Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

On the Efficient Computation of High-Dimensional Integrals and the Approximation by Exponential Sums

(revised version: November 2009)

by

Dietrich Braess, and Wolfgang Hackbusch

Preprint no.: 3 2009

On the efficient computation of high-dimensional integrals and the approximation by exponential sums

Dietrich Braess and Wolfgang Hackbusch

Abstract The approximation of the functions $1/x$ and $1/\sqrt{x}$ by exponential sums enables us to evaluate some high-dimensional integrals by products of one-dimensional integrals. The degree of approximation can be estimated via the study of rational approximation of the square root function. The latter has interesting connections with the Babylonian method and Gauss' arithmetic-geometric process.

Key words: exponential sums, rational functions, Chebyshev approximation, best approximation, completely monotone functions, Heron's algorithm, complete elliptic integrals, Landen transformation.

AMS Subject Classifications: 11L07, 41A20.

1 Introduction

The approximation of transcendental or other complicated functions by polynomials, rational functions, and spline functions is at the centre of classical approximation theory. In the last decade the numerical solution of partial differential equations gave rise to quite different problems in approximation theory. In this paper we will study the approximation of $x^{-\alpha}$ ($\alpha = 1/2$ or 1) by exponential sums. Here a simple function is approximated by a more complicated one, but it enables the fast com-

Dietrich Braess
Mathematisches Institut, Ruhr-Universität Bochum, 44780 Bochum, Germany,
e-mail: Dietrich.Braess@rub.de,
URL: http://homepage.ruhr-uni-bochum.de/Dietrich.Braess

Wolfgang Hackbusch
Max-Planck-Institut *Mathematik in den Naturwissenschaften*, Inselstr. 22, 04103 Leipzig, Germany, e-mail: wh@mis.mpg.de,
URL: http://www.mis.mpg.de/scicomp/hackbusch_e.html

R.A. DeVore, A. Kunoth (eds.), *Multiscale, Nonlinear and Adaptive Approximation*,
DOI 10.1007/978-3-642-03413-8_3,

putation of some high-dimensional integrals which occur in quantum physics and quantum chemistry.

A model example is an integral of the form

$$\int \frac{g_1(x_1)\cdots g_d(x_d)}{\|x-y\|_0}\mathrm{d}x \tag{1}$$

on a domain in $\mathbb{R}^d$, where $\|\cdot\|_0$ refers to the Euclidean norm. When we insert the approximation

$$\frac{1}{\sqrt{x}} \approx \sum_{j=1}^{n} \alpha_j \mathrm{e}^{-t_j x},$$

then the integral is reduced to a sum of products of one-dimensional integrals

$$\sum_{j=1}^{n} \alpha_j \prod_{i=1}^{d} \int g_i(x_i)\exp[-t_j(x_i-y_i)^2]\,\mathrm{d}x_i,$$

and a fast computation is now possible (at least in the domain, where the approximation is valid, see Section 6.2 for more details). Other examples will be discussed in Sections 5 and 6.

There are only a few problems in nonlinear approximation theory for which the degree of approximation can be estimated. Surprisingly, the problem under consideration belongs to those rare cases. The functions $x^{-\alpha}$ ($\alpha > 0$) are monosplines for the kernel e^{-tx}. For this reason, results for the rational approximation of the square root function provide good estimates for the degree of approximation by exponential sums.

In principle, rational approximation of the square root function is well known for more than a century from Zolotarov's results. Elliptic integrals play a central role in his investigations. We find it more interesting, direct, and less technical to derive approximation properties from the Babylonian method for the computation of a square root. Gauss' arithmetic-geometric process yields a fast computation of the decay rate of the approximation error.

The rest of the paper is organised as follows. Section 2 is devoted to the connection of the approximation of $x^{-\alpha}$ by exponential sums with the rational approximation of $\sqrt{x}$. The investigation of the latter with the help of the Babylonian method and Gauss' arithmetic-geometric mean is the main purpose of Section 3. The results for the approximation of $1/x$ by exponential sums on finite and infinite intervals are presented in Section 4. Numerical results show that the theory yields the correct asymptotic law, while an improvement is possible for infinite intervals. The role of the approximation problem for the computation of high-dimensional integrals is elucidated with several examples in Sections 5 and 6. The numerical computation of the best exponential sums is discussed in Section 7. Appendix A provides auxiliary results for small intervals. Properties of complete elliptic integrals that are required for the derivation of the asymptotic rates, are derived in Appendix B. Finally it is shown in Appendix C that a competing tool yields the same law for infinite intervals.

2 Approximation of completely monotone functions by exponential sums

Good estimates for the degree of approximation are available only for a few classes of nonlinear approximation problems. Fortunately, the asymptotic behaviour is known for the functions in which we are interested. The functions $1/x$ and $1/\sqrt{x}$ are *completely monotone* for $x > 0$, i.e., they are Laplace transforms of non-negative measures:

$$f(x) = \int_0^\infty \mathrm{e}^{-tx} \mathrm{d}\mu(t), \qquad \mathrm{d}\mu \geq 0.$$

In particular,

$$\frac{1}{x} = \int_0^\infty \mathrm{e}^{-tx} \mathrm{d}t, \quad \frac{1}{\sqrt{x}} = \int_0^\infty \mathrm{e}^{-tx} \frac{\mathrm{d}t}{\sqrt{\pi t}}.$$

In order to avoid degeneracies, we assume that the support of the measure is an infinite set. We will also restrict ourselves to $\Re e\, x \geq 0$.

We consider best approximations in the sense of Chebyshev, i.e., the error is to be minimised with respect to the supremum norm on a given interval. A unique best exponential sum of order n,

$$u_n(x) = \sum_{\nu=1}^{n} \alpha_\nu \mathrm{e}^{-t_\nu x} \tag{2}$$

exists for a given completely monotone function f, while this is not true for arbitrary continuous functions. Moreover, the coefficients α_ν in the best approximation are non-negative (cf. [4]).

Our error estimates require the solution of a nonlinear interpolation problem that is also solvable for completely monotone functions.

Theorem 2.1. *Let f be completely monotone for $x > 0$ and $0 < x_1 < x_2 < \ldots < x_{2n}$. Then there exists an exponential sum u_n such that*

$$u_n(x_j) = f(x_j), \qquad j = 1, 2, \ldots, 2n.$$

Moreover

$$u_n(x) < f(x) \qquad \text{for } 0 < x < x_1 \text{ and } x > x_{2n}.$$

If in addition f is continuous at $x = 0$, also $u_n(0) < f(0)$ holds.

Sketch of proof. The complete monotonicity allows us to apply deformation arguments from nonlinear analysis. The best approximation u_n on the interval $[\frac{1}{2}x_1, x_{2n} + 1]$ has an alternant of length $2n+1$ (see Definition 3.1). Therefore, $f - u_n$ has $2n$ zeros $y_1 < y_2 < \ldots, y_{2n}$. Set

$$x_j(s) := (1-s)y_j + sx_j, \qquad 0 \leq s \leq 1, \quad j = 1, 2, \ldots, 2n.$$

The set of numbers $s \in [0, 1]$ for which the interpolation at the points $x_j(s)$ is solvable, contains the point $s = 0$. The rules on the zeros of extended exponential sums

$\sum_j(\alpha_j+\beta_j x)\mathrm{e}^{-t_j x}$ and the Newton method imply that the set is open. It follows from compactness properties of exponential sums that the set is also closed. Hence, the value $s=1$ is included.

The given function f and the approximating exponential sums are analytic functions in the right half plane of $\mathbb{C}$, and the complete monotonicity provides some a priori bounds. For this reason, we can derive error bounds for our approximation problem in the interval $[a,b]$ from the knowledge of a function with small values in $[a,b]$ and symmetry properties. The latter is provided by the rational approximation of the square root function and is related to other fast computations, as we will see in the next section.

The extra assumption in the following lemma concerning the continuity of f at $x=0$ will be no drawback, since a shift $x\mapsto x+1/n$ will recover it.

Lemma 2.1. *Let f be completely monotone for $x>0$ and continuous at $x=0$. Assume that p_n and q_{n-1} are polynomials of degree n and $n-1$, respectively, and that*

$$\left|\frac{p_n(x)}{q_{n-1}(x)}-\sqrt{x}\right|\le\varepsilon\sqrt{x}\qquad \text{for } x\in[a^2,b^2] \tag{3}$$

$$\text{or}\quad \left|\frac{p_n(x)/q_{n-1}(x)-\sqrt{x}}{p_n(x)/q_{n-1}(x)+\sqrt{x}}\right|\le\varepsilon\qquad \text{for } x\in[a^2,b^2], \tag{4}$$

holds for some $\varepsilon>0$. Assume also that $p_n/q_{n-1}-\sqrt{x}$ has $2n$ zeros in the interval $[a^2,b^2]$. Then there exists an exponential sum u_n with n terms such that

$$|f(x)-u_n(x)|\le 2\varepsilon f(0)\quad \text{for } x\in[a,b].$$

Proof. Put $x=z^2$. Obviously, we can restrict ourselves to the case $\varepsilon<1$. Now (3) implies (4) and by assumption

$$\left|\frac{p_n(z^2)-zq_{n-1}(z^2)}{p_n(z^2)+zq_{n-1}(z^2)}\right|\le\varepsilon\quad \text{for } z\in[a,b].$$

Set $P_{2n}(z):=p_n(z^2)-zq_{n-1}(z^2)$ and write

$$\left|\frac{P_{2n}(z)}{P_{2n}(-z)}\right|\le\varepsilon\quad \text{for } z\in[a,b]. \tag{5}$$

Obviously,

$$\left|\frac{P_{2n}(z)}{P_{2n}(-z)}\right|=1\quad \text{for } \Re e\, z=0\quad \text{or } |z|\to\infty.$$

Let u_n be the interpolant of f at the $2n$ zeros of P_{2n}. The last inequality in

$$|f(z)|\le f(0),\quad |u_n(z)|\le u_n(\Re e\, z)\le u_n(0)<f(0)\quad \text{for } \Re e\, z\ge 0$$

follows from Theorem 2.1. Hence,

$$\left| \frac{P_{2n}(-z)}{P_{2n}(z)} \Big(f(z) - u_n(z) \Big) \right| \leq 2f(0) \tag{6}$$

holds at the boundary of the right half-plane. By the maximum principle for analytic functions (6) holds also in the interior, and

$$|f(z) - u_n(z)| \leq 2f(0) \left| \frac{P_{2n}(z)}{P_{2n}(-z)} \right|$$

completes the proof.

A similar method is sketched in Appendix 10. The maximum principle is applied to an analytic function on a sector of the complex plane and with properties different from (5). The inequality (5) is related to the capacity of a capacitor with the plates $[a,b]$ and $[-b,-a]$. We are looking for a rational function, whose absolute value is small on $[a,b]$ and large on $[-b,-a]$.

Lemma 2.1 provides only upper bounds for the degree of approximation. Surprisingly, numerical results in Section 3 lead to the conjecture that the asymptotic behaviour and the exponential decay is precisely described for finite intervals. The estimates for infinite intervals reflect the asymptotic behaviour, but are not optimal, although they are sharper than the estimate obtained via Sinc approximation methods [7] and §11.

3 Rational approximation of the square root function

3.1 Heron's algorithm and Gauss' arithmetic-geometric mean

At the beginning of the second century, Heron of Alexandria described a method to calculate the square root of a given positive number a using some initial approximation. The method was probably also known to the Babylonians. A modification – more precisely a rescaling – will help us to construct best rational approximations of the square root function in the sense of Chebyshev [22].

Let x_n be an approximation of $\sqrt{a}$. Obviously $\sqrt{a}$ is the *geometric mean* of x_n and a/x_n. Heron replaced it by the *arithmetic mean*, i.e., in modern notation:

$$x_{n+1} = \frac{1}{2}(x_n + \frac{a}{x_n}).$$

Convergence follows from the recursion relation for the error

$$x_{n+1} - \sqrt{a} = \frac{(x_n - \sqrt{a})^2}{2x_n}. \tag{7}$$

Gauss considered the two means in a different context. At an early age, he became enamoured of a sequential procedure that is now known as the arithmetic-geometric

process (see, e.g., [3]). Given two numbers $0 < a_0 < b_0$, one successively takes the arithmetic mean and the geometric mean:

$$a_{j+1} := \sqrt{a_j b_j}, \qquad b_{j+1} := \frac{1}{2}(a_j + b_j). \tag{8}$$

He expressed the common limit as an elliptic integral (see Appendix 9). The distance of the two numbers is characterised by $\lambda_j := b_j/a_j$. A direct calculation yields

$$\lambda_{j+1} = \frac{1}{2}\left(\sqrt{\lambda_j} + \frac{1}{\sqrt{\lambda_j}}\right) \quad \text{or} \quad \lambda_j = \left(\lambda_{j+1} + \sqrt{\lambda_{j+1}^2 - 1}\right)^2. \tag{9}$$

The mapping $\lambda \mapsto (\lambda + \sqrt{\lambda^2 - 1})^2$ is called the *Landen transformation*. The numbers in the table below show that a few steps forwards or backwards brings us either to large numbers or to numbers very close to 1. – Finally, we mention that the numbers $(\lambda+1)/(\lambda-1)$ and λ' with $(\lambda')^{-2} + \lambda^{-2} = 1$ are moved by the same rule, but in the opposite direction.

