initiated by Bernstein; see [28].) Now Gauss–Hermite quadrature corresponds to taking V_n as the space of polynomials of degree $n-1$, and this space is terribly inefficient for these weighted approximations since polynomials grow so fast as $|x|$ increases. For example, $x^n \exp(-x^2)$ reaches a maximum of $(n/2e)^{n/2}$ at $x = \sqrt{n/2}$ — about 10^6 for $n = 20$ and 10^{17} for $n = 40$. These huge numbers force a polynomial approximation to f to pay attention to the whole interval $[-O(n^{1/2}), O(n^{1/2})]$, just to keep (5.5) under control. By contrast, when we apply Gauss–Legendre quadrature on a truncated interval $[-O(n^{1/3}), O(n^{1/3})]$, we are changing the approximation space V_n to the set of polynomials of degree $n-1$ multiplied by the characteristic function of this interval. The approximation can focus on just the short interval, where much better accuracy is achievable.

6. Three More Examples. We now mention three further examples of quadrature formulas for which exactness proves to be an inaccurate guide to accuracy.

The first is Gauss–Laguerre quadrature, the analogue of (5.1) for integration over a semi-infinite interval:

$$I = \int_0^{\infty} e^{-x} f(x) dx. \tag{6.1}$$

The design principle is that the nodes and weights should be such that the formula is exact when f is a polynomial of degree $2n-1$. (In Chebfun, they can be computed by `[x,w] = lagpts(n)`, an implementation of a fast algorithm of Huybrechs and Opsomer [22].) As with Gauss–Hermite quadrature, one finds that many of the weights are so small that the corresponding terms contribute negligibly to the result. For $n = 100$, the minimal weight is 10^{-162} (decreasing exponentially with n), and just 38 weights lie above standard machine precision (a fraction increasing as $O(n^{1/2})$). For a bounded analytic function f on $[0, \infty)$ one has a domain-truncation error of size $\exp(-O(n))$ and a discretization error of order $\exp(-O(n^{1/2}))$, so the two are out of balance. By dropping nodes and weights or truncating to a shorter interval one can do much better, as illustrated in Figure 6.1. Truncated Laguerre formulas have been investigated in detail by Mastroianni and Monegato and their coauthors, though mostly for more general functions f with the property that $e^{-x} f(x)$ may decay only algebraically as $x \to \infty$ [29, 30].

The second example involves the trapezoidal rule, that is, the approximation of an integral of a function $f(x)$ by the integral of its piecewise linear interpolant in a given set of sample points. Here the exactness principle is that $I_n(f) = I(f)$ whenever f is a piecewise linear function with breaks at the nodes, which leads readily to an $O(h^2)$ accuracy bound for any f that is twice continuously differentiable, where h is the maximum distance between adjacent nodes. However, there is an important special case in which the convergence is much faster, which we exploited in Figure 5.1: when f is smooth and periodic and the nodes are equally spaced. If f is analytic, the convergence becomes exponential at a rate of $\exp(-O(n))$. The exactness principle based on piecewise linear interpolants fails to detect this, but it can be rescued (at least up to a Gauss factor of 2 attributable to aliasing) by the observation that for

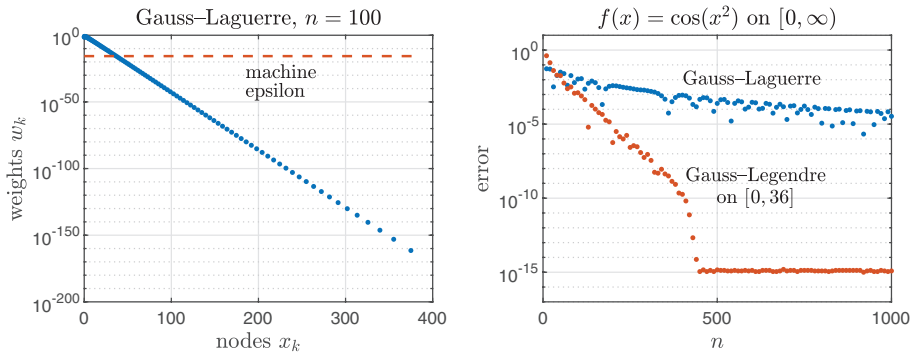

Fig. 6.1 *Analogue of Figure* 5.1 *for Gauss–Laguerre quadrature, now with* $f(x) = \cos(x^2)$. *As with Gauss–Hermite, many of the weights lie below machine precision, and the convergence rate is just root-exponential. One does much better by discarding nodes and weights or truncating to a finite interval and applying an unweighted formula like Gauss–Legendre.*

equispaced points, the piecewise linear interpolant has the same integral as a trigonometric interpolant. (Proof: in both cases the integral is equal to the mean of the sample data times the length of the interval.) For the trigonometric interpolant, fast convergence is readily proved via contour integrals or aliasing of Fourier series. Thus the special accuracy of the periodic equispaced trapezoidal rule can be understood by an exactness principle after all, so long as it is based on trigonometric interpolants instead of piecewise linear interpolants. For details, see [49].

