SIAM J. NUMER. ANAL.
Vol. 59, No. 6, pp. 2801–2822

RECIPROCAL-LOG APPROXIMATION AND PLANAR PDE SOLVERS*

YUJI NAKATSUKASA† AND LLOYD N. TREFETHEN†

Abstract. This article is about both approximation theory and the numerical solution of partial differential equations (PDEs). First we introduce the notion of *reciprocal-log* or *log-lightning approximation* of analytic functions with branch point singularities at points $\{z_k\}$ by functions of the form $g(z) = \sum_k c_k/(\log(z - z_k) - s_k)$, which have N poles potentially distributed on different sheets of a Riemann surface. We prove that the errors of minimax reciprocal-log approximations decrease exponentially with respect to N and that exponential or near-exponential convergence (i.e., at a rate $O(\exp(-CN/\log N))$) also holds for near-best approximations constructed by linear least-squares fitting on the boundary with suitably chosen preassigned singularities. We then apply these results to derive a "log-lightning method" for the numerical solution of Laplace and related PDEs in two-dimensional domains with corner singularities. The convergence is near-exponential, in contrast to the root-exponential convergence for the original lightning methods based on rational functions.

Key words. reciprocal-log approximation, rational approximation, lightning solver

AMS subject classifications. 65N35, 41A30, 65E05

DOI. 10.1137/20M1369555

1. Introduction.

In this paper we introduce a new problem in approximation theory. On the interval $(-\infty, 0]$, it is known that functions $F(s)$ like e^{as} $(\mathrm{Re}\, a > 0)$ with an essential singularity at $s = -\infty$ can be approximated by rational functions $r(s)$ with exponential convergence as a function of n, the degree of the rational function. Our starting point in section 2 is the change of variable $z = e^s$ from $s \in (-\infty, 0]$ to $z \in [0, 1]$. This gives us exponentially convergent approximations on $[0, 1]$ of functions $f(z)$ like z^a $(a > 0)$ with branch point singularities at $z = 0$ by functions of the form $g(z) = r(\log(z))$. In the generic case where the poles $\{s_k\}$ of r are distinct, the approximation can be written in the partial fractions form

$$g(z) = c_0 + \sum_{k=1}^{n} \frac{c_k}{\log(z) - s_k}. \tag{1.1}$$

We show that approximations of this kind can be computed by linear least-squares fitting, and we note the equioscillatory characterization of the error in the case of minimax (best ∞-norm) approximations. With both best and near-best approximations, one encounters the startling property that the oscillations typically cluster double-exponentially near the singularity, readily coming as close (in theory) as a distance of, say, 10^{-1000}. This is in contrast to the well-known case of rational approximation, where the clustering is just exponential [20].

Section 3 generalizes the discussion to domains in the complex z-plane with $m \geq 1$ singularities, typically at corners. Here we consider approximations of the generic form

*Received by the editors September 25, 2020; accepted for publication (in revised form) July 13, 2021; published electronically November 1, 2021. A paper written after this one but published earlier has confirmed the effectiveness of the log-lightning method for PDE problems. See P. Baddoo and L. N. Trefethen, Log-lightning computation of capacity and Green's function, *Maple Trans.*, 1 (2021), p. 1.

https://doi.org/10.1137/20M1369555

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

$$g(z) = \sum_{j=1}^{m} \sum_{k=1}^{n_j} \frac{c_{jk}}{\log(z - z_j) - s_{jk}} + p_0(z), \tag{1.2}$$

where p_0 is a polynomial of degree n_0 (possibly 0), with total number of degrees of freedom $N = 1 + \sum_{j=0}^{m} n_j$. The least-squares method extends to this case too and again gives exponentially or near-exponentially convergent approximations. (By near-exponential, we mean at a rate $O(\exp(-CN/\log N))$ with $C > 0$.) This and other claims are first explained heuristically and illustrated numerically. Then they are established rigorously by a succession of theorems in section 4, deriving accuracy estimates from Hermite integrals, and section 5, showing how a problem with several singularities can be decomposed into single-singularity problems by Cauchy integrals over open arcs.

As is typical with computations based on redundant expansions, numerical instability resulting from the use of ill-conditioned matrices may be a problem, and a crucial part of our numerical algorithm is the use of the Vandermonde with Arnoldi orthogonalization technique presented in [1]. The MATLAB code `VAorthog.m` we use for this, copied from [2], is listed in the appendix. This technique does not eliminate the ill-conditioning (unless one has just a single singularity and no polynomial term), but it diminishes its impact greatly, and in particular often keeps the norms of the coefficient vectors in reasonable bounds, an issue connected with the subject of regularization [7].