Table 1 Arithmetic-geometric process with $\lambda_0 = 1 + \sqrt{2}$ and $\lambda_j^{-2} + (\lambda_j')^{-2} = 1$

j	λ_j	$\frac{\lambda_{j+1}}{\lambda_j - 1}$	λ_j'
-4	$6.825 \cdot 10^{14}$	$1 + 2.930 \cdot 10^{-15}$	$1 + 1.07 \cdot 10^{-30}$
-3	$1.306 \cdot 10^{7}$	$1 + 1.531 \cdot 10^{-7}$	$1 + 2.930 \cdot 10^{-15}$
-2	1807.08	1.001107	$1 + 1.531 \cdot 10^{-7}$
-1	21.26	1.099	1.001107
0	2.414	2.414	1.099
1	1.099	21.26	2.414
2	1.001107	1807.08	21.26
3	$1 + 1.531 \cdot 10^{-7}$	$1.306 \cdot 10^{7}$	1807.08
4	$1 + 2.930 \cdot 10^{-15}$	$6.825 \cdot 10^{14}$	$1.306 \cdot 10^{7}$

3.2 Heron's method and best rational approximation

In view of Lemma 2.1 we are interested in the best relative Chebyshev approximation of $\sqrt{x}$ by rational functions in $R_{n,n-1}$. Specifically, v_n is called a best approximation if it yields the solution of the minimisation problem:

$$E_{n,n-1} := E_{n,n-1,[a,b]} := \inf_{v_n \in R_{n,n-1}} \left\| \frac{v_n - \sqrt{x}}{\sqrt{x}} \right\|_{L_\infty[a,b]}.$$

Definition 3.1. An error curve $\eta(x)$ has an alternant of length ℓ, if there are ℓ points $x_1 < x_2 < \ldots < x_\ell$ such that

$$\operatorname{sign}\eta(x_{i+1}) = -\operatorname{sign}\eta(x_i) \quad \text{for } i = 1,2,\ldots\ell-1 \tag{10}$$

and

$$|\eta(x_i)| = \|\eta\|_{L_\infty} \quad \text{for } i = 1,2,\ldots\ell \tag{11}$$

holds.

The following characterisation goes back to Chebyshev. Some degeneracies that are possible with rational approximation, cannot occur here.

Theorem 3.1 (characterisation theorem). *v_n is optimal in $R_{n,n-1}$ if and only if the error curve $(v_n - \sqrt{x})/\sqrt{x}$ has an alternant of length $2n+1$.*

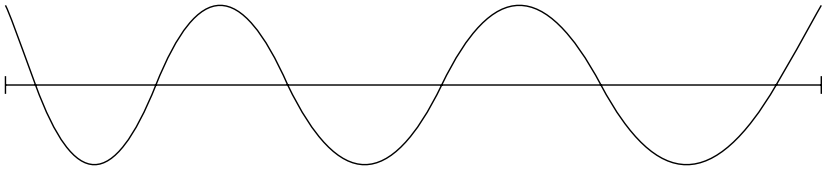

Fig. 1 Alternant of length 7

Let $p_n/q_{n-1} \in R_{n,n-1}$ be an approximation of $\sqrt{x}$. The application of one step of Heron's algorithm yields the rational function

$$\frac{1}{2}\left(\frac{p_n}{q_{n-1}} + \frac{x}{p_n/q_{n-1}}\right) = \frac{p_n^2 + xq_{n-1}^2}{2p_nq_{n-1}} \in R_{2n,2n-1}$$

From (7) we conclude that the associated error curve is non-negative and cannot be a best approximation; see Figure 2. A rescaling before and after the procedure, however, will yield a solution. This was already observed by Rutishauser [22], although he stopped at (12) and did not mention the connection with Gauss' arithmetic-geometric process.

Let v_n be the best approximation in $R_{n,n-1}$. By definition,

$$1 - E_{n,n-1} \le \frac{v_n(x)}{\sqrt{x}} \le 1 + E_{n,n-1}\,.$$

The corresponding relations for $w_n := \frac{1}{\sqrt{1-E_{n,n-1}^2}} v_n$ are

$$\sqrt{\frac{1-E_{n,n-1}}{1+E_{n,n-1}}} \le \frac{w_n(x)}{\sqrt{x}} \le \sqrt{\frac{1+E_{n,n-1}}{1-E_{n,n-1}}}\,.$$

The result of a Heron step is denoted by w_{2n} and

$$1 \le \frac{w_{2n}(x)}{\sqrt{x}} = \frac{1}{2}\left(\frac{w_n(x)}{\sqrt{x}} + \frac{\sqrt{x}}{w_n(x)}\right)$$
$$\le \frac{1}{2}\left(\sqrt{\frac{1+E_{n,n-1}}{1-E_{n,n-1}}} + \sqrt{\frac{1-E_{n,n-1}}{1+E_{n,n-1}}}\right) = \frac{1}{\sqrt{1-E_{n,n-1}^2}}.$$

We rescale the new rational function, set $v_{2n} := \frac{2\sqrt{1-E_{n,n-1}^2}}{1+\sqrt{1-E_{n,n-1}^2}} w_{2n}$, and obtain

$$\frac{2\sqrt{1-E_{n,n-1}^2}}{1+\sqrt{1-E_{n,n-1}^2}} \le \frac{v_{2n}(x)}{\sqrt{x}} \le \frac{2}{1+\sqrt{1-E_{n,n-1}^2}}.$$

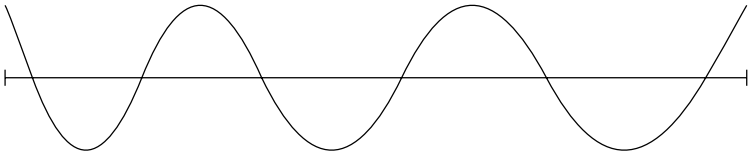

Fig. 2 Error curves and Heron's procedure

Figure 2 elucidates that the number of sign changes is doubled, and the equilibration above yields the desired alternant of length $4n+1$. Hence,

$$E_{2n,2n-1} = \frac{2}{1+\sqrt{1-E_{n,n-1}^2}} - 1 = \frac{1-\sqrt{1-E_{n,n-1}^2}}{1+\sqrt{1-E_{n,n-1}^2}} = \frac{E_{n,n-1}^2}{\left(1+\sqrt{1-E_{n,n-1}^2}\right)^2} \quad (12)$$

or

$$E_{2n,2n-1}^{-1} = \left(E_{n,n-1}^{-1} + \sqrt{E_{n,n-1}^{-2}-1}\right)^2. \quad (13)$$

Remark 3.1. The inverse $E_{2n,2n-1}^{-1}$ is obtained from $E_{n,n-1}^{-1}$ by the Landen transformation. In particular,

$$(\frac{1}{4}E_{n,n-1})^2 \le \frac{1}{4}E_{2n,2n-1}. \quad (14)$$

We will make repeated use of the following consequence: The inequality $E_{2n,2n-1} \le 4A^2$ with some $A > 0$ implies $E_{n,n-1} \le 4A$.

A start for the recursive procedure is the best constant function. The best constant for the interval $[a^2, b^2]$ and the approximation error follow from a simple optimisation of the constant:

$$v_{0,0} = \frac{2ab}{a+b}, \qquad E_{0,0,[a^2,b^2]} = \frac{b-a}{b+a} =: \frac{1}{\rho}. \quad (15)$$

Here ρ is the parameter of the ellipse on which the square root is an analytic function, if the interval $[a^2, b^2]$ is transformed into the interval $[-1, +1]$; cf. Appendix 8. Another important parameter is

$$\kappa := \frac{a}{b}.$$

Note that $v_{0,0}$ is the *harmonic mean* of the function values at the end points.

When Heron's method is applied to a constant function, a linear function with an alternant of length 3 is produced. Hence, $E_{1,0}^{-1} = \left(E_{0,0}^{-1} + \sqrt{E_{0,0}^{-2} - 1}\right)^2$, and Landen transformations provide the sequence

$$\rho = E_{0,0}^{-1} \to E_{1,0}^{-1} \to E_{2,1}^{-1} \to E_{4,3}^{-1} \to E_{8,7}^{-1} \to \dots \tag{16}$$

The asymptotic behaviour of $E_{n,n-1}$ for $n = 2^m$ can be determined already from this sequence. There are the trivial inequalities for the sequence (9)

$$4\lambda_j \le (4\lambda_{j+1})^2. \tag{17}$$

Let

$$\omega := \omega(\kappa) := \omega[a^2, b^2] := \lim_{m\to\infty} \left(\frac{1}{4} E_{2^m, 2^m-1, [a^2,b^2]}\right)^{-1/2^{m+1}}. \tag{18}$$

By (14) the sequence on the right-hand side is monotone, the limit exists, and the monotonicity also implies that

$$E_{n,n-1,[a^2,b^2]} \le 4\omega^{-2n} \tag{19}$$

holds for $n = 2^m$. We will establish the inequality for all $n \in \mathbb{N}$. Moreover, $\omega(\kappa)$ will be expressed in terms of elliptic integrals although the fast convergence of the arithmetic-geometric process is used for its fast computation, as we will see below.

Remark 3.2. We focus on upper bounds for the degree of rational approximation although lower bounds can be obtained by suitable modifications. We elucidate this for a bound corresponding to (19). Let $\lambda_j \ge \frac{1}{4}A + \frac{2}{A}$ for some $A > 1$. Hence,

$$\lambda_{j-1} \ge \left[\frac{1}{4}A + \frac{2}{A} + \sqrt{(\frac{1}{16}A^2 + 1 + \frac{4}{A^2}) - 1}\right]^2 \ge \left(\frac{1}{2}A + \frac{2}{A}\right)^2 \ge \frac{1}{4}A^2 + \frac{2}{A^2}.$$

The bound of λ_{j-1} has the same structure as the bound for λ_j. Now we obtain by induction and from (18)

$$E_{n,n-1,[a^2,b^2]} \ge \frac{4}{\omega^{2n} + 8\omega^{-2n}} \tag{20}$$

for $n = 2^m$. A comparison with (19) elucidates the fast convergence.

3.3 *Extension of the estimate* (19)

A transformation of the interval will be used for the extension of inequality (19) which was previously announced. It enables us to derive sharp estimates from the results for small intervals in Appendix 8. We encounter the arithmetic-geometric process once more.

Lemma 3.1. *Let $n \geq 1$ and (a_j, b_j) be a sequence according to the arithmetic-geometric mean process (8). Then*

$$E_{n,n-1,[a_{j+1}^2,b_{j+1}^2]} = E_{2n,2n-1,[a_j^2,b_j^2]}\,. \tag{21}$$

Proof. Set $r(x) := (x + a_j b_j)/2$. The function $r^2(x)/x$ maps the two subintervals $[a^2, ab]$ and $[ab, b^2]$ monotonously onto $[a_j b_j, (a_j + b_j)^2/4] = [a_{j+1}^2, b_{j+1}^2]$. Next, note that $\sqrt{x} = r(x)\sqrt{x/r^2(x)} = r(x)\sqrt{\xi}$ where $\xi = x/r(x)^2$.

Let $p/q \in R_{n,n-1}$ be the best approximation to $\sqrt{x}$ on $[a_{j+1}^2, b_{j+1}^2]$. Then

$$\frac{P(x)}{Q(x)} := r(x)\frac{p(x/r^2(x))}{q(x/r^2(x))} \in R_{2n,2n-1}$$

provides an approximation for the original interval with the same size of the maximal relative error as p/q on the smaller interval. The monotonicity of the mapping $r^2(x)/x$ assures that there is an alternant of length $4n+1$. Therefore, P/Q is the best approximation, and the proof is complete.

As a by-product we obtain a closed expression for the approximation by linear functions. For completeness, we also recall (15):

$$E_{1,0,[a^2,b^2]} = \Big(\frac{\sqrt{b}-\sqrt{a}}{\sqrt{b}+\sqrt{a}}\Big)^2, \qquad E_{0,0,[a^2,b^2]} = \frac{b-a}{b+a}\,. \tag{22}$$

Theorem 3.2. *Let ω be defined by (18). Then the degree of approximation is bounded by (19) for all $n \in \mathbb{N}$.*

Proof. Let $[a_0^2, b_0^2]$ be the interval for which the degree of approximation in $R_{n,n-1}$ is to be estimated and $\omega = \omega[a_0^2, b_0^2]$. Moreover, let $\ell = 2^k$. By Lemma 3.1 we know that

$$E_{0,0,[a_{k+1}^2,b_{k+1}^2]} = E_{1,0,[a_k^2,b_k^2]} = E_{\ell,\ell-1,[a_0^2,b_0^2]} \leq 4\omega^{-2\ell}.$$

From (45) it follows that the parameter of the regularity ellipse associated to the interval $[a_{k+1}^2, b_{k+1}^2]$ is the inverse, i.e.,

$$\rho = \frac{1}{4}\omega^{2\ell}.$$

By (44) we have

$$E_{n,n-1,[a_{k+1}^2,b_{k+1}^2]} \le 4(4\rho-3)^{-2n} \le 4(\omega^{2\ell}-3)^{-2n}.$$

By using the preceding lemma once more we return to the original interval,

$$E_{2\ell n,2\ell n-1,[a_0^2,b_0^2]} \le 4(\omega^{2\ell}-3)^{-2n}.$$

The degree of the numerator is $2\ell n = 2^{k+1}n$. Now we perform $k+1$ Landen transformations in the opposite direction and recall Remark 3.1 to obtain

$$E_{n,n-1,[a_0^2,b_0^2]} \le 4(\omega^{2\ell}-3)^{-2n/2\ell} = 4\omega^{-2n}(1-3\omega^{-2\ell})^{-2n/2\ell}.$$

Since we may choose an arbitrarily large ℓ, the proof is complete.

3.4 An explicit formula

The asymptotic behaviour of the degree of approximation was determined for finite intervals without the knowledge of elliptic integrals – in contrast to [32]. Explicit formulae will be useful for the treatment of the approximation with exponential sums on infinite intervals. The relevant properties of complete elliptic integrals are provided in Appendix 9.

