

A trapezoidal rule error bound unifying the Euler–Maclaurin formula and geometric convergence for periodic functions¹

Mohsin Javed² and Lloyd N. Trefethen³
Mathematical Institute, Oxford University

Abstract

The error in the trapezoidal rule quadrature formula can be attributed to discretization in the interior and non-periodicity at the boundary. Using a contour integral, we derive a unified bound for the combined error from both sources for analytic integrands. The bound gives the Euler–Maclaurin formula in one limit and the geometric convergence of the trapezoidal rule for periodic analytic functions in another.

Keywords: trapezoidal rule, Euler–Maclaurin formula

1 Introduction

Let f be continuous on $[0, 1]$ and let n be a positive integer. The (composite) trapezoidal rule approximates the integral

$$I = \int_0^1 f(x) dx \tag{1}$$

by the sum

$$I_n = n^{-1} \sum_{k=0}^n {}' f(k/n), \tag{2}$$

where the prime indicates that the terms $k = 0$ and $k = n$ are multiplied by $1/2$. Throughout this paper, f may be real or complex, and “periodic” means periodic with period 1.

The approximation of I by I_n has many interesting properties. One is that if f is periodic and analytic, the convergence is geometric. This observation in some sense goes back to Poisson in the 1820s [9], though it

¹Supported 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.

²javed@maths.ox.ac.uk

³trefethen@maths.ox.ac.uk

seems to have been Davis in 1959 who first stated a theorem [2, 16]. Another is that for non-periodic f , the accuracy is $O(n^{-2})$, and this can be improved to $O(n^{-4})$, $O(n^{-6})$, and so on by subtracting appropriate multiples of $f'(1) - f'(0)$, $f'''(1) - f'''(0)$, and so on, if f is sufficiently smooth. The latter process is described by the Euler–Maclaurin formula, published independently by Euler and Maclaurin around 1740 [3, 7].

A standard derivation of Davis’s result involves contour integrals in the complex plane, and contour integrals can also be used to derive the Bernoulli numbers that appear in the Euler–Maclaurin formula. With these facts in mind, we have attempted to develop a unified formulation based on contour integrals that would make it possible to derive both kinds of results at once for analytic integrands. It is hard to believe that our theorem can be new, but we have been unable to find such a result in the literature. The closest we have found is [13, Appendix B], which sets up the problem in the same way without deriving an explicit estimate.

The theorem is stated in Section 2. Various existing results are derived as corollaries in Section 3, and the proof of the theorem is presented in Section 4. Section 5 mentions the variation of the midpoint rather than trapezoidal formula, and the discussion in Section 6 points to connections to rational approximation and the theory of hyperfunctions.

2 Theorem

The theorem is stated in terms of the following Euler–Maclaurin correction sum. For any $m \geq 0$ and sufficiently smooth f , we define

$$Q_{m,n} = \sum_{\substack{k=1 \\ k \text{ odd}}}^m \frac{f^{(k)}(1) - f^{(k)}(0)}{n^{k+1}(k+1)!} B_{k+1}, \quad (3)$$

where B_k is the k th Bernoulli number ($B_2 = 1/6$, $B_4 = -1/30$, $B_6 = 1/42, \dots$). We further define

$$\Delta_m = \sum_{\substack{k=1 \\ k \text{ odd}}}^m |f^{(k)}(1) - f^{(k)}(0)| \quad (4)$$

and

$$\Delta^{(m+1)} = \sup_{-a < y < a} |f^{(m+1)}(1 + iy) - f^{(m+1)}(iy)|. \quad (5)$$

Note that $Q_{m,n}$, Δ_m , and $\Delta^{(m+1)}$ all depend on f , though the notation does not make this explicit.

Here is the theorem. The region of analyticity is sketched in Figure 1 (Section 4).