Another numerical matter, more specific to the reciprocal-log approximation context, is mentioned in section 6. This is a consequence of the "10^{-1000}" effect mentioned above, and we propose that it may be a good idea to constrain the singular part of a reciprocal-log approximation to approach zero more rapidly at each singularity.

Up to this point, the paper is all approximation theory and algorithms. The motivation, however, is numerical solution of PDEs by a "log-lightning" analogue of the lightning solvers for Laplace, Helmholtz, and biharmonic equations presented in [2, 5, 6]. An initial experiment is presented in section 7 to establish that, as intended, these approximations can be used to derive exponentially convergent new methods for solving PDEs in planar domains with corner singularities.

In a word, this is a paper about approximation by functions with logarithmic branch points rather than just pole singularities. What starts out as a seemingly arbitrary rearrangement of familiar rational approximation theory turns out to have potentially significant consequences for scientific computing. Much more will need to be done to make the method into a practical PDE solver, but we have included PDEs in the title of the paper because this motivation is at the heart of this work.

A publication of Driscoll and Fornberg in 2001 proposed the introduction of a logarithmic term to approximate functions with jumps, with coefficients computed in the manner of quadratic Hermite–Padé approximation [4]. Their technique is different from what is being proposed here, but the use of logarithms to resolve singularities gives a family resemblance.

2. Approximation on $[0, 1]$. A classical problem in approximation theory is the rational minimax approximation of $F(s) = e^s$ on the interval $(-\infty, 0]$. To be precise, for each n, let us consider rational functions $r(s)$ of degree n, meaning that $r(s)$ can be written as $p(s)/q(s)$ for degree n polynomials p and q. In the generic case where the poles are distinct and finite, r can be written in partial fractions form as

$$r(s) = c_0 + \sum_{k=1}^{n} \frac{c_k}{s - s_k} \tag{2.1}$$

with coefficients $\{c_k\}$ and poles $\{s_k\}$. Of course, some rational functions have confluent poles (i.e., poles of order > 1) and cannot be written in this form, and we shall sometimes make use of such approximations.

Cody, Meinardus, and Varga showed in 1969 [3] that approximations $r(s)$ to $F(s) = e^s$ exist with exponential convergence, meaning $\|F - r\| = O(C^{-n})$ as $n \to \infty$ for some $C > 1$, where $\|\cdot\|$ is the ∞-norm on $(-\infty, 0]$. The optimal constant is $C = 1/H = 9.28903\ldots$, where H is a number related to elliptic integrals known as Halphen's constant; for a discussion with references, see Chapter 25 of [18]. Exponential convergence of best rational approximations at the same asymptotic rate applies to $F(s) = e^{as}$ for any a with $\mathrm{Re}\, a > 0$ and to certain other functions $F(s)$ related to e^{as} by rational transformations [15]. Moreover, convergence at this rate occurs not just on $(-\infty, 0]$ but throughout the complex plane [15].

By the change of variables $z = e^s$ and the definitions $f(z) = F(s)$ and $g(z) = r(s)$, we transplant this theory to a new setting. Now the domain is $z \in [0, 1]$, and f is a function such as z^a that may have a branch point singularity at $z = 0$. What is new is that (2.1) now takes the form

$$g(z) = c_0 + \sum_{k=1}^{n} \frac{c_k}{\log(z) - s_k}, \tag{2.2}$$

where log denotes the standard branch of the logarithm. Thus, by the transplantation of existing theory, we know that a wide class of functions on $[0, 1]$ with branch point singularities at $z = 0$ can be approximated by functions of the form (2.2) with exponential convergence at the rate $O((9.28903\ldots)^{-n})$.

More important for scientific computing are approximations that are not best but readily computable, still with exponential convergence though not at the optimal rate. One way to derive such approximations is by expressing $F(s)$ as an integral along a Hankel contour winding around $(-\infty, 0]$ in the s-plane. If the integral is approximated by a truncation of the transplanted equispaced trapezoidal rule on the real axis, one obtains rational approximations (2.1) to $F(s)$, hence reciprocal-log approximations (2.2) to $f(z)$, which converge at rates $O(C^{-n})$ with values of C on the order of 3 or 4 depending on details of contours and parameters. The leader in developing these methods has been Weideman [22, 26]; for a review, see section 15 of [21].