Theorem 3.3. *Let $k = a/b$, then*

$$E_{n,n-1,[a^2,b^2]} \le 4\omega^{-2n}, \quad \textit{for } n = 1,2,3,\ldots \tag{23}$$

with

$$\omega(k) = \exp\left[\frac{\pi \mathbf{K}(k)}{\mathbf{K}'(k)}\right]. \tag{24}$$

Proof. Set $\kappa_{-j} := E_{2^j,2^j-1}$ and $\lambda_{-j} := 1/\kappa_{-j} :=$ for $j = 0,1,2,\ldots$ and extend the two sequences by the backward Landen transformation. From (16) we know that λ_{-j} obeys the rule of the arithmetic-geometric process, and $\kappa_{-j+1} = 2\sqrt{\kappa_{-j}}/(1+\kappa_{-j})$. By Lemma 9.1 and (53) we obtain

$$\begin{aligned}\lim_{j\to\infty}\left(\frac{1}{4}E_{2^j,2^j-1}\right)^{-1/2^j} &\ge \exp\left[\frac{\pi \mathbf{K}'(\kappa_0)}{2\mathbf{K}(\kappa_0)}\right] = \exp\left[\frac{2\pi \mathbf{K}'(\kappa_2)}{\mathbf{K}(\kappa_2)}\right] \\ &= \exp\left[\frac{2\pi \mathbf{K}(\kappa_2')}{\mathbf{K}'(\kappa_2)}\right] \\ &= \exp\left[2\pi \mathbf{K}\left(\frac{1-\kappa_1}{1+\kappa_1}\right)\Big/\mathbf{K}'\left(\frac{1-\kappa_1}{1+\kappa_1}\right)\right]. \end{aligned} \tag{25}$$

It follows from $\kappa_1 = E_{0,0}$ and (22) that

$$\frac{1-\kappa_1}{1+\kappa_1} = \frac{a}{b}.$$

Now the left-hand side of (25) can be identified with ω^2, and the proof is complete.

Example 3.1. We consider the approximation problem on the interval $[1,2]$, and from (22) we know that $E_{0,0} = (\sqrt{2}-1)/(\sqrt{2}+1)$. The sequence (16) and the successive calculation of square roots for modelling (18) yields the tableau

$$\begin{array}{r} \frac{\sqrt{2}-1}{\sqrt{2}+1} = 5.828427 \to 133.87475 \to 4613.84 \to 71687.79 \\ \times 4 \downarrow \\ 23.140689 \leftarrow 535.4915 \leftarrow 286751.2 \end{array}$$

Since $\mathbf{K}(1/\sqrt{2}) = \mathbf{K}'(1/\sqrt{2})$, the evaluation of ω by formula (18) is easy. We get $\omega = \exp(\pi) = 23.1406924$ in accordance with the result in the tableau above.

4 Approximation of $1/x^\alpha$ by exponential sums

4.1 Approximation of $1/x$ on finite intervals

The symbol $E_{n,[a,b]}(f)$ with only one integer index refers to the approximation by exponential sums of order n. In order to have a short notation we start with the approximation of $1/x$: First we note that

$$E_{n,[a,b]}(1/x) = \frac{1}{a} E_{n,[1,b/a]}(1/x). \tag{26}$$

Indeed, let u_n be the best approximation of $1/x$ on the interval $[1,b/a]$. The transformation $x = at$ yields

$$\frac{1}{x} - \frac{1}{a} u_n\left(\frac{x}{a}\right) = \frac{1}{a}\left[\frac{1}{t} - u_n(t)\right]. \tag{27}$$

Since the alternant is transformed into an alternant, we have (26).

Theorem 4.1. *Let $0 < a < b$ and $k = a/b$. Then*

$$E_{n,[a,b]}(1/x) \le \frac{c(k)}{a}\, n\, \omega(k)^{-2n}$$

with $\omega(k)$ given by (24) and $c(k)$ depending only on k.

Proof. By (26) it is sufficient to study approximation on the interval $[1,1/k]$. To this end, we consider the approximation of $f(x) := \frac{1}{x+1/n}$ on the interval $[1-1/n, 1/k-$

$1/n]$. Since we are interested in upper bounds, we may enlarge the interval to $[1-1/n, 1/k]$. It follows from Lemma 2.1, Theorem 3.3, and $f(0) = n$ that

$$E_{n,[1,1/k]}(1/x) \le E_{n,[1-1/n,1/k]}(1/(x+1/n)) \le 2n4\left[\omega\left(\frac{1-1/n}{1/k}\right)\right]^{-2n}.$$

Since the function $k \mapsto \omega(k)$ is differentiable, we have

$$\omega\left(\frac{1-1/n}{1/k}\right) = \omega(k[1-1/n]) \ge \omega(k)(1-\frac{c}{n})$$

with $c = c(k)$ being a bound of the derivative in a neighbourhood of k. We complete the proof by recalling $\lim_{n\to\infty}(1-c/n)^{2n} = \mathrm{e}^{-2c}$.

Theorem 4.1 provides only an upper bound. The following examples for small and large intervals, respectively, show that the order of exponential decay proved there is sharp. The numerical results give rise to the conjecture that the polynomial term is too conservative and that

$$E_{n,[a,b]}(1/x) \approx n^{1/2}\omega(k)^{-2n}.$$

Example 4.1. The parameter for the (small) interval $[1,2]$, i.e., $[a^2,b^2] = [1,4]$ is evaluated in the following tableau and is to be compared with the numbers in the third column of Table 2.

$$\begin{array}{|rrrrr|}\hline 3 & \to 33.970 & \to 4613.84 & \to & 85150133 \\ & & & & \times 4 \downarrow \\ 11.655 & \leftarrow 135.85 & \leftarrow 18445.3 & \leftarrow & 340600530 \\ \hline \end{array}$$

Example 4.2. The parameter for the large interval $[1,1000]$, i.e., $[a^2,b^2] = [1,10^6]$ is evaluated in the following tableau and is to be compared with the numbers in the third column of Table 3.

$$\begin{array}{|rrrrr|}\hline 1001/999 & \to 1.13488 & \to 2.79396 & \to 29.1906 \to & 3406.37 \\ & & & & \times 4 \downarrow \\ & 1.813 \leftarrow & 3.2869 \leftarrow & 10.804 \leftarrow & 116.728 \leftarrow 13625 \\ \hline \end{array}$$

4.2 *Approximation of* $1/x$ *on* $[1,\infty)$

If we fix n and consider the approximation problem on the interval $[1,R]$, then the bound in Theorem 4.1 increases with R. This does not reflect the right asymptotic behaviour.

Table 2 Numerical results for $1/x$ (left) and $1/\sqrt{x}$ (right) on $[1,2]$

f	$1/x$		$1/\sqrt{x}$
n	E_n	$\frac{2n}{2n-1}\frac{E_{n-1}}{E_n}$	E_n
1	$2.12794\cdot 10^{-2}$		$1.26035\cdot 10^{-2}$
2	$2.07958\cdot 10^{-4}$	136.43	$9.28688\cdot 10^{-5}$
3	$1.83414\cdot 10^{-6}$	136.06	$6.83882\cdot 10^{-7}$
4	$1.54170\cdot 10^{-8}$	135.96	$5.03516\cdot 10^{-9}$
5	$1.26034\cdot 10^{-10}$	135.92	$3.70688\cdot 10^{-11}$
6	$1.01179\cdot 10^{-12}$	135.89	$2.72889\cdot 10^{-13}$

Table 3 Numerical results for $1/x$ (left) and $1/\sqrt{x}$ (right) on $[1,1000]$

f	$1/x$		$1/\sqrt{x}$
n	E_n	$\frac{2n}{2n-1}\frac{E_{n-1}}{E_n}$	E_n
5	$6.38478\cdot 10^{-4}$		$1.21681\cdot 10^{-3}$
6	$2.17693\cdot 10^{-4}$	3.1995	$3.68730\cdot 10^{-4}$
7	$7.15300\cdot 10^{-5}$	3.2776	$1.11788\cdot 10^{-4}$
8	$2.32088\cdot 10^{-5}$	3.2875	$3.39264\cdot 10^{-5}$
9	$7.46801\cdot 10^{-6}$	3.2905	$1.03020\cdot 10^{-5}$
10	$2.38880\cdot 10^{-6}$	3.2908	$3.12940\cdot 10^{-6}$
11	$7.60494\cdot 10^{-7}$	3.2907	$9.50867\cdot 10^{-7}$
12	$2.41164\cdot 10^{-7}$	3.2905	$2.88981\cdot 10^{-7}$
13	$7.62271\cdot 10^{-8}$	3.2903	$8.78389\cdot 10^{-8}$

The error curve for the best approximation u_n has $2n$ zeros in $[1,R]$. It follows from Theorem 2.1 that $u_n(x) < 1/x$ and

$$\left|\frac{1}{x} - u_n(x)\right| < \frac{1}{x} < \frac{1}{R}$$

holds for $x > R$. Hence, for all $R > 1$,

$$E_{n,[1,\infty]}(1/x) \le \max\left\{E_{n,[1,R]}(1/x), \frac{1}{R}\right\}. \tag{28}$$

It is our aim to choose R such that the right-hand side of (28) is close to the minimal value.

Table 4 Numerical results for $1/x$ (left) and $1/\sqrt{x}$ (right) on $[1,\infty)$

f	$1/x$			$1/\sqrt{x}$
n	R_n	E_n	$E_n e^{\pi\sqrt{2n}}/\log(2+n)$	E_n
1	8.667	$8.55641\cdot 10^{-2}$	6.62	$1.399\cdot 10^{-1}$
2	41.54	$1.78498\cdot 10^{-2}$	6.89	$4.087\cdot 10^{-2}$
5	1153	$6.42813\cdot 10^{-4}$	6.82	$3.297\cdot 10^{-3}$
10	56502	$1.31219\cdot 10^{-5}$	6.67	$1.852\cdot 10^{-4}$
15	$1.175\cdot 10^{6}$	$6.31072\cdot 10^{-7}$	6.62	$2.011\cdot 10^{-5}$
20	$1.547\cdot 10^{7}$	$4.79366\cdot 10^{-8}$	6.60	$3.083\cdot 10^{-6}$
25	$1.514\cdot 10^{8}$	$4.89759\cdot 10^{-9}$	6.60	$5.898\cdot 10^{-7}$
30	$1.198\cdot 10^{9}$	$6.18824\cdot 10^{-10}$	6.61	$1.321\cdot 10^{-7}$
35	$8.064\cdot 10^{9}$	$9.19413\cdot 10^{-11}$	6.62	$3.336\cdot 10^{-8}$
40	$4.771\cdot 10^{10}$	$1.55388\cdot 10^{-11}$	6.64	$9.264\cdot 10^{-9}$
45	$2.540\cdot 10^{11}$	$2.91895\cdot 10^{-12}$	6.66	$2.780\cdot 10^{-9}$
50	$1.237\cdot 10^{12}$	$5.99210\cdot 10^{-13}$	6.68	$8.901\cdot 10^{-10}$

In order to avoid the singularity at $x=0$, we consider the approximation of $f(x) := 1/(x+1/2)$ on the interval $[\frac{1}{2}, R-\frac{1}{2}]$. The constant shift of $1/2$ is better suited for estimates on large intervals. Now it follows from Lemma 2.1, Theorem 3.3, and $f(0)=2$ that

$$E_{n,[1,R]}(1/x) \le 2\cdot 2\cdot 4\exp\left[-\frac{2n\pi\mathbf{K}(k)}{\mathbf{K}'(k)}\right]$$

with $k = 1/(2R-1)$. From (48) we know that $\mathbf{K}(k) \ge \pi/2$. This inequality and (50) imply

$$E_{n,[1,R]}(1/x) \le 16\exp\left[-\frac{\pi^2 n}{\log(\frac{4}{k}+2)}\right] \le 16\exp\left[-\frac{\pi^2 n}{\log(8R)}\right]. \tag{29}$$

The choice $R = \frac{1}{8}\exp[\pi\sqrt{n}]$ yields the final result:

$$E_{n,[1,\infty]}(1/x) \le 16\exp\left[-\pi\sqrt{n}\right]. \tag{30}$$

The results in Table 4 are based on numerically computed best approximations. They lead to the conjecture that

$$E_{n,[1,\infty]}(1/x) \approx \log n\cdot\exp\left[-\pi\sqrt{2n}\right]. \tag{31}$$

In particular, the exponents in (30) and (31) differ by a factor of $\sqrt{2}$. The same gap is found with the method discussed in Appendix 10. (The approximation by sinc functions leads even to a larger gap [7] and §11.)

Table 5 Comparison of the approximation of $\sqrt{x}$ by rational functions and $1/x$ by exponential sums

k^{-1}	$E_{4,3,[1,k^{-2}]}(\sqrt{x})$	$E_{4,[1,k^{-1}]}(1/x)$
2	$1.174\cdot 10^{-8}$	$1.542\cdot 10^{-8}$
10	$8.935\cdot 10^{-5}$	$5.577\cdot 10^{-5}$
100	$9.781\cdot 10^{-3}$	$1.066\cdot 10^{-3}$
500	$2.220\cdot 10^{-2}$	$1.700\cdot 10^{-3}$

The gap may be surprising since the numerical results in the Tables 2 and 3 show that the theory provides sharp estimates for the asymptotic behaviour for large n. It is the factor in front of the exponential term in Theorem 4.1 that is responsible. We have compared the data for $n = 4$, i.e., for a small n in Table 5. They show that the application of Lemma 2.1 provides estimates which are too conservative on large intervals, although the behaviour for large n is well modelled.

The logarithmic factor in front of (31) also shows that it will not be easy to establish sharper estimates for the infinite interval.