THEOREM 1. *Given real numbers $a > 0$ and $M \geq 0$ and an integer $m \geq 0$, let f satisfy $|f(z)| \leq M$ and have a continuous $(m+1)$ st derivative in the region defined by $0 \leq \operatorname{Re} z \leq 1$, $-a < \operatorname{Im} z < a$, and be analytic in the interior of this region. Then*

$$I_n - I - Q_{m,n} = E_{\text{interior}} + E_{\text{boundary}} + E_{\text{tail}} \quad (6)$$

with

$$|E_{\text{interior}}| \leq \frac{2M}{e^{2\pi an} - 1}, \quad (7)$$

$$|E_{\text{boundary}}| \leq \frac{\Delta^{(m+1)}}{3\pi^m (2n)^{m+2}}, \quad (8)$$

and

$$|E_{\text{tail}}| \leq \frac{\Delta_m (2\pi na + 1)^m}{\pi^2 e^{2\pi na}}. \quad (9)$$

In words, Theorem 1 breaks the error of the trapezoidal rule adjusted by $Q_{m,n}$ into three terms, the first related to discretization error in the interior and the others to boundary effects. Two of these are exponentially small as $n \rightarrow \infty$, and the other is algebraically small. We use the labels “boundary” and “tail” for reasons that will become evident in the proof.

Although Theorem 1 is valid for any $m \geq 0$, one would not normally apply it for odd values of m , since $Q_{m,n} = Q_{m+1,n}$ and $\Delta_m = \Delta_{m+1}$ when m is odd. This means that if m is odd, then increasing it to the next even number yields the same bound except with $O(n^{-m-2}\Delta^{(m+1)})$ improved to $O(n^{-m-3}\Delta^{(m+2)})$, assuming $f^{(m+2)}$ exists and is continuous.

3 Corollaries

By considering special cases of Theorem 1, we obtain various familiar results. The first is Davis’s theorem for periodic integrands; see [2] and [16, §4].

COROLLARY 1. *Let f be analytic and 1-periodic with $|f(z)| \leq M$ in the region $0 \leq \operatorname{Re} z \leq 1$, $-a < \operatorname{Im} z < a$ for some $a > 0$. Then*

$$|I_n - I| \leq \frac{2M}{e^{2\pi an} - 1}. \quad (10)$$

Proof. By (3)–(5), $Q_{m,n}$, Δ_m , and $\Delta^{(m+1)}$ are zero when f is periodic. It follows from (8) and (9) that E_{boundary} and E_{tail} are zero in this case too. The bound (10) now follows from (6) and (7). ■

The second corollary is one version of the Euler–Maclaurin formula.

COROLLARY 2. *Let f be analytic on $[0, 1]$. Then for any $m \geq 0$,*

$$I_n - I - Q_{m,n} = O(n^{-m-2}) \quad (11)$$

as $n \rightarrow \infty$.

Proof. If f is analytic on $[0, 1]$, it is analytic and bounded in the strip around $[0, 1]$ of half-width a for some $a > 0$. The result now follows from (6)–(9) since as $n \rightarrow \infty$, $E_{\text{boundary}} = O(n^{-m-2})$ and both E_{interior} and E_{tail} are of asymptotically smaller order, $O(n^m e^{-2\pi an})$. ■

If f is a polynomial of degree at most $m + 1$, the m th Euler–Maclaurin approximation is exact.

COROLLARY 3. *Let f be a polynomial of degree at most $m + 1$ for some $m \geq 0$. Then for any n ,*

$$I_n = I + Q_{m,n}. \quad (12)$$

Proof. If f is a polynomial of degree at most $m + 1$, then $f^{(m+1)}$ is a constant, implying $\Delta^{(m+1)} = 0$, so (8) implies that $E_{\text{boundary}} = 0$ in (6). The bounds for the other terms E_{interior} and E_{tail} contain the factor $e^{-2\pi an}$, where n is fixed but a can be taken as large as we want since a polynomial is an entire function. In the case of E_{tail} , Δ_m is fixed, so (9) implies that E_{tail} becomes arbitrarily small as $a \rightarrow \infty$. In the case of E_{interior} , M grows as $a \rightarrow \infty$, but only at a polynomial rate, so (7) implies that this term too becomes arbitrarily small as $a \rightarrow \infty$. Thus $I_n - I - Q_{m,n}$ must be equal to zero. ■

Corollary 3 leads to the identity known as *Faulhaber’s formula*.