The third example concerns *cubature formulas* for integration in the s-dimensional hypercube $D = [0,1]^s$. Following an idea introduced by James Clerk Maxwell in 1877 [31], one may approximate I by the integral of a degree d multivariate polynomial interpolant through a set of data values at n points $x_j \in D$. For each $d \geq 0$, the dimension of the space of polynomials is $N = \binom{s+d}{s}$, and one would hope to be able to have $n \leq N$ nodes with positive weights. A theorem of Tchakaloff asserts that this is indeed possible (in greater generality, not just for uniformly weighted integrals over a hypercube), even though a suitable set of nodes may be difficult to determine [43]. For extensive information about cubature formulas, see the survey article by Cools [7].

The difficulty with the exactness principle for cubature formulas concerns the limitations of multivariate polynomials as a guide to accuracy. The standard definition of the degree d of a multivariate polynomial is the maximum 1-norm of its exponents; thus $p(x,y) = x^3y^4$, for example, has degree $d = 7$. This definition has the property of isotropy (i.e., rotation-invariance) in s-space: if the variables are transformed by a rotation, the degree of a polynomial p is unchanged. However, the hypercube is itself far from isotropic. The diagonals are $\sqrt{s}$ times longer than the diameters, and there are 2^d of them, so most of its volume is "in the corners" in the sense of lying outside the inscribed hypersphere. As a consequence, cubature formulas designed on Maxwell's principle will behave anisotropically in $[0,1]^2$, giving better accuracy for an integrand $f(x) = \varphi(x_1)$ aligned along one axis, say, than for the same function rotated to $f(x) = \varphi((x_1 + \cdots + x_s)/\sqrt{s})$. For angle-independent resolution in the hypercube it would be necessary to base cubature formulas instead on the *Euclidean degree*, defined in terms of the 2-norm of the exponents. (For example, the Euclidean degree of x^3y^4 is 5.) This effect was first pointed out in [46], and a theorem making it precise was published in [47]. Note the analogy to the discussion of Gauss quadrature

on $[-1,1]$ in section 4. There the issue was translation-invariance, whereas here it is rotation-invariance.

It would be interesting to provide a numerical illustration of the suboptimality of standard cubature formulas. In preparing this article, however, I have come to realize that cubature formulas are not used much in practice, making it unclear exactly what nodes and weights might be appropriate for such a comparison. It appears that most multiple integrals are handled by essentially one-dimensional methods such as tensor products, or by Monte Carlo methods and their relatives, or by other more specialized tools [10]. Nevertheless, we can estimate how much efficiency should be lost, in principle, by building cubature formulas in the standard manner on the total degree. The number of coefficients needed to specify a polynomial of Euclidean degree d is d^s times the volume of an orthant of the s-hypersphere,

$$N_{\text{euclidean}} = \frac{d^s \pi^{s/2}}{2^s (s/2)!} \sim \frac{d^s}{\sqrt{\pi s}} \left(\frac{\pi e}{2s}\right)^{s/2},$$

with the asymptotic approximation $\sim$ referring to the limit $d \to \infty$. To get the same resolution via the total degree would require the number of coefficients to be d^s times the volume of the s-simplex expanded by $\sqrt{s}$ in each direction,

$$N_{\text{total}} = \frac{d^s s^{s/2}}{s!} \sim \frac{d^s e^s}{s^{s/2}\sqrt{2\pi s}}.$$

The inefficiency ratio associated with the standard design principle of cubature formulas is accordingly

$$\frac{N_{\text{total}}}{N_{\text{euclidean}}} = \frac{(s/2)!}{s!} \left(\frac{4s}{\pi}\right)^{s/2} \sim \frac{1}{\sqrt{2}} \left(\frac{2e}{\pi}\right)^{s/2} \approx (1.3)^{s-1}.$$

For dimensions $s = 1$, 2, 5, 20, and 40 the ratios are about 1, 1.27, 2.83, 171, and 41104.