4.3 Approximation of $1/x^{\alpha}$, $\alpha > 0$

When more freedom in the exponent of the given function is admitted, there are no substantial changes on finite intervals. Proceeding as in the proof of Theorem 4.1 we obtain with $k = a/b$:

$$E_{n,[a,b]}(x^{-\alpha}) \le \frac{c(k)}{a^{\alpha}}\, n^{\alpha}\, \omega(k)^{-2n} \tag{32}$$

with $\omega(k)$ given in Theorem 4.1. The exponential term on the right-hand side that dominates the asymptotic behaviour for large n is unchanged.

The situation on infinite intervals is different. Given $R > 1$, we obtain with $k = 1/(2R-1)$ in analogy to (29)

$$E_{n,[1,R]}(x^{-\alpha}) \le 2^{\alpha}\, 8 \exp\left[-\frac{\pi^2 n}{\log(\frac{4}{k}+2)}\right] \le 2^{\alpha}\, 8 \exp\left[-\frac{\pi^2 n}{\log(8R)}\right] \tag{33}$$

Moreover, we have $E_{n,[1,\infty]}(x^{-\alpha}) \le \max\left\{E_{n,[1,R]}(x^{-\alpha}), R^{-\alpha}\right\}$ in analogy to (28). A suitable choice is $R = \frac{1}{8}\exp[\pi\sqrt{n/\alpha}]$. It yields

$$E_{n,[1,\infty]}(x^{-\alpha}) \le 2^{\alpha}\, 8 \exp[-\pi\sqrt{\alpha n}]. \tag{34}$$

The asymptotic decay depends heavily on α.

5 Applications of $1/x$ approximations

5.1 About the exponential sums

Let $[a,b] \subset (0,\infty]$ be a possibly semi-infinite interval, e.g. $b = \infty$ is allowed. The best approximation in $[a,b]$ is denoted by

$$\frac{1}{x} \approx u_{n,[a,b]}(x) = \sum_{\nu=1}^{n} \alpha_{\nu,[a,b]} \exp(-t_{\nu,[a,b]}\, x).$$

The rule (27) is inherited by the coefficients,

$$\alpha_{\nu,[a,b]} := \frac{1}{a}\alpha_{\nu,[1,b/a]}, \quad t_{\nu,[a,b]} := \frac{1}{a}t_{\nu,[1,b/a]},$$

and allows us to reduce the considerations to intervals of the form $[1,R]$. Due to (26) the approximation errors $E_{n,[a,b]}$ are related by $E_{n,[a,b]} = \frac{1}{a}E_{n,[1,b/a]}$. The coefficients of $v_{n,[1,R]}$ for various n and R can be found in [31].

5.2 Application in quantum chemistry

The so-called Coupled Cluster (CC) approaches are rather accurate but expensive numerical methods for solving the electronic many-body problems. The cost may be $\mathcal{O}(N^7)$, where N is the number of electrons. One of the bottlenecks is an expression of the form

$$\frac{numerator}{\varepsilon_a + \varepsilon_b + \ldots - \varepsilon_j - \varepsilon_i},$$

where $\varepsilon_i, \varepsilon_j, \ldots < 0$ are energies related to occupied orbitals $i, j, \ldots$, while $\varepsilon_a, \varepsilon_b, \ldots > 0$ are energies related to virtual orbitals $a, b, \ldots$ The denominator belongs to an interval $[E_{\min}, E_{\max}]$, where the critical lower energy bound $E_{\min}$ depends on the so-called HOMO-LUMO gap.

The denominator leads to a coupling of all orbitals $a, b, \ldots, i, j, \ldots$, whereas the numerator possesses a partial separation of variables. Therefore one tries to replace $1/(\varepsilon_a + \varepsilon_b + \ldots - \varepsilon_i - \varepsilon_j)$ by a separable expression. Such a separation saves one order in the complexity.

Any exponential sum approximation $\frac{1}{x} \approx \sum_{\nu=1}^{n} \alpha_\nu \exp(-t_\nu x)$ leads to the separable expression

$$1/(\varepsilon_a + \varepsilon_b + \ldots - \varepsilon_i - \varepsilon_j) \approx \sum_{\nu=1}^{n} \alpha_\nu \mathrm{e}^{-t_\nu \varepsilon_a} \mathrm{e}^{-t_\nu \varepsilon_b} \cdots \mathrm{e}^{t_\nu \varepsilon_j} \mathrm{e}^{t_\nu \varepsilon_i}.$$

In quantum chemistry, Almlöf [1] used the representation

$$\frac{1}{x} = \int_0^\infty \mathrm{e}^{-sx} \mathrm{d}s$$

together with a quadrature formula $\sum_{\nu=1}^{n} \alpha_\nu e^{-t_\nu x}$. This ansatz has been used in many places like the Møller-Plesset second order perturbation theory (MP2, cf. [1, 15, 16, 30]), computation of connected triples contribution in MP4 (cf. [15]), atomic orbital (AO)-MP2 (cf. [16, 2, 19]), AO-MP2 energy gradient (cf. [16, 24]), combinations with the resolution of the identity (RI)-MP2 (cf. [10, 9]), and the density-matrix-based MP2 (cf. [27, 17]).

It is hard to adapt the quadrature to the interval $[E_{\min}, E_{\max}]$ where the approximation is needed. The favourite choice among those used in quantum chemistry is the Gauss-Legendre quadrature applied to the transformed integral

$$\int_0^\infty e^{-sx} ds = \int_0^1 e^{-tx/(1-t)} \frac{dt}{(t-1)^2} \qquad (s = t/(1-t)).$$

Best approximations are only considered with respect to a weighted L^2-norm (cf. [23]). Best approximations in the supremum norm has not been considered in this community. The recent paper [28] contains a comparison between the Gauss-Legendre approach and the best approximation $u_{n,[E_{\min},E_{\max}]}$ for various applications. For instance, an error of size ≈ 0.005 of the MP2 energies for benzene with the aug-cc-pCVTZ basis set is obtained by the Gauss-Legendre quadrature with 14 terms, while the same accuracy is already obtained by the best approximation with 4 terms (best approximations with 14 terms yield an accuracy of $2 \cdot 10^{-10}$). The value $R = E_{\max}/E_{\min}$ for this example is about 278.

5.3 Inverse matrix

The previous application refers to the scalar function $1/x$. Now we consider its matrix-valued version M^{-1} for a matrix M with positive spectrum $\sigma(M) \subset [a,b] \subset (0,\infty]$. Formally, we have

$$M^{-1} \approx u_{n,[a,b]}(M) = \sum_{\nu=1}^{n} \alpha_{\nu,[a,b]} \exp(-t_{\nu,[a,b]} M).$$

Additionally, we assume that M is diagonalisable: $M = T^{-1} D T$. Then a simple calculation shows the estimate

$$\left\| M^{-1} - u_{n,[a,b]}(M) \right\|_2 \le cond_2(T)\, E_{n,[a,b]}$$

with respect to the spectral norm. We emphasize that the spectral norm estimate hinges on a *uniform* estimate of $\frac{1}{x} - u_{n,[a,b]}$ on the spectral interval $[a,b]$. Approximations of $1/x$ by exponential sums with respect to the L^2-norm would not be helpful.

The approximation of M^{-1} seems to be rather impractical since now matrix exponentials $\exp(-t_\nu M)$ have to be evaluated. The interesting applications, however, are matrices which are sums of Kronecker products.

We recall that a differential operator L is called separable in $x_1, \ldots, x_d$, if $L = \sum_{i=1}^{d} L_i$, where the operator L_i applies only to the variable x_i and the coefficients of L_i depend only on x_i. Let the domain of the boundary value problem be of product form: $\Omega = \Omega_1 \times \ldots \times \Omega_d$. Then a suitable discretisation leads to an index set I of product form: $I = \prod_{i=1}^{d} I_i$, where I_i contains the indices of the i-th coordinate direction.

The system matrix for a suitable discretisation has the form

$$\mathbf{M} = \sum_{i=1}^{d} I \otimes \cdots \otimes M^{(i)} \otimes \cdots \otimes I, \quad M^{(i)} \in \mathbb{R}^{I_i \times I_i} \tag{35}$$

(factor $M^{(i)}$ at i-th position). We assume that all $M^{(i)}$ are positive definite with smallest eigenvalue $\lambda_{\min}^{(i)}$. Since the spectrum of $\mathbf{M}$ is the sum $\sum_{i=1}^{d} \lambda^{(i)}$ of all $\lambda^{(i)} \in \sigma(M^{(i)})$, the minimal eigenvalue of $\mathbf{M}$ is $\lambda_{\min} := \sum_{i=1}^{d} \lambda_{\min}^{(i)}$. Since $\lambda_{\min}^{(i)}$ approximates the smallest eigenvalue of L_i, we have $\lambda_{\min} = \mathcal{O}(1)$.

Now we take the best approximation E_n^* with respect $[\lambda_{\min}, b]$ ($b = \sum_{i=1}^{d} \lambda_{\max}^{(i)}$ or $b = \infty$). We know that for the symmetric matrices

$$\left\| u_n(\mathbf{M}) - \mathbf{M}^{-1} \right\| \leq E_{n,[\lambda_{\min}, b]}.$$

For the evaluation of $u_n(\mathbf{M}) = \sum_{\nu=1}^{n} \alpha_\nu \exp(-t_\nu \mathbf{M})$ we make use of the identity

$$\exp(-t_\nu \mathbf{M}) = \bigotimes_{i=1}^{d} \exp\left(-t_\nu M^{(i)}\right)$$

with $M^{(i)}$ from (35) (cf. [14, §15.5.2]) and obtain

$$\mathbf{M}^{-1} \approx \sum_{\nu=1}^{n} \alpha_\nu \bigotimes_{i=1}^{d} \exp\left(-t_\nu M^{(i)}\right).$$

As described in [14, §13.3.1], the hierarchical matrix format allows us to approximate $\exp\left(-t_\nu M^{(i)}\right)$ with a cost almost linear in the size of $M^{(i)}$. The total number of arithmetical operations is $\mathcal{O}(n \sum_{i=1}^{d} \#I_i \log^* \#I_i)$. For $\#I_i = N$ $(1 \leq i \leq d)$ this expression is $\mathcal{O}(ndN \log^* N)$ and depends only linearly on d.

Therefore, it is possible to treat cases with large N and d. In [11], examples can be found with $N = 1024$ and $d \approx 1000$. Note that in this case $\mathbf{M}^{-1} \in \mathbb{R}^{M \times M}$ with $M \approx 10^{3000}$.

6 Applications of $1/\sqrt{x}$ approximations

6.1 Basic facts

Let $[a,b] \subset (0,\infty]$ be as above. We consider the best approximation of $1/\sqrt{x}$ in $[a,b]$:

$$\frac{1}{\sqrt{x}} \approx u_{n,[a,b]}(x) = \sum_{\nu=1}^{n} \alpha_{\nu,[a,b]} \exp(-t_{\nu,[a,b]}\, x).$$

In this case the relations

$$\alpha_{\nu,[a,b]} = \frac{1}{\sqrt{a}}\alpha_{\nu,[1,b/a]}, \quad t_{\nu,[a,b]} = \frac{1}{a}t_{\nu,[1,b/a]}, \quad E_{n,[a,b]} = \frac{1}{\sqrt{a}}E_{n,[1,b/a]} \tag{36}$$

hold, and again it is sufficient to determine $v_{n,[1,R]}$ with $R := b/a$. The coefficients of $v_{n,[1,R]}$ for various n and R can be obtained from [31].

The standard application uses the substitution $x = \|y\|^2 = \sum_{i=1}^{d} y_i^2$ with a vector $y \in \mathbb{R}^d$. Then we obtain the sum

$$G_{n,[a,b]}(y) := \sum_{\nu=1}^{n} \alpha_{\nu,[a^2,b^2]} \prod_{i=1}^{d} \exp(-t_{\nu,[a^2,b^2]}\, y_i^2)$$

of Gaussians which is the best approximation of $1/\|y\|$ for $\|y\| \in [a,b]$. Since, in 3D, $1/\|y\|$ is the Newton potential or Coulomb potential, this function appears in many problems.

6.2 Application to convolution

A further application refers to the convolution integral

$$\Phi(x) := \int_{\mathbb{R}^3} \frac{f(y)}{\|x-y\|}\,\mathrm{d}y.$$

We assume that f can be written as a sum of simple products. For simplicity we consider only one term:

$$f(y) = f_1(y_1)\, f_2(y_2)\, f_3(y_3). \tag{37}$$

When we replace $1/\|x-y\|$ by an approximation of the form

$$G_n(x-y) = \sum_{\nu=1}^{n} \alpha_\nu \prod_{i=1}^{3} \exp(-t_\nu\,(x_i-y_i)^2),$$

the convolution integral becomes

$$\Phi_n(x) := \int_{\mathbb{R}^3} G_n(x-y) f(y) \mathrm{d}y = \sum_{\nu=1}^{n} \alpha_\nu \prod_{i=1}^{3} \int_{\mathbb{R}} \exp(-t_\nu (x_i - y_i)^2) f_i(y_i) \mathrm{d}y_i,$$

and the 3D convolution is reduced to three 1D convolutions. This fact reduces the computational cost substantially. In the paper [13] this technique is applied for the case that the functions f_i are piecewise polynomials. However, there still remains a gap is to be closed. We have used some best approximation $G_n = G_{n,[a,b]}$ of $1/\|\cdot\|$. The value of b may be infinite or finite, if the support of f is finite and the evaluation of $\Phi(x)$ is required for x in a bounded domain. The lower bound a may be small but positive. Therefore the difference $\Phi(x) - \Phi_n(x)$ contains the term $\delta\Phi_n(x) := \int_{\|x-y\|\le A} \big(\frac{1}{\|x-y\|} - G_n(x-y)\big) f(y) \mathrm{d}y$, where the approximation fails. This contribution can be treated separately to obtain $\Phi_n + \delta\Phi_n \approx \Phi$. As shown in [13] the numerical cost arising from the extra term, is low.