COROLLARY 4. *For any $n \geq 1$ and $m \geq 1$,*

$$\sum_{k=1}^n k^m = \frac{n^{m+1}}{m+1} + \frac{n^m}{2} + \sum_{\substack{k=1 \\ k \text{ odd}}}^{m-1} \binom{m}{k} \frac{B_{k+1}}{k+1} n^{m-k}. \quad (13)$$

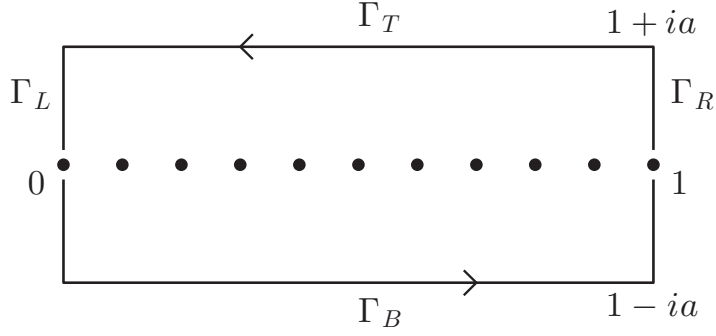


Figure 1: *Contour of integration for the proof of Theorem 1. At $z = 0$ and 1 the integral is defined in the principal value sense, as suggested by the gaps in the contour. The dots show the sample points of the trapezoidal rule, which become the poles of the characteristic function $(2i)^{-1} \cot(\pi n z)$. The theorem bounds the error by the sum of three terms: E_{interior} corresponding to Γ_T and Γ_B , E_{boundary} corresponding to Γ_L and Γ_R , and E_{tail} associated with extending the integrals along Γ_L and Γ_R to infinite intervals.*

Proof. This is equation (12) in the special case $f(x) = n^{m+1}x^m$; the term $n^{m+1}/(m+1)$ is the integral I , and the term $n^m/2$ appears because I_n is defined in (2) with a factor $1/2$ multiplying the term $k = n$. The reason for the assumption $m \geq 1$ is that in the (trivial) case $m = 0$, the sum on the left is missing a nonzero contribution $1/2$ corresponding to the $k = 0$ term in the trapezoidal sum. ■

4 Proof

We now prove Theorem 1.

As sketched in Figure 1, let Γ be the boundary of the rectangle of analyticity of f , oriented in the positive sense, and let Γ_L , Γ_R , Γ_T , and Γ_B be the left, right, top and bottom boundary segments. In the following argument we suppose for simplicity that f extends continuously to Γ_T and Γ_B . If it does not, then the required result can be obtained by replacing Γ_T by $\Gamma_T - \varepsilon i$ and Γ_B by $\Gamma_B + \varepsilon i$ and considering $\varepsilon \rightarrow 0$.

The “characteristic function” $(2i)^{-1} \cot(\pi n z)$ has simple poles at $z = k/n$ for each integer k with residue $(2\pi i n)^{-1}$. It follows from residue calculus that the trapezoidal approximation (2) can be represented by the contour integral

$$I_n = \int_{\Gamma} f(z) (2i)^{-1} \cot(\pi n z) dz, \quad (14)$$

where the integrals over Γ_L and Γ_R are taken in the principal value sense so

as to introduce the necessary factors of $1/2$.

The true integral I can also be represented by a contour integral over Γ :

$$I = \int_{\Gamma} f(z)u(z)dz, \quad (15)$$

where u is defined by

$$u(z) = \begin{cases} -1/2 & \text{Im} z > 0, \\ +1/2 & \text{Im} z < 0. \end{cases} \quad (16)$$

We can derive this formula by noting that I can be regarded as half the integral of f from $0 - 0i$ to $1 - 0i$ minus half the integral of f from $1 + 0i$ to $0 + 0i$. Since f is analytic in the rectangular region and u is analytic in the upper and lower half-planes, these two contours of integration can be deformed to the upper and lower halves of Γ without changing the value of the integral.

Combining (14)–(16), we find

$$I_n - I = \int_{\Gamma} f(z)S(z)dz, \quad (17)$$

with $S(z) = (2i)^{-1} \cot(\pi n z) - u(z)$, which simplifies to

$$S(z) = \begin{cases} \frac{1}{1 - e^{-2i\pi n z}}, & \text{Im} z > 0, \\ \frac{1}{e^{2i\pi n z} - 1}, & \text{Im} z < 0. \end{cases} \quad (18)$$

We now establish (6) by breaking (17) into four terms,

$$I_n - I = I_L + I_R + I_T + I_B, \quad (19)$$

corresponding to the integrals of $f(z)S(z)$ over the four segments of the boundary (Figure 1). We note immediately that by (17) and (18), I_T and I_B are each bounded by $M/(e^{2\pi a n} - 1)$, which implies the bound (7) with E_{interior} defined by

$$E_{\text{interior}} = I_T + I_B. \quad (20)$$

To complete the derivation of (6)–(9), by (6) and (19) and (20), we must show that $I_L + I_R$ can be broken into the pieces

$$I_L + I_R = Q_{m,n} + E_{\text{boundary}} + E_{\text{tail}} \quad (21)$$

with E_{boundary} and E_{tail} satisfying (8) and (9).