7. Discussion. Quadrature theory is an edifice built over 200 years, featuring both detailed estimates and general theories. Often it may be hard to extract the important points, not least because valid estimates have a way of turning out to be far from sharp. In this article I have taken the exactness principle as the unifying theme, collecting both known results and new ones to show how this principle may mislead even in the cases we know best and that get the most attention in our textbooks. Of course the treatment has not been exhaustive, and one could note other related effects such as Simpson's rule having order 4 when one might have expected 3 (similarly for other odd-order Newton–Cotes formulas) and surprises in the accuracy of Gauss–Kronrod formulas. It is worth mentioning that although polynomials dominate this kind of analysis, other bases come into play too, including piecewise polynomials (for composite rules) and trigonometric functions (for periodic integrands, as in the discussion of the trapezoidal rule in the last section). Periodic integrands have received special attention in the context of high-dimensional integrals, for example, in the use of lattice methods and other quasi-Monte Carlo rules [10].

Quadrature theory is not a hot research area nowadays; like complex analysis, it is a field that we use all the time but which has a way of seeming finished. A book that I have found particularly valuable is the 2011 monograph by Brass and Petras [4]. The opening chapter presents a "standard estimation framework" that

elegantly makes precise the central question: given a quadrature formula, how can we speak quantitatively of its degree of optimality in integrating functions of particular classes? The book is full of results and references to detailed work on these problems, which show in case after case how the exactness principle is always a part of the analysis of accuracy but is never the whole story. A definite integral is just one example of a linear functional, and the estimation framework for quadrature is related to wide-ranging theories including n-widths, information-based complexity, and optimal recovery. A recent contribution in this area is [9], and a textbook by Simon Foucart is forthcoming that will present links to data science and machine learning [14].

The exactness principle for designing quadrature formulas is algebraic, a matter of whether or not certain quantities are exactly zero. This link to algebra is the reason exactness has proved to be such a fruitful tool for deriving formulas. Ultimately, however, quadrature is a problem of analysis, concerned with whether or not certain quantities are small. It is to be expected that there will be some discrepancy between the two points of view.[8] Still, it is surprising that sometimes the discrepancy can be huge without our having fully noticed it, as in the case of Gauss–Hermite quadrature.

A referee asks, do the themes we have discussed lead to any advice on how to select the right quadrature formula for an application? Perhaps the advice is that if you are told that a certain formula is optimal, do not assume this is the end of the discussion.

Appendix. Proof of Theorem 5.1.

Proof. Write $g(x) = \exp(-x^2)f(x)$, and let $L > 0$ be fixed arbitrarily as indicated. For each $n \geq 1$, let J_n denote the interval $J_n = [-Ln^{1/3}, Ln^{1/3}]$. By assumption, $f(x)$ is bounded for $x \in (-\infty, \infty)$. Therefore, truncating $(-\infty, \infty)$ to J_n introduces an error in the integral of g of size $O(\exp(-Cn^{2/3}))$ for some C. To prove the theorem, it will accordingly be enough to show that the Gauss–Legendre, Clenshaw–Curtis, and trapezoidal approximations to the integral of g over J_n also have accuracy of this order $O(\exp(-Cn^{2/3}))$.

For the trapezoidal rule, this follows from a minor adaptation of the contour integral argument used to prove Theorem 3.2 of [49]. That theorem asserts that for a 2π-periodic function analytic and bounded in the strip of half-width $\beta > 0$ about the real axis, the n-point periodic trapezoidal rule has errors of size $O(\exp(-an))$ as $n \to \infty$. This is proved by writing the trapezoidal rule estimate in terms of a contour integral over the rectangle about $[-\pi, \pi]$ of half-width β (more precisely of half-width $\beta' < \beta$ for β' arbitrarily close to β). The contributions from the end segments of the rectangle cancel by periodicity, and the accuracy bound then comes from estimating the integral over the sides of the rectangle, with imaginary parts $\pm\beta$. For our application, the interval is of length $2Ln^{1/3}$, so rescaling to length 2π gives an estimate of the integral along the sides of size $O(\exp(-a\pi n/n^{1/3}L))$, i.e., $O(\exp(-Cn^{2/3}))$ for some $C > 0$. This time the integrand is not periodic, so the contributions from the endpoints do not cancel, but the assumption on f implies that they too are of size $O(\exp(-Cn^{2/3}))$.