A related problem is the integral $I := \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{g(x)f(y)}{\|x-y\|} \mathrm{d}x\mathrm{d}y$ which appears for example as "two-electron integral" in Quantum Chemistry. It can be considered and computed as the scalar product of g with the function Φ from above. Another approach is the replacement of $1/\|\cdot\|$ by the exponential sum G_n. Assuming again that f and g are simple products like in (37), the identity

$$\begin{aligned} I_n &:= \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} G_n(x-y) g(x) f(y) \mathrm{d}x\mathrm{d}y \\ &= \sum_{\nu=1}^{n} \alpha_\nu \prod_{i=1}^{3} \int_{\mathbb{R}} \int_{\mathbb{R}} \exp(-t_\nu (x_i - y_i)^2) g_i(x_i) f_i(y_i) \mathrm{d}y_i \end{aligned}$$

shows that the six-dimensional integral is reduced to three two-dimensional ones. Concerning the error analysis, we split the integral $I = I_{\mathrm{near}} + I_{\mathrm{far}}$ into the near-field and far-field parts

$$I_{\mathrm{near}} := \int_{\|z\|\le r} \int_{\mathbb{R}^3} \frac{g(z+y)f(y)}{\|z\|} \mathrm{d}z\mathrm{d}y, \quad I_{\mathrm{far}} := \int_{\|z\|\ge r} \int_{\mathbb{R}^3} \frac{g(z+y)f(y)}{\|z\|} \mathrm{d}z\mathrm{d}y.$$

Let $I_n = I_{\mathrm{near},n} + I_{\mathrm{far},n}$ be the corresponding splitting with $1/\|\cdot\|$ replaced by G_n. We assume that $f, g \in C(\mathbb{R}^3)$ have bounded support[1]. Then for $\|z\| \le r$ the error can be bounded by $|I_{\mathrm{near},n}| + |I_{\mathrm{near}}| \lesssim \int_{\|z\|\le r} \frac{\mathrm{d}z}{\|z\|} = \mathcal{O}(r^2)$. If an error ε is desired, we need $r \sim \sqrt{\varepsilon}$. This requires the choice $G_n = G_{n,[\sqrt{\varepsilon},\infty)}$. The approximation error of G_n is $\left\| 1/\|\cdot\| - G_{n,[\sqrt{\varepsilon},\infty)} \right\|_{\infty,\|z\|\ge\sqrt{\varepsilon}} = E_{n,[\varepsilon,\infty)} = \frac{1}{\sqrt{\varepsilon}} E_{n,[1,\infty)} = \mathcal{O}(\frac{1}{\sqrt{\varepsilon}} \exp(-c\sqrt{n}))$. To equilibrate both terms, we have to choose $n = \mathcal{O}(\log^2 \varepsilon)$.

[1] In Quantum Chemistry, the functions have infinite support but decay exponentially. Therefore, similar error estimates hold.

6.3 Modification for wavelet applications

Let f be the function which is to be approximated by an exponential sum E_n. There are wavelet applications, where scalar products $\langle f, \psi \rangle$ with wavelets ψ appear. Wavelets have a certain number of vanishing moments, i.e., $\langle p, \psi \rangle = 0$ for all polynomials of degree $\le \ell$ for some $\ell \in \mathbb{N}_0$. In order to keep the moments, one can approximate f by the mixed ansatz

$$\sum_{\nu=1}^{n} \alpha_\nu \exp(-t_\nu x) + \sum_{\nu=0}^{\ell} \beta_\nu x^\nu .$$

Let $u_n^*(x) + p_\ell^*(x)$ be the best approximation of this form in the interval $[a,b] \supset \mathit{support}(\psi)$. By definition we have

$$\langle f, \psi \rangle \approx \langle u_n^* + p_\ell^*, \psi \rangle = \langle v_n^*, \psi \rangle .$$

Therefore, the polynomial part p_ℓ^* need not be stored, and the storage and quadrature costs of $\langle u_n^*, \psi \rangle$ are the same as for the usual best approximation u_n. Of course, the approximation is improved: $\|f - (u_n^* + p_\ell^*)\|_{\infty,[a,b]} \le \|f - u_n\|_{\infty,[a,b]}$.

For an illustration we give the approximation accuracy for $f(x) = 1/\sqrt{x}$ and $n = 4$, $\ell = 1$ in the interval $[1,10]$. The standard approximation is $E_{4,[1,10]} = 2.856 \cdot 10^{-5}$, while the new approach yields the better result $E^*_{4,[1,10]} = \|f - (u_4^* + p_1^*)\|_{\infty,[1,10]} = 2.157 \cdot 10^{-6}$. When these approximations are used after the substitution $x = \|y\|^2 = \sum_{i=1}^d y_i^2$, one has to take into account that $p_\ell^*(\|y\|^2)$ is a polynomial of degree 2ℓ, i.e., a corresponding number of vanishing moments is required. More details can be found in [12].

6.4 Expectation values of the H-atom

In [8, 18] the reduction similar to (1) is applied to the evaluation of expectation values of the H-atom at the ground state. The error is given in terms of the integral

$$4\alpha^2 \int_0^\infty \{v_n(r^2) - r^{-1}\} \, \mathrm{e}^{-2\alpha r} r^2 \mathrm{d}r,$$

where $v(x) = v_n(r^2)$ is an exponential sum that approximates $1/\sqrt{x}$. It is independent of α, if v_n is adapted for each α in the spirit of (36). According to [8, p.138] the asymptotic behaviour is

$$An^{1/2} \exp\left[-\pi\sqrt{\frac{4}{3}n}\right]. \tag{38}$$

We will estimate the more conservative integral

$$\varepsilon_n := 4\int_0^\infty |v_n(r^2) - r^{-1}|\,\mathrm{e}^{-2r} r^2 \mathrm{d}r. \tag{39}$$

for (almost) best approximations v_n. Without loss of generality we set $\alpha = 1$. Specifically, (39) is a weighted L_1 norm, and the treatment is typical for the estimation of weighted L_1 norms of the error [6]. The infinite interval is split into three parts $[0,a]$, $[a,b]$, and $[b,\infty)$. The points a and b are chosen such that the contributions of the first and the third interval are small. A bound for the contribution of $[a,b]$ is determined from the maximal error on this subinterval. Here the results of Section 4 are applied.

We set $a := \beta\sqrt{2n}\exp[-\frac{1}{2}\beta\sqrt{n}]$ and $b := \frac{1}{2}\beta\sqrt{n}$ with β to be fixed later with $\beta \geq 1$. Let v_n be the best approximation or, more generally, be determined by a procedure such that it interpolates $\sqrt{x}$ at $2n$ points in $[a,b]$). In these cases $|v_n - r^{-1}| < r^{-1}$ holds for $x < a$ and $x > b$. Hence,

$$4\int_0^a |v_n(r^2) - r^{-1}|\,\mathrm{e}^{-2r} r^2 \mathrm{d}r \leq 4\int_0^a r\mathrm{d}r = 2a^2 = 4\beta^2 n\exp[-\beta\sqrt{n}].$$

Similarly,

$$\begin{aligned} 4\int_b^\infty |v_n(r^2) - r^{-1}|\,\mathrm{e}^{-2r} r^2 \mathrm{d}r &\leq 4\int_b^\infty \mathrm{e}^{-2r} r\mathrm{d}r = (2b+1)\mathrm{e}^{-2b} \\ &\leq 2\beta\sqrt{n}\exp[-\beta\sqrt{n}]. \end{aligned}$$

Next, set $E := \max_{a\leq r\leq b} |v_n(r^2) - r^{-1}|$ and observe that

$$4\int_a^b |v_n(r^2) - r^{-1}|\,\mathrm{e}^{-2r} r^2 \mathrm{d}r \leq 4E\int_0^\infty \mathrm{e}^{-2r} r^2 \mathrm{d}r = E.$$

The substitution $x = r^2$ shows that we have to consider the approximation on the interval $[a^2, b^2]$. Recalling (36) we apply the guaranteed bound (33) to the best approximation for $R = (b/a)^2 = \frac{1}{8}\exp[-\beta\sqrt{n}]$:

$$\begin{aligned} E = \max_{a^2\leq x\leq b^2} |v_n(x) - \frac{1}{\sqrt{x}}| &\leq \frac{1}{a} 12\exp\left[-\frac{\pi^2 n}{\log(8R)}\right] \\ &\leq \exp[\frac{1}{2}\beta\sqrt{n}]\frac{1}{\beta\sqrt{2n}}\,12\exp\left[-\frac{\pi^2 n}{\beta\sqrt{n}}\right] \\ &\leq 12\exp\left[\frac{1}{2}\beta\sqrt{n} - \frac{\pi^2 n}{\beta}\right]. \end{aligned}$$

Finally we set $\beta = \pi\sqrt{\frac{2}{3}}$ to obtain $E \leq 12\mathrm{e}^{-\beta\sqrt{n}}$. The collection of the integrals yields

$$\varepsilon_n \leq cn\exp\left[-\pi\sqrt{\frac{2}{3}n}\right].$$

This bound is not as good as (38), while the sinc method yields bounds for norms of the error that are not as sharp as the results in Section 4; cf. [7] and §11.

7 Computation of the best approximation

Let $f(x)$ be the function $1/x$ or $1/\sqrt{x}$ to be approximated. We make the ansatz $u_n(x;\{\alpha_\nu\},\{t_\nu\}) = \sum_{\nu=1}^{n} \alpha_\nu \exp(-t_\nu x)$ and define the error

$$\eta_n(x;\{\alpha_\nu\},\{t_\nu\}) := u_n(x;\{\alpha_\nu\},\{t_\nu\}) - f(x).$$

As described in Definition 3.1, the best approximation in the interval $[1,R]$ is characterised by an *alternant* consisting of $2n+1$ points $x_0 < x_1 < \cdots < x_{2n}$ in the interval satisfying the equi-oscillation conditions (10) and (11).

Then $E_{n,[1,R]} := |\eta_n(x_i,\{\alpha_\nu\},\{t_\nu\})|$ is the optimal error $\|u_n(\cdot;\{\alpha_\nu\},\{t_\nu\}) - f\|_{\infty,[1,R]}$ over all $\{\alpha_\nu\},\{t_\nu\}$. The Remez algorithm determines the $2n$ unknown coefficients $\{\alpha_\nu\}$ and $\{t_\nu\}$ from the $2n$ equations $\eta_n(x_i) = -\eta_n(x_{i+1})$. Details of the implementation which we use will follow below.

There is a specific difference between best approximations by polynomials and exponential sums. For polynomials, the error $|\eta_n|$ approaches ∞ as $|x| \to \infty$. Since in our setting $f(x) \to 0$ and $E_n(x;\{\alpha_\nu\},\{t_\nu\}) \to 0$ as $x \to \infty$, the error satisfies $|\eta_n| \to 0$ for $x \to \infty$ (cf. §4.2). As a consequence, for each n there is a unique $R_n > 0$ such that all best approximations in intervals $[1,R]$ with $R \geq R_n$ have the same alternants. In particular, $x_{2n} = R_n$ holds. Hence, best approximations in $[1,R_n]$ are already best approximations in $[1,\infty)$. On the other hand, best approximations in $[1,R]$ with $R < R_n$ satisfy $x_{2n} = R$ and lead to larger errors $|\eta_n(x)| > E_{n,[1,R]}$ for $x > R$ beyond the end of the interval.

From the equi-oscillation property (10) we conclude that there are zeros $\xi_i \in (x_{i-1},x_i)$ of η_n for $1 \leq i \leq 2n$. Formally, we set $\xi_0 := 1$ and $\xi_{2n+1} := R$. Then $\eta_n(x_i)$ is the (local) extremum in the interval $[\xi_i,\xi_{i+1}]$ for $0 \leq i \leq 2n$. Remez-like algorithms start from *quasi-alternants*, i.e., sets of points which satisfy (10), but not yet (11). They replace x_i by the true extrema in $[\xi_i,\xi_{i+1}]$ and try to satisfy $\eta_n(x_i) = -\eta_n(x_{i+1})$ or a relaxed version with updated exponential sums (cf. Remez [21]).

Since the underlying equations are nonlinear, one must use Newton-like methods. The natural choice of parameters of u_n are the coefficients $\{\alpha_\nu\}$ and $\{t_\nu\}$. However variations in these parameters may change the sign structure of $\eta_n = u_n - f$ completely[2], but the Remez algorithms relies on the condition (10). Therefore, we use the zeros ξ_i $(1 \leq i \leq 2n)$ as parameters: $u_n(x;\{\xi_\nu\})$. Since by definition $u_n(\xi_i;\{\xi_\nu\}) = f(\xi_i)$, the function $u_n(\cdot;\{\xi_\nu\})$ can be considered as the interpolating exponential sum.

[2] Note that η_n might be very small, say 1E-12, for good approximations. Then tiny variations of t_ν may yield a new η_n which is completely positive.

In the case of polynomials we have explicit formulae (Lagrange representation) for the interpolating polynomial. Here, we need a secondary Newton process to compute the coefficients $\{\alpha_\nu\} = \{\alpha_\nu(\xi_1,\ldots,\xi_{2n})\}$ and $\{t_\nu\} = \{t_\nu(\xi_1,\ldots,\xi_{2n})\}$. This makes the algorithm more costly, but stability has priority. For the implementation leading to the results in [31] extended precision is used. Then it is possible to determine, e.g., the approximation u_7 of $1/x$ in $[1,2]$, which leads to the error $E_{7,[1,2]} = 8.020\text{E-}15$, which is rather close to machine precision.