The approach we shall take is a variation on this theme motivated by the lightning PDE solvers of [2, 5, 6] and section 7. Rather than deriving coefficients from the trapezoidal rule, we will obtain them by solving a linear least-squares problem. Typically this will be posed on a discrete subset of $[0, 1]$, with exponential clustering of sample points at $z = 0$. We have not attempted an analysis of the conditions that make for effective sampling points, since in practice this part of the problem poses little difficulty so long as there are plenty of points and they are clustered somewhat more densely near corners than the singularities being employed in the approximations, as discussed in [5].

Experiments show that an effective choice of singularities $\{s_k\}$ for our approximations lies on a parabolic contour in the style of [22, 26],

$$s_k = \frac{n}{4}(1 + i\theta_k)^2, \quad \theta_k = -\pi + 2\pi(k - \tfrac{1}{2})/n, \quad 1 \le k \le n. \tag{2.3}$$

For simplicity, however, we will also sometimes use a configuration in which all the points s_k are taken to be confluent at a single value s_0 of scale $O(n)$; we typically take $s_0 = n/2$ or $n/6$. In this case (2.2) changes to the confluent form

$$g(z) = p\left(\frac{1}{\log(z) - s_0}\right), \tag{2.4}$$

where p is a polynomial of degree n. As we shall prove in section 4, these methods give exponential convergence as $n \to \infty$.

We illustrate these ideas with a MATLAB code segment for approximation of $\sqrt{z}$ on $[0,1]$ whose fifth line realizes the distribution (2.3). Similar results are obtained for other functions with branch point singularities such as $z^{3/4}$, $\sqrt{z}\cos(10z)$, $z\log z$, etc. The line invoking `VAorthog` (see the appendix) is a numerically stable alternative to `A = [z.^0 1./(log(z)-sk)]`: both matrices A have the same column space, but with `VAorthog` the columns are orthogonal.

```
M = 2000;
z = logspace(-50,0,M)';
f = sqrt(z);
n = 10;
kk = 1:n; tk = -pi + 2*pi*(kk-.5)/n; sk = (n/4)*(1+1i*tk).^2;
[~,A] = VAorthog(log(z),0,{sk});
c = A\f;
g = A*c;
```

The maximum error on $[0,1]$ is 1.3×10^{-4}, and this number decreases exponentially as $n \to \infty$, as shown in the first image of Figure 2.1. There is a marked improvement in comparison with the root-exponential behavior of the lightning rational approximation of [5]. The second image shows the singularities $\{s_k\}$ in the case $n = 32$. Note that although the geometry in the s-plane is just an arc, the corresponding points $z_k = \exp(s_k)$ in the z domain lie along a spiral on the Riemann surface of $g(z)$ associated with its definition in terms of the function $\log(z)$, clustering exponentially at both ends at $z = 0$. With $n = 32$, the last value in the upper half s-plane is $s_{32} \approx -66.1 + 48.7i$. The corresponding value z_{32} has modulus 1.0×10^{-31} and argument 15.5π, putting it 8 sheets above the main branch of the Riemann surface of $g(z)$.

The second row of Figure 2.1 shows corresponding results for the confluent-pole approximation (2.4) with $s_0 = n/2$. These poles are further from optimal, and the exponential convergence is about 15% slower. Certain fine points associated with confluent computations are discussed in section 6. To minimize the effects very near the singularity discussed there and make the comparison of Figure 2.1 a fair one, the least-squares fitting for this figure was done on a grid of 2000 points logarithmically spaced in $[10^{-100}, 1]$.

Figure 2.2 shows error curves $g(z) - \sqrt{z}$, $z \in [10^{-20}, 1]$ for approximations (2.2). On the left these are the approximations computed by least-squares fitting with the singularities $\{s_k\}$ of (2.3), as in the first row of Figure 2.1. Note that the maximal error decreases exponentially with n and that the error is much smaller for $z \approx 0$ than for $z \approx 1$, because the clustered grid puts more weight there. The right column shows error curves for approximations with the same $\{s_k\}$ whose coefficients have been modified to achieve L^∞ optimality by 20 steps of a linear Lawson iteration, also known as iteratively reweighted least-squares [9]. These equioscillatory error curves are elegant, but we doubt whether the "optimal" approximations are superior in practice. Their maximal error is lower by only about a factor of 2, at the cost of greatly worsening the accuracy for $z \approx 0$. See Chapter 16 and Myth 3 of Appendix A in [18].