For the Gauss–Legendre and Clenshaw–Curtis formulas, the error estimate follows from Theorem 19.3 of [48], which is also proved by estimating contour integrals. We start with the integral over $[-Ln^{1/3}, Ln^{1/3}]$ of a function bounded and analytic in the strip around the real axis of half-width a. Rescaling to $[-1, 1]$ gives an integrand bounded and analytic in the strip of half-width $a/Ln^{1/3}$. This strip contains the

[8] See the discussion of sampling theory and approximation theory on page 2118 of [1].

Bernstein ρ-ellipse with foci ± 1 and parameter $\rho = 1 + a/Ln^{1/3}$ (the sum of the semiminor and semimajor axes). Theorem 19.3 of [48] now asserts that the errors are of size $O(\rho^{-n})$, i.e., $O(\exp(-na/Ln^{1/3}))$, that is, $O(\exp(-Cn^{2/3}))$ as required. ∎

Acknowledgments. I am grateful for advice and suggestions from four anonymous referees as well as from Jim Bremer, Ron DeVore, Simon Foucart, Walter Gautschi, Andrew Gibbs, Abi Gopal, Karlheinz Gröchenig, Nick Hale, Daan Huybrechs, Giuseppe Mastroianni, Giovanni Monegato, Yuji Nakatsukasa, Incoronata Notarangelo, Sotiris Notaris, Sheehan Olver, Allan Pinkus, Lothar Reichel, Kirill Serkh, Ian Sloan, Alex Townsend, and André Weideman. Particularly important contributions were made by Bremer and Hale in section 4 (Gauss) and Weideman in section 5 (Gauss–Hermite).

REFERENCES

[1] A. P. Austin and L. N. Trefethen, *Trigonometric interpolation and quadrature in perturbed points*, SIAM J. Numer. Anal., 55 (2017), pp. 2113–2122, https://doi.org/10.1137/16M1107760. (Cited on p. 147)

[2] N. S. Bakhvalov, *On the optimal speed of integrating analytic functions*, Comput. Math. Math. Phys., 7 (1967), pp. 63–75. (Cited on p. 138)

[3] I. Bogaert, *Iteration-free computation of Gauss–Legendre quadrature nodes and weights*, SIAM J. Sci. Comput., 36 (2014), pp. A1008–A1026, https://doi.org/10.1137/140954969. (Cited on p. 137)

[4] H. Brass and K. Petras, *Quadrature Theory: The Theory of Numerical Integration on a Compact Interval*, Amer. Math. Soc., 2011. (Cited on pp. 133, 137, 140, 146)

[5] J. Bremer, Z. Gimbutas, and V. Rokhlin, *A nonlinear optimization procedure for generalized Gaussian quadratures*, SIAM J. Sci. Comput., 32 (2010), pp. 1761–1788, https://doi.org/10.1137/080737046. (Cited on p. 139)

[6] C. W. Clenshaw and A. R. Curtis, *A method for numerical integration on an automatic computer*, Numer. Math., 2 (1960), pp. 197–205. (Cited on p. 135)

[7] R. Cools, *Constructing cubature formulae: The science behind the art*, Acta Numer., 6 (1997), pp. 1–54. (Cited on p. 145)

[8] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration*, Academic Press, 1975. (Cited on pp. 133, 134)

[9] R. DeVore, S. Foucart, G. Petrova, and P. Wojtaszczyk, *Computing a quantity of interest from observational data*, Constr. Approx., 49 (2019), pp. 461–508. (Cited on p. 147)

[10] J. Dick, F. Y. Kuo, and I. H. Sloan, *High-dimensional integration: The quasi-Monte Carlo way*, Acta Numer., 22 (2013), pp. 133–288. (Cited on p. 146)

[11] Z. Ditzian and V. Totik, *Moduli of Smoothness*, Springer Science & Business Media, 2012. (Cited on p. 138)

[12] T. A. Driscoll, N. Hale, and L. N. Trefethen, *Chebfun Guide*, Pafnuty Publications, Oxford, 2014; see also www.chebfun.org. (Cited on pp. 133, 135, 137)

[13] G. Evans, *Practical Numerical Integration*, Wiley, Chichester, UK, 1993. (Cited on p. 135)

[14] S. Foucart, *Mathematical Pictures at a Data Science Exhibition*, to appear. (Cited on p. 147)