We conclude this section with some practical remarks concerning the computation. Once a best approximation u_n is known in some interval $[1,R]$, it can be used as a good starting value for a next interval $[1,R']$ with R' sufficiently close[3] to R. In general, computations with larger R are easier than those with smaller R because the corresponding size of the error η_n. To determine the first u_n for a value n, one should proceed as follows. For rather small n, it is not so difficult to get convergence from reasonable starting values. Assume that u_{n-1} is known (preferably for a larger value of R). The structure of the coefficients $\{\alpha_\nu\},\{t_\nu\}$ and of the zeros ξ_i allow to "extrapolate" for the missing starting values α_n, t_n and ξ_{2n-1}, ξ_{2n}. The search for reasonable starting values becomes extremely simple, if one makes use of the precomputed values in [31].

Appendices

8 Rational approximation of $\sqrt{x}$ on small intervals

The rule for the transformation of the intervals allows us to extend the error bound (3.1) from all powers of 2 to all $n \in \mathbb{N}$, if we verify them for small intervals. Here we can use Newman's trick that was first applied to the approximation of e^x (see [20]). It is based on the following observation. The special product of linear polynomials is a linear and not a quadratic function if considered on the unit circle in $\mathbb{C}$:

$$(z+\beta)(\bar{z}+\beta) = 2\beta\,\Re e\, z + (r^2+\beta^2) \quad \text{if } |z| = 1.$$

Moreover, the winding number of functions on a circle provide additional information that gives rise to estimates from below.

In particular, given $\rho > 1$, we observe that

$$(\rho+z)(\rho+\bar{z}) = \rho^2+1+2\rho x = 2\rho(a+x) \qquad \text{for } |z| = 1,\ x = \Re e\, z,$$

where $a := \frac{1}{2}(\rho+\rho^{-1})$. Setting $f(z) := \sqrt{\rho+z}$, the induced function in the sense of the lemma below is $F(x) = 2\rho\sqrt{a+x}$. The quotient of the arguments at the left

[3] Let ξ_{2n} belong to $[1,R]$. Then $R' > \xi_{2n}$ is required to maintain the quasi-alternant condition (10). If one wants to get immediately the results for $R' < \xi_{2n}$, also the interpolation points ξ_ν must be diminished (e.g., by $\xi'_\nu := (\xi_\nu - 1)\,\frac{R'-1}{R-1} + 1$).

and the right boundary of the unit interval $[-1,+1]$ is

$$\frac{a-1}{a+1} = \left(\frac{\rho-1}{\rho+1}\right)^2. \tag{40}$$

We note that ρ equals the sum of the semi-axes of that ellipse in $\mathbb{C}$ with foci $+1$ and -1 in which $F(x)$ is an analytic function.

We emphasize that the symbols a and b are generic parameters in this appendix. Next, we recall a simple formula for complex numbers: $f\bar{f} - g\bar{g} = 2\Re e[\bar{f}(f-g)] - |f-g|^2$.

Lemma 8.1 (Newman's trick). *Let $r > 0$. Assume that f is a real analytic function in the disk $|z| < 1$ and that $qf - p$ with $p/q \in R_{mn}$ has $m+n+1$ zeros in the disk while q and f have none. Moreover, let $F(x) = f(z)f(\bar{z})$ where $|z| = r$, $\Re e\, z = rx$, Then*

$$2\min_{|z|=r}\left|f\left(f-\frac{p}{q}\right)\right| \le E_{m,n}(F)\ (1+o(1)) \le 2\max_{|z|=r}\left|f\left(f-\frac{p}{q}\right)\right|. \tag{41}$$

Proof. Since we are concerned with the case $|f - p/q| \ll |f|$, we write

$$\begin{aligned} \bar{f}f - \frac{\bar{p}}{\bar{q}}\frac{p}{q} &= 2\Re e[\bar{f}\left(f-\frac{p}{q}\right)] - \left|f-\frac{p}{q}\right|^2 \\ &= 2\Re e[\bar{f}\left(f-\frac{p}{q}\right)]\,(1+o(1)), \end{aligned} \tag{42}$$

and the upper bound follows from the fact that $\bar{p}p/\bar{q}q$ defines a function in $R_{m,n}$.

The lower bound will be derived by using de la Vallée–Poussin's theorem. Note that

$$\Re e\left[\bar{f}\left(f-\frac{p}{q}\right)\right] = \begin{cases} +\left|f\left(f-\frac{p}{q}\right)\right| & \text{if } \arg\left[\bar{f}\left(f-\frac{p}{q}\right)\right] \equiv 0\ (mod\ 2\pi), \\ -\left|f\left(f-\frac{p}{q}\right)\right| & \text{if } \arg\left[\bar{f}\left(f-\frac{p}{q}\right)\right] \equiv \pi\ (mod\ 2\pi). \end{cases} \tag{43}$$

By assumption $f^{-1}q^{-1}(qf-p)$ has $m+n+1$ zeros counting multiplicities but no pole in the disk $|z| < r$. The winding number of this function is $m+n+1$. The argument of $\bar{f}(f-p/q) = f^{-1}q^{-1}(qf-p)|f|^2$ is increased by $(m+n+1)\,2\pi$ when an entire circuit is performed. The argument is increased by $(m+n+1)\,\pi$ as z traverses the upper half of the circle. Since the function is real for $z = +1$ and $z = -1$, we get a set of $m+n+2$ points with sign changes as in (10). By de la Vallée–Poussin's theorem, the degree of approximation cannot be smaller than the minimum of the absolute values at those points, and the proof is complete.

The trick was invented by Newman [20] for deriving an upper bound of the error when e^x is approximated. The application to lower bounds may be traced back to [5]. The treatment of the square root function followed in [3].

A rational approximant $p_n/q_{n-1} \in R_{n,n-1}$ to $f(z) := \sqrt{\rho+z}$ is given by

$$p_n(z) = \tfrac{1}{2}\left\{\left(\sqrt{\rho}+\sqrt{\rho+z}\right)^{2n} + \left(\sqrt{\rho}-\sqrt{\rho+z}\right)^{2n}\right\},$$
$$q_{n-1}(z) = \frac{1}{2\sqrt{\rho+z}}\left\{\left(\sqrt{\rho}+\sqrt{\rho+z}\right)^{2n} - \left(\sqrt{\rho}-\sqrt{\rho+z}\right)^{2n}\right\},$$

and the error can be written in the form

$$\sqrt{\rho+z} - \frac{p_n}{q_{n-1}} = -\frac{\left(\sqrt{\rho}-\sqrt{\rho+z}\right)^{2n}}{q_{n-1}(z)}.$$

The error curve has a zero of order $2n$ at $z=0$. Therefore, p_n/q_{n-1} is a Padé approximant and Newman's trick with $x = \Re e\, z$ (and $r=1$) yields

$$2\rho\sqrt{a+x} - \frac{p_n(z)p_n(\bar{z})}{q_{n-1}(z)q_{n-1}(\bar{z})} = -2\,\Re e\left[\sqrt{\rho+\bar{z}}\,\frac{\left(\sqrt{\rho}-\sqrt{\rho+z}\right)^{2n}}{q_{n-1}(z)}\right](1+o(1))$$
$$= 8\rho\sqrt{a+x}\,\Re e\frac{z^{2n}}{\left(\sqrt{\rho}-\sqrt{\rho+z}\right)^{4n} - z^{2n}}(1+o(1)).$$

Note that $4\rho - 3 \le |(\sqrt{\rho}+\sqrt{\rho+z})^2| \le \rho+3$. Having upper and lower bounds, the winding number $2n$ yields $2n+1$ points close to an alternant. The relative error is

$$E_{n,n-1}\left(\sqrt{a+x}\right) = \frac{4}{(4\rho+\delta)^{2n}}(1+o(1)) \tag{44}$$

with some $|\delta| \le 3$. The parameters a and ρ are related as given by (40). The approximation of $\sqrt{a+x}$ on the unit interval describes the approximation of $\sqrt{x}$ on $[a-1, a+1]$. From (22) and (40) it follows that

$$E_{0,0}(\sqrt{a+x}) = \frac{1}{\rho}. \tag{45}$$

9 The arithmetic-geometric mean and elliptic integrals

Given two numbers $0 < a_0 < b_0$, the common limit $\lim_{j\to\infty} a_j = \lim_{j\to\infty} b_j$ of the double sequence (8) is called the *arithmetic-geometric mean* of a_0 and b_0 and is denoted as $m(a_0, b_0)$. It can be expressed in terms of a complete elliptic integral

$$I(a,b) = \int_0^\infty \frac{dt}{\sqrt{(a^2+t^2)(b^2+t^2)}}. \tag{46}$$

Gauss' crucial observation for establishing the relation between $m(a,b)$ and $I(a,b)$ is that $I(a,b)$ is invariant under the transformation $(a,b) \mapsto (a_1,b_1) = (\sqrt{ab}, \frac{a+b}{2})$.

We see this by the substitution $t = \frac{1}{2}(x - \frac{ab}{x})$. As x goes from 0 to ∞, the variable t increases from $-\infty$ to ∞. Moreover,

$$dt = \frac{x^2+ab}{2x^2}\,dx, \quad t^2 + \left(\frac{a+b}{2}\right)^2 = \frac{(x^2+a^2)(x^2+b^2)}{4x^2}, \quad t^2+ab = \frac{(x^2+ab)}{4x^2}.$$

Hence,

$$I(a_1,b_1) = \frac{1}{2}\int_{-\infty}^{\infty} \frac{dt}{\sqrt{(a_1^2+t^2)(b_1^2+t^2)}} = \int_0^{\infty} \frac{dx}{\sqrt{(a^2+x^2)(b^2+x^2)}} = I(a,b) \quad (47)$$

yields the invariance.

Let $m = m(a,b)$, and set $a_0 = a$, $b_0 = b$. By induction it follows that $I(a_0,b_0) = I(a_j,b_j)$ for all j, and by continuity $I(a_0,b_0) = I(m,m)$. Obviously, $I(m,m) = \int_0^{\infty} \frac{dt}{m^2+t^2} = \frac{\pi}{2m}$, and we conclude that

$$m(a,b) = \frac{\pi}{2I(a,b)}.$$

The *elliptic integrals* are defined by $\mathbf{K}'(k) := I(k,1)$ and $\mathbf{K}'(k) = \mathbf{K}(k')$. Here the module k and the complementary module k' are related by $k^2 + (k')^2 = 1$. A scaling argument shows that

$$I(a,b) = b^{-1}\mathbf{K}'(a/b) \qquad \text{for } 0 < a \le b. \quad (48)$$

Since the arithmetic-geometric mean of 1 and k lies between the arithmetic mean and the geometric mean, we get an estimate that is good for $k \approx 1$.

$$\frac{\pi}{1+k} \le \mathbf{K}'(k) \le \frac{\pi}{2\sqrt{k}}. \quad (49)$$

An estimate that is good for small k is more involved:

$$\begin{aligned}\mathbf{K}'(k) &= 2\int_0^{\sqrt{k}} \frac{dt}{\sqrt{(1+t^2)(k^2+t^2)}} \le 2\int_0^{\sqrt{k}} \frac{dt}{\sqrt{k^2+t^2}} = 2\int_0^{1/\sqrt{k}} \frac{dt}{\sqrt{1+t^2}} \\ &= 2\log\left(\sqrt{\frac{1}{k}} + \sqrt{\frac{1}{k}+1}\right) \le \log\left(4\left(\frac{1}{k}+\frac{1}{2}\right)\right). \end{aligned} \quad (50)$$

As a consequence, we have $(\pi/2)\mathbf{K}'(k)/\mathbf{K}(k) \le \log(\frac{4}{k}+2)$ and

$$\frac{1}{k} \ge \frac{1}{4}\exp\left[\frac{\pi\mathbf{K}'(k)}{2\mathbf{K}(k)}\right] - \frac{1}{2}. \quad (51)$$

Lemma 9.1. *Let λ_0, $\lambda_{-1},\lambda_{-2},\ldots$ be a sequence generated by the Landen transformation and $\kappa_0 := 1/\lambda_0$. Then*

$$\lambda_{-j} \geq \frac{1}{4} \exp\left[2^j \frac{\pi \mathbf{K}'(\kappa_0)}{2\mathbf{K}(\kappa_0)}\right]. \tag{52}$$

Proof. Let $0 < \kappa < 1$ and $\kappa_1 = \frac{2\sqrt{\kappa}}{1+\kappa}$. Note that

$$\kappa_1{}' = \frac{1-\kappa}{1+\kappa}. \tag{53}$$

From (47) and (48) it follows that

$$\mathbf{K}'(\kappa) = I(\kappa, 1) = I\left(\sqrt{\kappa}, \frac{1+\kappa}{2}\right) = \frac{2}{1+\kappa}\mathbf{K}'\left(\frac{2\sqrt{\kappa}}{1+\kappa}\right) = \frac{2}{1+\kappa}\mathbf{K}'(\kappa_1)$$

and with the two means of $1-\kappa$ and $1+\kappa$:

$$\begin{aligned}\mathbf{K}(\kappa_1) = I(\kappa_1', 1) = I\left(\frac{1-\kappa}{1+\kappa}, 1\right) &= (1+\kappa) I(1-\kappa, 1+\kappa) \\ &= (1+\kappa) I\left(\sqrt{1-\kappa^2}, 1\right) = (1+\kappa)\mathbf{K}(\kappa).\end{aligned}$$

Hence,

$$\frac{\mathbf{K}'(\kappa)}{\mathbf{K}(\kappa)} = 2\frac{\mathbf{K}'(\kappa_1)}{\mathbf{K}(\kappa_1)} \tag{54}$$

and by induction $\mathbf{K}'(\kappa_{-j})/\mathbf{K}(\kappa_{-j}) = 2^j \mathbf{K}'/(\kappa_0)\mathbf{K}(\kappa_0)$. Now (51) yields the preliminary estimate

$$\lambda_{-j} \geq \frac{1}{4} \exp\left[2^j \frac{\pi \mathbf{K}'(\kappa_0)}{2\mathbf{K}(\kappa_0)}\right] - \frac{1}{2}.$$

If we apply the estimate to $j+1$ instead of j, return to j noting that $\sqrt{A^2-2} \geq A + 2/A$, we see that we can drop the extra term $1/2$, and the proof is complete.