For $y \in (-a, a)$, define

$$g(y) = f(1 + iy) - f(iy). \quad (22)$$

This definition simplifies (4) and (5) to

$$\Delta_m = \sum_{\substack{k=1 \\ k \text{ odd}}}^m |g^{(k)}(0)| \quad (23)$$

and

$$\Delta^{(m+1)} = \sup_{-a < y < a} |g^{(m+1)}(y)|. \quad (24)$$

Since $S(1 + iy) = S(iy)$, our task is to estimate

$$I_L + I_R = i \int_{-a}^a g(y) S(iy) dy, \quad (25)$$

where the integral is taken in the principal value sense. Since g has a continuous $(m + 1)$ st derivative on $(-a, a)$, one of the standard forms of Taylor's theorem with remainder gives

$$\left| g(y) - \sum_{k=0}^m \frac{g^{(k)}(0)}{k!} y^k \right| \leq \frac{\Delta^{(m+1)}}{(m+1)!} |y|^{m+1} \quad (26)$$

for $y \in (-a, a)$ [6, Theorem 1.36]. In (25), we note that $S(iy)$ is an odd function of y . Therefore when the sum on the left-hand side of (26) is inserted in (25) as an approximation to $g(y)$, the contributions from even values of k vanish (in the case $k = 0$, we use the fact that it is a principal value integral). The result is

$$\left| I_L + I_R + 2i \sum_{\substack{k=1 \\ k \text{ odd}}}^m \frac{g^{(k)}(0)}{k!} \int_0^a \frac{y^k}{e^{2\pi ny} - 1} dy \right| \leq \frac{2\Delta^{(m+1)}}{(m+1)!} \int_0^a \frac{y^{m+1}}{e^{2\pi ny} - 1} dy \quad (27)$$

since $S(iy) = -1/(\exp(2\pi ny) - 1)$ for $y > 0$.

We can now identify the pieces $Q_{m,n}$, E_{boundary} , and E_{tail} . The quantity E_{boundary} is the number inside absolute value signs on the left of (27), satisfying

$$|E_{\text{boundary}}| \leq \frac{2\Delta^{(m+1)}}{(m+1)!} \int_0^a \frac{y^{m+1}}{e^{2\pi ny} - 1} dy. \quad (28)$$

The quantity $Q_{m,n}$ is the negative of the sum on the left in (27), but with the integrals running from 0 to ∞ instead of a :

$$Q_{m,n} = -2i \sum_{\substack{k=1 \\ k \text{ odd}}}^m \frac{g^{(k)}(0)}{k!} \int_0^\infty \frac{y^k}{e^{2\pi ny} - 1} dy. \quad (29)$$

And E_{tail} is the error just introduced in extending those integrals to ∞ :

$$E_{\text{tail}} = -2i \sum_{\substack{k=1 \\ k \text{ odd}}}^m \frac{g^{(k)}(0)}{k!} \int_a^\infty \frac{y^k}{e^{2\pi ny} - 1} dy. \quad (30)$$

These definitions ensure that (21) holds as required; what remains is to derive (8) from (28) and (9) from (30) and to confirm that (29) matches the definition of $Q_{m,n}$ given in (3).

That (29) matches (3) follows from an identity that goes back to the late 19th century [1, 10],

$$\int_0^\infty \frac{u^k du}{1 - e^{2\pi u}} = i^{k+1} \frac{B_{k+1}}{2k+2}, \quad k \text{ odd}, \quad (31)$$

since $g^{(k)}(0) = i^k (f^{(k)}(1) - f^{(k)}(0))$ and $i^{2k+2} = 1$ for k odd.