[15] C. F. Gauss, *Methodus nova integralium valores per approximationem inveniendi*, Comment. Soc. Reg. Gotting. Recent., 1814 (also in Gauss's Werke, Vol. III, pp. 163–196). (Cited on p. 137)

[16] W. Gautschi, *A survey of Gauss–Christoffel quadrature formulae*, in EB Christoffel: The Influence of his Work on Mathematics and the Physical Sciences, P. L. Butzer and F. Fehér, eds., Birkhäuser, Basel, 1981, pp. 72–147. (Cited on p. 140)

[17] W. Gautschi, *How and how not to check Gaussian quadrature formulae*, BIT, 23 (1983), pp. 209–216. (Cited on p. 135)

[18] A. Gil, J. Segura, and N. M. Temme, *Fast, reliable and unrestricted iterative computation of Gauss–Hermite and Gauss–Laguerre quadratures*, Numer. Math., 143 (2019), pp. 649–682. (Cited on p. 140)

[19] G. Gourier, *Sur une méthode capable de fournir une valeur approchée de l'intégrale* $\int_{-\infty}^{\infty} F(x)dx$, C. R. Acad. Sci. Paris, 97 (1883), pp. 79–82. (Cited on p. 140)

[20] N. Hale and A. Townsend, *Fast and accurate computation of Gauss–Legendre and Gauss–Jacobi quadrature nodes and weights*, SIAM J. Sci. Comput., 35 (2013), pp. A652–A674, https://doi.org/10.1137/120889873. (Cited on p. 137)