10 A direct approach to the infinite interval

There is also a one-step proof for the special function $1/x$. It is based on a result of Vjačeslavov [29] which in turn requires complicated evaluations of some special integrals; see also [25]. Since constructions on finite intervals are circumvented, it supports the argument that the non-optimal bound (30) is not induced by the limit process with large intervals.

Given $\alpha > 0$ and $n \in \mathbb{N}$, there exists a polynomial p of degree n with n zeros in $[0,1]$ such that

$$\left| x^\alpha \frac{p(x)}{p(-x)} \right| \leq c_0(\alpha) \cdot \mathrm{e}^{-\pi\sqrt{\alpha n}} \qquad \text{for } 0 \leq x \leq 1.$$

Let p be the polynomial for $\alpha = 1/4$ as stated above. Since $p(\bar{z}) = \bar{p}(z)$, it follows that $p(z)/p(-z) = 1$ for $\Re e\, z = 0$ and

$$\left| \frac{p(z^2)}{p(-z^2)} \right| = 1 \qquad \text{for } \Re e\, z = |\Im m\, z| \geq 0. \tag{55}$$

We consider $P(z) := p^2(1/z^2)$ on the sector $\mathscr{S} := \{z \in \mathbb{C} : \ |\arg z| \leq \pi/4\}$. By construction P has n double zeros in $[1,\infty)$, and from (55) it follows that

$$\left| \frac{P(z)}{P(-z)} \right| = 1 \quad \text{for } z \in \partial\mathscr{S}, \qquad \frac{P(x)}{xP(-x)} \leq \left(c_0(1/4) \cdot \mathrm{e}^{-\pi\sqrt{n/4}} \right)^2 \quad \text{for } x \geq 1.$$

Now let u_n be the exponential sum interpolating $1/x$ and its first derivative at the (double) zeros of P. Since $1/x - u_n$ has no more zeros than the specified ones, we have $u_n(x) \leq 1/x$ for $x \geq 0$. Hence,

$$|u_n(z)| \leq u_n(\Re e\, z) \leq 1/\Re e\, z \leq \sqrt{2}/|z| \qquad \text{on the boundary of } \mathscr{S}.$$

Arguing as in Section 2, we introduce the auxiliary function $g(z) := (\frac{1}{z} - u_n(z)) z \frac{P(-z)}{P(z)}$. We know that $|g(z)| \leq 1 + \sqrt{2}$ holds on the boundary of $\mathscr{S}$ and therefore in $\mathscr{S}$. Finally,

$$\left| \frac{1}{z} - u_n(z) \right| = \left| g(z) \frac{P(z)}{zP(-z)} \right| \leq (1+\sqrt{2})\, c_0^2(1/4) \mathrm{e}^{-\pi\sqrt{n}}.$$

11 Sinc quadrature derived approximations

The sinc quadrature discussed in this section approximates integrals of the form $\int_{-\infty}^{\infty} F(t)\mathrm{d}t$ under certain conditions on F. In particular, we are interested in functions that depend on a further parameter x like $F(t,x) = F_1(t)\exp[F_2(t)x]$, and the evaluation at a quadrature point $t = \tau_\nu$ yields $\alpha_\nu \mathrm{e}^{-t_\nu x}$ with $\alpha_\nu := F_1(\tau_\nu)$ and $t_\nu := F_2(\tau_\nu)$. Therefore the sinc quadrature applied to

$$f(x) := \int_{-\infty}^{\infty} F(t,x)\mathrm{d}t \tag{56}$$

is a popular method to obtain exponential sums even with guaranteed upper bounds [8, 18]. Concerning literature we refer to the monograph of Stenger [26] or [14, Anhang D]. Next, we introduce the sinc quadrature rule $T(F,h)$, its truncated form $T_N(f,h)$, and its application to $1/x$ (the application to $1/\sqrt{x}$ is quite similar).

The sinc function $\mathrm{sinc}(x) := \frac{\sin(\pi x)}{\pi x}$ is an analytic functions with the value one at $x = 0$ and zero at $x \in \mathbb{Z}\backslash\{0\}$. Given a step size $h > 0$, the family of functions

$$S_{k,h}(x) := \mathrm{sinc}(\frac{x}{h} - k) \qquad (k \in \mathbb{Z}),$$

satisfies $S_{k,h}(\nu h) = \delta_{k\nu}$ ($\delta_{k\nu}$: Kronecker symbol). Let $F \in C(\mathbb{R})$ decay sufficiently fast for $x \to \pm\infty$. Then the sum

$$F_h(x) := \sum_{k\in\mathbb{Z}} F(kh) S_{k,h}(x)$$

converges and interpolates F at all grid points $x = \nu h \in h\mathbb{Z}$. This fact suggests the interpolatory quadrature rule $\int_{-\infty}^{\infty} F(t)\mathrm{d}t \approx \int_{-\infty}^{\infty} F_h(t)\mathrm{d}t$. Since $\int_{-\infty}^{\infty} \mathrm{sinc}(t)\mathrm{d}t = 1$, the right-hand side leads to the *sinc quadrature rule*

$$T(F,h) := h \sum_{k\in\mathbb{Z}} F(kh)$$

for $\int_{-\infty}^{\infty} F(t)\mathrm{d}t$, and $T(f,h)$ can be considered as the infinite trapezoidal rule. The next step is the truncation (cut-off) of the infinite series to the finite sum

$$T_N(f,h) := h \sum_{k=-N}^{N} F(kh).$$

For convenience, we will use N as truncation parameter. It will be related to the number n of terms in (2) by $n = 2N+1$.

Before we discuss the quadrature error of $T(F,h)$, we show how to get exponential sums from this approach. As example we consider the representation of $\frac{1}{x}$ by $\int_0^{\infty} \exp(-xs)\mathrm{d}s$. Let $s = \varphi(t)$ be any differentiable transformation of $(-\infty,\infty)$ onto $(0,\infty)$. This yields the integral

$$\frac{1}{x} = \int_{-\infty}^{\infty} \exp(-x\varphi(t))\varphi'(t)\mathrm{d}t \tag{57}$$

to which the sinc quadrature $T_N(f,h)$ can be applied:

$$\frac{1}{x} \approx T_N(\exp(-x\varphi(\cdot))\varphi'(\cdot),h) = h \sum_{k=-N}^{N} \varphi'(kh)\mathrm{e}^{-x\varphi(kh)}.$$

Obviously, the right-hand side is the exponential sum (2) with $\alpha_\nu = h\varphi'((\nu-1-N)h)$ and $t_\nu = \varphi((\nu-1-N)h)$ for $1 \le \nu \le n = 2N+1$. Note that different transformations φ yield different exponential sums.

A good candidate for φ is $\varphi(t) := \exp(t)$ leading to[4]

$$\frac{1}{x} = \int_{-\infty}^{\infty} \exp(-x\mathrm{e}^t)\mathrm{e}^t\mathrm{d}t. \tag{58}$$

The exponential behaviour $t_\nu = const \cdot \mathrm{e}^{\nu h}$ of the coefficients is sometimes used as an explicit ansatz for (2). Indeed, the coefficients t_ν of the best approximations from

[4] Also $\varphi(t) = \exp(At)$ for $A > 0$ is possible. The reader may try to analyse the influence of A to the error analysis.

Section 4 lead to similar quotients $t_{\nu+1}/t_\nu$ for ν in the middle range with deviations for ν close to 1 and n.

Next we study the quadrature error of T_N. It is the sum of $\int_{-\infty}^{\infty} F(t)\mathrm{d}t - T(F,h)$ and $T(F,h) - T_N(F,h)$. The quadrature error of the sinc quadrature

$$\eta(F,h) := \left| \int_{-\infty}^{\infty} F(t)\mathrm{d}t - T(F,h) \right|$$

tends to zero as $h \to 0$. The truncation error $|T(F,h) - T_N(F,h)|$ vanishes as $N \to \infty$. Both discretisation parameters h and N will be related in such a way that both errors are (asymptotically) equal.

The analysis of $\eta(F,h)$ requires holomorphy of F in a stripe. Let

$$D_d := \{z = x + iy : x \in \mathbb{R}, |y| < d\} \subset \mathbb{C}$$

be the open stripe along the real axis with width $2d$. The function F is assumed to be holomorphically extendable to D_d such that the L^1 integral

$$\|F\|_{D_d} := \int_{\mathbb{R}} \{|F(x+id| + |F(x-id|\}\,\mathrm{d}x$$

over the boundary of D_d exists and is finite. As proved in [26, p. 144 f] the error $\eta(F,h)$ of the infinite quadrature rule $T(F,h)$ is bounded by

$$\eta(F,h) \le \|F\|_{D_d} \exp[-2\pi d/h]. \tag{59}$$

The truncation error $|T(F,h) - T_N(F,h)|$ equals $h\left|\sum_{|k|>N} F(kh)\right|$ and depends on the decay of F as $x \to \pm\infty$ (note that this concerns only the behaviour on the real axis, not in the stripe D_d).
If, for instance, $|F(t)| \le c \cdot \mathrm{e}^{-\alpha|t|}$ holds, then $|T(F,h) - T_N(F,h)| \le (2c/\alpha)\mathrm{e}^{-\alpha Nh}$ follows. In this case, the error $\eta(F,h) = \mathcal{O}(\mathrm{e}^{-2\pi d/h})$ and the truncation error $\mathcal{O}(\mathrm{e}^{-\alpha Nh})$ are asymptotically equal if $-2\pi d/h = -\alpha Nh$, i.e., $h = \sqrt{2\alpha\pi dN}$. This leads to the estimate of the total error

$$\left| \int_{-\infty}^{\infty} F(t)\mathrm{d}t - T_N(F,h) \right| \le \left(\|F\|_{D_d} + \frac{2c}{\alpha} \right) \exp[-\sqrt{2\pi d/(\alpha N)}] \quad \text{for } h = \sqrt{2\alpha\pi dN}. \tag{60}$$

So far, F is a function of t only and the integral $\int_{-\infty}^{\infty} F(t)\mathrm{d}t$ is a real number. Now we replace F by $F(t,x)$ as in (56) such that the integral defines a function $f : D \to \mathbb{R}$ on a domain D. The error estimate (60) is still correct, but it holds only pointwise for $x \in D$. We note that $\|F\|_{D_d}$ becomes a function $\|F(\cdot,x)\|_{D_d}$ of x, and even the width d of the stripe may change with x. Moreover, if decay inequality $|F(t,x)| \le c \cdot \mathrm{e}^{-\alpha|t|}$ holds with x-dependent factors $c(x)$ and $\alpha(x)$, also these quantities in (60) become variable. We have to take care that the estimate (60) (with $\|F(\cdot,x)\|_{D_d}$ replaced by an upper bound) is *uniform* in $x \in D$, and the error $|f(x) - T_N(F(\cdot,x),h)|$ is uniform too.

In the following, we will simply write $F(t)$ instead of $F(t,x)$, i.e., $F(t)$ is understood to be function-valued.

We apply this strategy to the error estimation of the integral in (58). The integrand $F(t) = \exp(-x\mathrm{e}^t)\mathrm{e}^t$ is an entire function in t, and to obtain a bounded norm $\|F\|_{D_d}$ we choose $d < \pi/2$. Then $|F(t \pm id)| = \exp(-x\mathrm{e}^t \cos(d))\mathrm{e}^t$ implies $\|F\|_{D_d} = \frac{1}{x\cos(d)}$. Inequality (59) yields

$$|\eta(F,h)| \leq \frac{\exp(-2\pi d/h)}{x\cos(d)} \qquad \text{for all } 0 < d < \pi/2.$$

Optimisation with respect to d yields $d = \arctan(2\pi/h) < \pi/2$ and

$$|\eta(F,h)| \leq \frac{\sqrt{1+(2\pi/h)^2}}{x} \exp\Big(\frac{-2\pi \arctan(2\pi/h)}{h}\Big).$$

Concerning $|T(F,h) - T_N(F,h)| = h\left|\sum_{|k|>N} F(kh)\right|$ notice the different behaviour of $F(kh)$ for $k \to \infty$ and $k \to -\infty$. As $k \to -\infty$, the factor e^{kh} describes a uniformly exponential decay, while $\exp(-x\mathrm{e}^{kh}) \to 1$. For $k \to +\infty$, the factor $\exp(-x\mathrm{e}^{kh})$ shows a doubly exponential behaviour which, however, depends on the value of x. The precise asymptotics are given by

$$\begin{aligned} h\left|\sum_{k<-N} F(kh)\right| &\leq h \sum_{k<-N} \mathrm{e}^{kh} \leq \int_{-\infty}^{-Nh} \exp(t)\mathrm{d}t = \mathrm{e}^{-Nh}, \\ h\left|\sum_{k>+N} F(kh)\right| &\leq h \sum_{k>N} \exp(-x\mathrm{e}^{kh})\mathrm{e}^{kh} \\ &\leq \int_{Nh}^{\infty} \exp(-x\mathrm{e}^t)\mathrm{e}^t \mathrm{d}t = \frac{1}{x}\exp(-x\mathrm{e}^{Nh}). \end{aligned}$$