To derive (8) from (28) we note that the change of variables $t = 2\pi ny$ and extension of the upper limit of integration to ∞ in (28) gives

$$|E_{\text{boundary}}| \leq \frac{2\Delta^{(m+1)}}{(2\pi n)^{m+2}} \frac{1}{(m+1)!} \int_0^\infty \frac{t^{m+1}}{e^t - 1} dt. \quad (32)$$

The bound (8) follows from this together with an identity closely related to (31) [8, (25.5.1)],

$$\frac{1}{(m+1)!} \int_0^\infty \frac{t^{m+1}}{e^t - 1} dt = \zeta(m+2) \quad (m \geq 0), \quad (33)$$

where ζ is the Riemann zeta function. The numbers $\zeta(m+2)$ decrease monotonically from their maximum $\zeta(2) = \pi^2/6$ in the case $m = 0$.

To derive (9) from (30), we note that by (23) and (30),

$$|E_{\text{tail}}| \leq 2\Delta_m \max_{\substack{1 \leq k \leq m \\ k \text{ odd}}} \frac{1}{k!} \int_a^\infty \frac{y^k}{e^{2\pi ny} - 1} dy, \quad (34)$$

or after the change of variables $t = 2\pi ny$,

$$|E_{\text{tail}}| \leq 2\Delta_m \max_{\substack{1 \leq k \leq m \\ k \text{ odd}}} \frac{1}{k!(2\pi n)^{k+1}} \int_{2\pi na}^\infty \frac{t^k}{e^t - 1} dt, \quad (35)$$

Applying Lemma 1 of the appendix with $b = 2\pi na$ and using the inequalities $(2\pi n)^{-k-1} \leq (2\pi)^{-2}$ and $(b+1)^k \leq (b+1)^m$, we obtain (9). ■

5 Midpoint rule variant

A close cousin of the trapezoidal rule (2) is the *midpoint rule*,

$$\tilde{I}_n = n^{-1} \sum_{k=1}^n f((k - \frac{1}{2})/n), \quad (36)$$

where now no prime is needed in the sum since all the terms have exactly the same weight. All the arguments of this paper go through for this case with very little change. The term $\cot(\pi n z)$ in (14) becomes $-\tan(\pi n z)$, and the correction sum $Q_{m,n}$ is adjusted to a new sum $\tilde{Q}_{m,n}$. The final result is a bound very close to Theorem 1; details will be reported elsewhere.

6 Discussion

In the theory of *hyperfunctions*, delta functions and other distributions are realized not by the test functions and linear functionals of real analysis, but by methods of complex analysis. Specifically, a distribution on a real interval is defined as a difference of analytic functions in the upper and lower half-planes, or more precisely, an equivalence class of such differences [4, 11, 12]. Our arguments have exactly this flavor, and in particular, the function $S(z)$ is expressed in (18) in hyperfunction form. The reason hyperfunction theory is relevant is that it provides a convenient framework in which to compare the integral I with the trapezoidal approximation I_n , which is regarded essentially as an integral whose integrand is a string of delta functions. The present paper is a contribution toward a longer-term goal of strengthening the links between hyperfunction theory and numerical analysis.

Going beyond the trapezoidal rule, it may be noted that whenever an integral I of an analytic function f is approximated by a quadrature formula $I_n = \sum_{k=0}^n w_k f(x_k)$ defined by nodes $\{x_k\}$ and weights $\{w_k\}$, I_n can be written as a contour integral involving the product $r(z)f(z)$, where r is the type $(n, n+1)$ rational function with poles x_k and residues w_k . Writing I itself as a contour integral of f times a hyperfunction such as $u(z)$ in (16) makes it possible to estimate $I_n - I$ by contour integrals. This technique was pioneered by Takahasi and Mori [14], who had the vision of connecting numerical analysis and hyperfunctions long before we did, and it was applied to the comparison of Gauss and Clenshaw–Curtis quadrature formulas in [15]. Such analyses highlight the fact that every quadrature formula implicitly makes use of a rational approximation, and the properties of these rational approximations are investigated in [16, §14] and [17].

Appendix

The following inequality was used in the proof of Section 4.