[21] N. Hale and L. N. Trefethen, *New quadrature formulas from conformal maps*, SIAM J. Numer. Anal., 46 (2008), pp. 930–948, https://doi.org/10.1137/07068607X. (Cited on p. 138)
[22] D. Huybrechs and P. Opsomer, *Construction and implementation of asymptotic expansions for Laguerre-type orthogonal polynomials*, IMA J. Numer. Anal., 38 (2018), pp. 1085–1118. (Cited on p. 144)
[23] D. Kosloff and H. Tal-Ezer, *A modified Chebyshev pseudospectral method with an $O(N^{-1})$ time step restriction*, J. Comput. Phys., 104 (1993), pp. 457–469. (Cited on p. 138)
[24] P. K. Kythe and M. R. Schäferkotter, *Handbook of Computational Methods for Integration*, Chapman and Hall/CRC, Boca Raton, FL, 2005. (Cited on p. 135)
[25] H. J. Landau and H. O. Pollak, *Prolate spheroidal wave functions, Fourier analysis and uncertainty*—II, Bell Syst. Tech. J., 40 (1961), pp. 65–84. (Cited on p. 139)
[26] H. J. Landau and H. O. Pollak, *Prolate spheroidal wave functions, Fourier analysis and uncertainty*—III*: The dimension of the space of essentially time- and band-limited signals*, Bell Syst. Tech. J., 41 (1962), pp. 1295–1336. (Cited on p. 139)
[27] H. J. Landau and H. Widom, *Eigenvalue distribution of time and frequency limiting*, J. Math. Anal. Appl., 77 (1980), pp. 469–481. (Cited on p. 139)
[28] D. S. Lubinsky, *A Survey of Weighted Approximation for Exponential Weights*, preprint, https://arxiv.org/abs/math/0701099, 2007. (Cited on p. 144)
[29] G. Mastroianni and G. Monegato, *Truncated Gauss–Laguerre quadrature rules*, in Recent Trends in Numerical Analysis, D. Trigiante, ed., Nova Science, Hauppauge, NY, 2000, pp. 213–221. (Cited on pp. 143, 144)
[30] G. Mastroianni and G. Monegato, *Truncated quadrature rules over $(0,\infty)$ and Nyström-type methods*, SIAM J. Numer. Anal., 41 (2003), pp. 1870–1892, https://doi.org/10.1137/S0036142901391475. (Cited on pp. 143, 144)
[31] J. C. Maxwell, *On approximate multiple integration between limits of summation*, Proc. Cambridge Philos. Soc., 3 (1877), pp. 39–47. (Cited on p. 145)
[32] D. J. Newman and T. J. Rivlin, *Approximation of monomials by lower degree polynomials*, Aequationes Math., 14 (1976), pp. 451–455. (Cited on p. 135)
[33] H. O'Hara and F. J. Smith, *Error estimation in the Clenshaw–Curtis quadrature formula*, Comput. J., 11 (1968), pp. 213–219. (Cited on pp. 135, 136)
[34] S. Olver, R. M. Slevinsky, and A. Townsend, *Fast algorithms using orthogonal polynomials*, Acta Numer., 29 (2020), pp. 573–699; see also github.com/JuliaApproximation/ApproxFun.jl. (Cited on p. 133)
[35] A. Osipov and V. Rokhlin, *On the evaluation of prolate spheroidal wave functions and associated quadrature rules*, Appl. Comput. Harmon. Anal., 36 (2014), pp. 108–142. (Cited on p. 139)
[36] A. Osipov, V. Rokhlin, and H. Xiao, *Prolate Spheroidal Wavefunctions of Order Zero: Mathematical Tools for Bandlimited Approximation*, Springer, 2013. (Cited on p. 139)
[37] J. Ouspensky, *Sur les valeurs asymptotiques des coefficients de Cotes*, Bull. Amer. Math. Soc., 31 (1925), pp. 145–156. (Cited on p. 134)
[38] K. Petras, *Gaussian versus optimal integration of analytic functions*, Constr. Approx., 14 (1998), pp. 231–245. (Cited on p. 137)
[39] G. Pólya, *Über die Konvergenz von Quadraturverfahren*, Math. Z., 37 (1933), pp. 264–286. (Cited on p. 134)
[40] C. Runge, *Über empirische Funktionen und die Interpolation zwischen äquidistanten Ordinaten*, Z. Math. Phys., 46 (1901), pp. 224–243. (Cited on p. 133)
[41] D. Slepian and H. O. Pollak, *Prolate spheroidal wave functions, Fourier analysis and uncertainty*—I, Bell Syst. Tech. J., 40 (1961), pp. 43–63. (Cited on p. 139)
[42] G. Szegő, *Orthogonal Polynomials*, 4th ed., Amer. Math. Soc., 1975. (Cited on p. 140)
[43] V. Tchakaloff, *Formules de cubatures mécaniques à coefficients non négatifs*, Bull. Sci. Math. (2), 81 (1957), pp. 123–134. (Cited on p. 145)
[44] A. Townsend, T. Trogdon, and S. Olver, *Fast computation of Gauss quadrature nodes and weights on the whole real line*, IMA J. Numer. Anal., 36 (2016), pp. 337–358. (Cited on pp. 140, 143)
[45] L. N. Trefethen, *Is Gauss quadrature better than Clenshaw–Curtis?*, SIAM Rev., 50 (2008), pp. 67–87, https://doi.org/10.1137/060659831. (Cited on pp. 135, 137)
[46] L. N. Trefethen, *Cubature, approximation, and isotropy in the hypercube*, SIAM Rev., 59 (2017), pp. 469–491, https://doi.org/10.1137/16M1066312. (Cited on pp. 139, 145)
[47] L. N. Trefethen, *Multivariate polynomial approximation in the hypercube*, Proc. Amer. Math. Soc., 145 (2017), pp. 4837–4844. (Cited on p. 145)
[48] L. N. Trefethen, *Approximation Theory and Approximation Practice, Extended Edition*, SIAM, 2019, https://doi.org/10.1137/1.9781611975949. (Cited on pp. 135, 147, 148)

[49] L. N. Trefethen and J. A. C. Weideman, *The exponentially convergent trapezoidal rule*, SIAM Rev., 56 (2014), pp. 385–458, https://doi.org/10.1137/130932132. (Cited on pp. 142, 145, 147)
[50] J. Waldvogel, *Fast construction of the Fejér and Clenshaw–Curtis quadrature rules*, BIT Numer. Math., 46 (2006), pp. 195–202. (Cited on p. 135)
[51] J. A. C. Weideman, *Gauss–Hermite vs. trapezoidal rule*, talk at SIAM Annual Meeting, Portland, OR, 2008. (Cited on p. 143)
[52] J. A. C. Weideman and L. N. Trefethen, *The kink phenomenon in Fejér and Clenshaw–Curtis quadrature*, Numer. Math., 107 (2007), pp. 707–727. (Cited on p. 137)
[53] 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, https://doi.org/10.1137/120869845. (Cited on p. 137)
[54] H. Xiao, V. Rokhlin, and N. Yarvin, *Prolate spheroidal wavefunctions, quadrature and interpolation*, Inverse Problems, 17 (2001), pp. 805–828. (Cited on p. 139)