Here we assume $x\mathrm{e}^{Nh} \geq 1$ for the second inequality, so that the function $\exp(-x\mathrm{e}^t)\mathrm{e}^t$ is monotonously decreasing in $[Nh,\infty)$. Altogether, we get the following error estimate between the integral (58) and the exponential sum $T_N(F,h)$

$$\begin{aligned} \left|\frac{1}{x} - T_N(F,h)\right| \leq{}& \frac{\sqrt{1+(2\pi/h)^2}}{x} \exp\left(\frac{-2\pi \arctan(2\pi/h)}{h}\right) \\ &+ \mathrm{e}^{-Nh} + \frac{1}{x}\exp(-x\mathrm{e}^{Nh}). \end{aligned} \tag{61}$$

To simplify the analysis, we assume $x \in [1,R]$, which implies the relation

$$\frac{1}{x}\exp(-x\mathrm{e}^{Nh}) \leq \mathrm{e}^{-Nh-1},$$

i.e., the last term in (61) is smaller than the second one. Further, we use the asymptotic behaviour $2\pi \arctan(2\pi/h) = \pi^2 - h + \mathscr{O}(h^3)$ to show that

$$\exp\frac{-2\pi\arctan(2\pi/h)}{h} = \mathcal{O}(\exp(-\pi^2/h)).$$

Therefore, the right-hand side in (61) becomes $\mathcal{O}(\sqrt{1+(2\pi/h)^2}\exp(-\pi^2/h)) + \mathcal{O}(\exp(-Nh))$. The asymptotically best choice of h is $h=\pi/\sqrt{N}$ which leads to equal exponents: $-\pi^2/h = -Nh = -\pi\sqrt{N}$. Inserting this choice of h, we get the uniform estimate

$$\left|\frac{1}{x} - T_N(F,h)\right| \le \left(\mathcal{O}(1)+2\sqrt{N}\right)\mathrm{e}^{-\pi\sqrt{N}} \le \mathcal{O}\left(\sqrt{n}\exp\left[-\frac{\pi}{\sqrt{2}}\sqrt{n}\right]\right) \quad \text{for all } x \ge 1, \tag{62}$$

where the last expression uses the number $n = 2N+1$ of terms in $T_N(F,h)$. The exponential behaviour $\exp[-\pi\sqrt{n/2}]$ is not as good as $\exp[-\pi\sqrt{n}]$ from (30).

Although we get the same behaviour $\exp[-const\sqrt{n}]$ of the error as in (30), the reason is a different one. In the case of the best approximation, we could show an error behaviour $\exp[-const\cdot n]$ for finite intervals $[1,R]$, whereas $\exp[-\pi\sqrt{n}]$ was caused by the unboundedness of $[1,\infty)$. The $\exp[-\pi\sqrt{n/2}]$ behaviour of the sinc quadrature is independent of the choice $x\in[1,R]$, R finite, or $x\in[1,\infty)$. Even if we restrict x to a single point x_0, the error is like in (62). The reason for $\exp[-const\sqrt{n}]$ in the sinc case is due to the fact that we have to equalise the exponents in $\mathcal{O}(\exp(-const/h)) + \mathcal{O}(\exp(-const\cdot Nh))$. The error $\mathcal{O}(\exp(-const/h))$ of the infinite sinc quadrature can hardly be improved (see (59)), but the truncation error $\mathcal{O}(\exp(-const\cdot Nh))$ of $|T(F,h)-T_N(F,h)|$ depends on the decay behaviour of F. If, for instance, $|F(t)| \le c\cdot\exp(-\alpha|t|^\gamma)$ holds for some $\gamma>1$, this faster decay yields the smaller truncation error $\mathcal{O}(\exp(-\alpha(Nh)^\gamma))$. Finally, $h=\mathcal{O}(N^{-\gamma/(\gamma+1)})$ leads to the total error $\exp\left[-const\cdot n^{\gamma/(\gamma+1)}\right]$. For large γ, the exponent comes close to $-const\cdot n$.

An even better decay behaviour is the doubly exponential decrease $|F(t)| \le c_1\cdot\exp(-c_2\mathrm{e}^{c_3|t|})$. In this case, the total error can be estimated by $\mathcal{O}\left(\|F\|_{D_d}\exp\left(\frac{-2\pi dc_3N}{\log(2\pi dc_3N)}\right)\right)$ (cf. [14, Satz D.4.3]). To obtain a doubly exponential decay, one can follow the following lines: Start with an integral $\int_{-\infty}^{\infty}F(t)\mathrm{d}t$, where F has the usual exponential asymptotic $|F(t)| \le c\cdot\exp(-\alpha|t|)$. Then apply the transformation $t=\sinh\tau$. The new integral is $\int_{-\infty}^{\infty}G(\tau)\mathrm{d}\tau$ with the doubly exponential integrand $G(\tau)=F(\sinh\tau)\cosh\tau$. The drawback is that one must ensure that G is still holomorphic in a stripe D_d and that $\|F\|_{D_d}$ is finite. The mentioned transformation applied to $F(t)=\exp(-x\mathrm{e}^t)\mathrm{e}^t$ from above does not succeed. For any $d>0$ the real part of $\mathrm{e}^{\sinh\tau}$ may be negative in D_d and, because of the exponentially increasing function $\exp(-x\mathrm{e}^{\sinh(\tau+id)})$, the integral with respect to $\tau\in\mathbb{R}$ does not exist, i.e. $\|F\|_{D_d}=\infty$.

A possible approach is to replace the first transformation $\varphi(t) := \exp(t) : [0,\infty) \to (-\infty,\infty)$ in (57) by $\varphi(t) := \log(1+\exp(\sinh t))$, which yields

$$\frac{1}{x} = \int_{-\infty}^{\infty} \exp\left(-x\log\left(1+\mathrm{e}^{t}\right)\right) \frac{\mathrm{d}t}{1+\mathrm{e}^{-t}}; \tag{63}$$

cf. [14, §D.4.3.2]. The integrand $F = \exp\left(-x\log\left(1+\mathrm{e}^{t}\right)\right)/\left(1+\mathrm{e}^{-t}\right)$ in (63) behaves simply exponential for $t \to \infty$ and $t \to -\infty$. Thanks to this property, the second transformation $t = \sinh\tau$ succeeds in providing an integrand G which is holomorphic in D_d for $d < \pi/2$ with finite norm $\|G\|_{D_d}$. However, pointwise finiteness $\|G(\cdot,x)\|_{D_d} < \infty$ is not enough. It turns out that in general $\|G(\cdot,x)\|_{D_d} \leq \mathcal{O}(\mathrm{e}^{x})$, which destroys the error estimates. For $x \in [1,R]$ one has to reduce the stripe D_d to the width $d = d(R) := \mathcal{O}(1/\log R)$. Then involved estimates show that the error $\left|\frac{1}{x} - T_N(G,h)\right|$ in $1 \leq x \leq R$ is bounded by

$$\mathcal{O}\left(\exp\left(-\frac{2\pi d(R)N}{\log(2\pi d(R)N}\right)\right) \qquad \text{with} \quad d(R) := \mathcal{O}(1/\log R)$$

(cf. [14, §D.4.3.2]). Since a detailed analysis shows $d(R) = \frac{\pi}{2}\frac{1}{\log(3R)} - \mathcal{O}(\log^{-2}(3R))$, this estimate is almost of the form $\exp(-Cn)$ with $C := \frac{2\pi^2}{\log(3R)\log(2\pi^2 n/\log(3R))}$ and may be compared with $\exp(-C^* n)$ from (29) with $C^* = \frac{\pi^2}{\log(8R)}$. Obviously, $C < C^*$ holds for sufficiently large n, but even for small n, $C < C^*$ holds, e.g., for $R \leq 1600$ ($n = 4$), $R \leq 24700$ ($n = 5$), $R \leq 3.7_{10}5$ ($n = 6$), $R \leq 5.5_{10}6$ ($n = 7$), and $R \leq 8.1_{10}7$ ($n = 8$). The latter bounds of R are (much) larger than the value $R = \frac{1}{8}\exp[\pi\sqrt{n}]$ introduced in the line before (30). Hence, for values of R for which the best approximation on $[1,R]$ is not already a best approximation for $[1,\infty)$, (29) gives a better result than the sinc estimate from above.

References

1. J. Almlöf: *Elimination of energy denominators in Møller-Plesset perturbation theory by a Laplace transform approach*. Chem. Phys. Lett. **176** (1991), 319–320
2. P.Y. Ayala and G. Scuseria: *Linear scaling second-order Moller-Plesset theory in the atomic orbital basis for large molecular systems*. J. Chem. Phys. **110** (1999), 3660
3. J.M. Borwein and P.B. Borwein: *Pi and the AGM*. John Wiley & Sons, 1987.
4. D. Braess: *Nonlinear Approximation Theory*. Springer-Verlag, Berlin, 1986.
5. D. Braess: *On the conjecture of Meinardus on the rational approximation of e^x*. J. Approximation Theory **36** (1982), 317–320.
6. D. Braess: *Asymptotics for the approximation of wave functions by exponential sums*. J. Approximation Theory **83** (1995), 93–103.
7. D. Braess and W. Hackbusch: *Approximation of 1/x by exponential sums in [1,∞)*. IMA J. Numer. Anal. **25** (2005), 685–697
8. E. Cancès, M. Defranceschi, W. Kutzelnigg, C. Le Bris, and Y. Maday: *Computational quantum chemistry: a primer*. In: 'Handbook of Numerical Analysis', **X**, pp. 3–270, C. Le Bris (ed.), Elsevier, Amsterdam 2003
9. A.F. Izmaylov and G.E. Scuseria: *Resolution of the identity atomic orbital Laplace transformed second order Møller–Plesset theory for nonconducting periodic systems*. Phys. Chem. Chem. Phys. **10** (2008), 3421–3429

10. Y. Jung, R.C. Lochan, A.D. Dutoi, and M. Head-Gordon: *Scaled opposite-spin second order Møller–Plesset correlation energy: An economical electronic structure method*. J. Chem. Phys. **121** (2004), 9793
11. L. Grasedyck: *Existence and computation of low Kronecker-rank approximations for large linear systems of tensor product structure*. Computing, **72** (2004), 247–265
12. W. Hackbusch: *Approximation of* $1/\|x-y\|$ *by exponentials for wavelet applications*. Computing, **76** (2006), 359–366
13. W. Hackbusch: *Efficient convolution with the Newton potential in d dimensions*. Numer. Math. **110** (2008), 449–489
14. W. Hackbusch: *Hierarchische Matrizen – Algorithmen und Analysis*. Springer-Verlag, Berlin (to appear in 2009)
15. M. Häser and J. Almlöf: *Laplace transform techniques in Møller–Plesset perturbation theory*. J. Chem. Phys. **96** (1992), 489
16. M. Häser: *Møller-Plesset (MP2) perturbation theory for large molecules*. Theor. Chim. Acta **87** (1993), 147–173
17. M. Kobayashi and H. Nakai: *Implementation of Surján's density matrix formulae for calculating second-order Møller–Plesset energy*. Chem. Phys. Lett. **420** (2006), 250–255
18. W. Kutzelnigg: *Theory of the expansion of wave functions in a Gaussian basis*. Int. J. of Quantum Chemistry **51** (1994), 447–463.
19. D.S. Lambrecht, B. Doser, and C. Ochsenfeld: *Rigorous integral screening for electron correlation methods*. J. Chem. Phys. **123** (2005), 184102
20. D.J. Newman: *Rational approximation to* e^x. J. Approximation Theory **27** (1979), 234–235
21. E.J. Remez: *Sur un procédé convergent d'approximations successives pour déterminer les polynômes d'approximation*. Compt. Rend. Acad. Sc. **198** (1934), 2063–2065
22. R. Rutishauser: *Betrachtungen zur Quadratwurzeliteration*. Monatshefte Math. **67** (1963), 452–464
23. D. Kats, D. Usvyat and M. Schütz: *On the use of the Laplace transform in local correlation methods*. Phys. Chem. Chem. Phys., 2008, DOI: 10.1039/b802993h
24. S. Schweizer, B. Doser, and C. Ochsenfeld: *An atomic orbital-based reformulation of energy gradients in second-order Møller–Plesset perturbation theory*. J. Chem. Phys. **128** (2008), 154101
25. H.R. Stahl: *Best uniform rational approximation of* x^α *on* $[0,1]$. Acta Math. **190** (2003), 241–306.
26. F. Stenger: *Numerical Methods Based of Sinc and Analytic Functions*. Springer-Verlag, New York 1993
27. P.R. Surján: *The MP2 energy as a functional of the Hartree–Fock density matrix*. Chem. Phys. Lett. **406** (2005), 318–320
28. A. Takatsuka, S. Ten-no, and W. Hackbusch: *Minimax approximation for the decomposition of energy denominators in Laplace-transformed Møller-Plesset perturbation theories*. J. Chem. Phys. **129** (2008), 044112
29. N.S. Vjačeslavov: *On the least deviation of the function* sign$x$ *and its primitives from the rational functions in the* L_p *metrics,* $0 < p < \infty$. Math. USSR Sbornik **32** (1977), 19–31
30. A.K. Wilson and J. Almlöf: *Møller-Plesset correlation energies in a localized orbital basis using a Laplace transform technique*. Theor. Chim. Acta **95** (1997), 49–62
31. Webpages `www.mis.mpg.de/scicomp/EXP_SUM/1_x/` and `ldots/1_sqrtx/` with explanations in `.../1_x/tabelle` and `.../1_sqrtx/tabelle`
32. E.I. Zolotarov: *Application of elliptic functions to questions of functions deviating least and most from zero* (Russian). Zap. Imp. Akad. Nauk (1877). St. Petersburg 30 no. 5; reprinted in collected works II, pp. 1–59. Izdat, Akad. Nauk SSSR, Moscow 1932.