LEMMA 1. *For any real number $b \geq 0$ and integer $k \geq 1$,*

$$\frac{1}{k!} \int_b^\infty \frac{t^k}{e^t - 1} dt \leq 2(b+1)^k e^{-b}. \quad (37)$$

Proof. As observed below (33), the left-hand side of (37) is $\leq \pi^2/6 \approx 1.64$, whereas it can be verified that the right-hand side is greater than this value for $b \leq 0.75$. To complete the proof we may accordingly assume $b > 0.75$. Since $e^{0.75} > 2$, replacing the denominator in (37) by e^t decreases the integral by less than a factor of 2, so it is enough to show

$$\frac{1}{k!} \int_b^\infty \frac{t^k}{e^t} dt \leq (b+1)^k e^{-b} \quad (38)$$

for $b > 0.75$. We can do this by induction in k . For $k = 1$, the inequality holds as an equality, as can be verified by integration by parts. Assume then that it holds for some $k \geq 1$ and consider the case $k + 1$. Integration by parts gives

$$\begin{aligned} \frac{1}{(k+1)!} \int_b^\infty \frac{t^{k+1}}{e^t} dt &= \frac{-1}{(k+1)!} t^{k+1} e^{-t} \Big|_b^\infty + \frac{1}{k!} \int_b^\infty \frac{t^k}{e^t} dt \\ &\leq \frac{1}{(k+1)!} b^{k+1} e^{-b} + (b+1)^k e^{-b} \end{aligned}$$

by the inductive hypothesis. Cancelling the common factor of e^{-b} , this leaves us with the problem of establishing

$$\frac{1}{(k+1)!} b^{k+1} + (b+1)^k \leq (b+1)^{k+1},$$

which follows since the left-hand side is less than $b^{k+1} + (b+1)^k$ and the right-hand side is equal to $b(b+1)^k + (b+1)^k$. ■

Acknowledgments

We have benefited from good advice from Endre Süli and André Weideman.

References

- [1] Carda, K. 1894 Zur Darstellung der Bernoullischen Zahlen durch bestimmte Integrale. *Monatshefte für Mathematik* **5**, 321–324.
- [2] Davis, P. J. 1959 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, pp. 45–59.
- [3] Euler, L. 1736 Methodus universalis series summandi ulterius promota. *Comm. Acad. Sci. Imp. Petrop.* **8**, 147–158.
- [4] Graf, U. 2000 *Introduction to Hyperfunctions and Their Integral Transforms: An Applied and Computational Approach*. Birkhäuser. (doi:10.1007/978-3-0346-0408-6.)
- [5] Hale, N., Higham, N. J. & Trefethen, L. N. 2008 Computing A^α , $\log(A)$, and related matrix functions by contour integrals. *SIAM J. Numer. Anal.* **46**, 2505–2523. (doi:10.1137/070700607.)
- [6] Knapp, A. W. 2005 *Basic Real Analysis*. Birkhäuser.
- [7] Maclaurin, C. 1742 *A Treatise of Fluxions*, v. II. Edinburgh: Ruddiman.
- [8] Olver, F. W. J., Lozier, D. W., Boisvert, R. F. & Clark, C. W., eds. 2010 *NIST Handbook of Mathematical Functions*. Cambridge U. Press.
- [9] Poisson, S.-D. 1827 Sur le calcul numérique des intégrales définies. *Mémoires de l'Académie Royale des Sciences de l'Institut de France* **4**, 571–602.
- [10] Saalschütz, L. 1893 *Vorlesungen über die Bernoullischen Zahlen*. Berlin: Springer.
- [11] Sato, M. 1959 Theory of hyperfunctions. I. *J. Fac. Sci. Univ. Tokyo Sect. I* **8**, 139–193.
- [12] Sato, M. 1960 Theory of hyperfunctions. II. *J. Fac. Sci. Univ. Tokyo Sect. I* **8**, 387–437.
- [13] Schwartz, C. 1969 Numerical integration of analytic functions. *J. Comp. Phys.* **4**, 19–29.

- [14] Takahasi, H. & Mori, M. 1971 Estimation of errors in the numerical quadrature of analytic functions. *Applicable Anal.* **1**, 201–229.
- [15] Trefethen, L. N. 2008 Is Gauss quadrature better than Clenshaw–Curtis? *SIAM Rev.* **50**, 67–87. (doi:10.1137/060659831.)
- [16] Trefethen, L. N. & Weideman, J. A. C. 2013 The exponentially convergent trapezoidal rule. *SIAM Rev.*, submitted.
- [17] Trefethen, L. N., Weideman, J. A. C. & Schmelzer, T. 2006 Talbot quadratures and rational approximations. *BIT Numer. Math.* **46**, 653–670. (doi:10.1007/s10543-006-0077-9.)