The poles of our approximations $g(z)$ lie exponentially close to the origin in the z-plane, at least when the formula (2.3) is used. Exponential clustering of poles is familiar in rational approximation theory [20]. Another phenomenon appears in Figure 2.2 that is unfamiliar, however, and that is *doubly exponential* clustering of

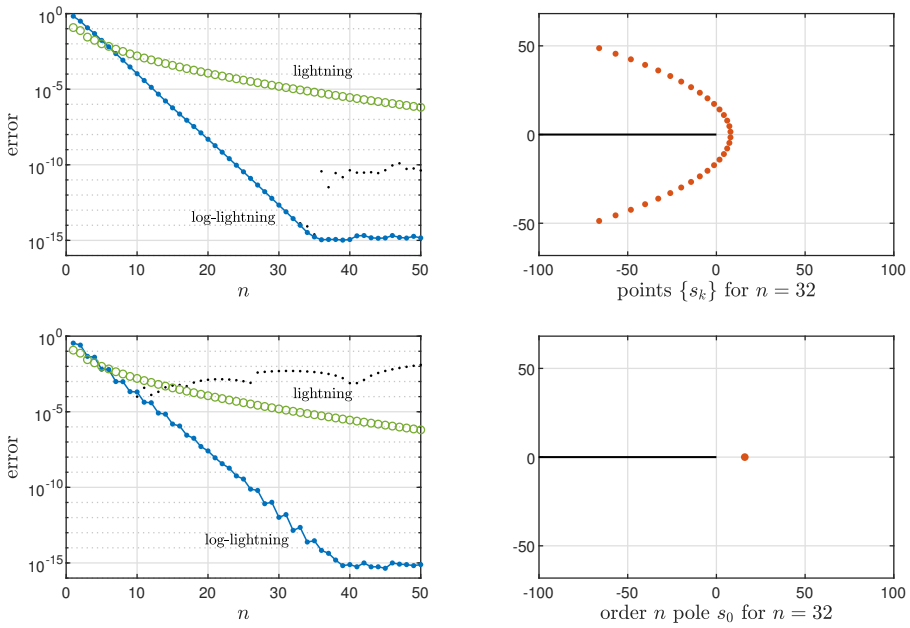

FIG. 2.1. *Exponential convergence of approximations* (2.2)–(2.3) *(top row) and* (2.4) *(bottom row) to* $\sqrt{z}$ *computed by least-squares fitting on an exponentially clustered grid in* $[0, 1]$. *(These data are obtained with Arnoldi stabilization* [1]*; the numbers obtained without stabilization are shown by the small dots.) A strong contrast is apparent with the root-exponential convergence of lightning rational approximations* [5] *for the same problem. The right column shows the points* $\{s_k\}$ *or* s_0 *in the* s*-plane for* $n = 32$. *Since the imaginary parts* $\operatorname{Im} s_k$ *extend far beyond* $[-\pi, \pi]$ *in the first case, the values* $\exp(s_k)$*, which are the poles of* $g(z)$*, lie on many different sheets of a Riemann surface.*

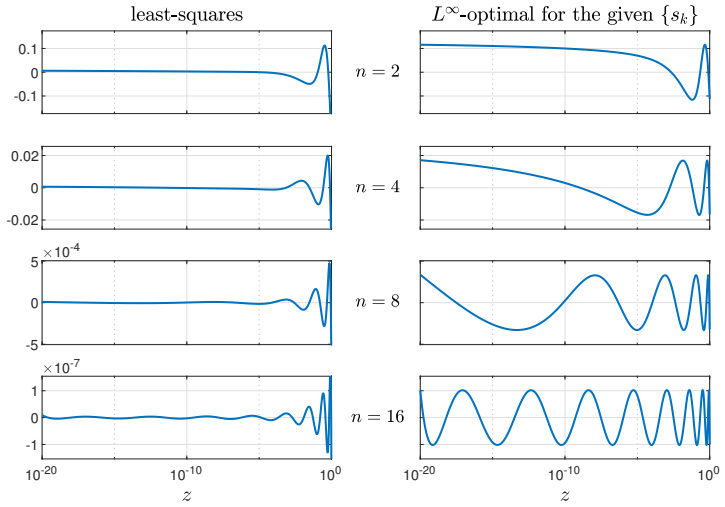

FIG. 2.2. *Error curves for reciprocal-log approximations* (2.2) *of* $\sqrt{z}$ *on* $[0, 1]$ *with* $n = 2, 4, 8, 16$ *with singularities* s_k *given by* (2.3). *In the left column, coefficients* $\{c_k\}$ *are computed by least-squares fitting on an exponentially clustered grid of* 1000 *points from* 10^{-20} *to* 1. *In the right column, the fits are then improved to* L^∞ *optimal for the given* $\{s_k\}$ *by a Lawson iteration (iteratively reweighted least-squares)* [9]. *This improves the maximal error by about a factor of* 2 *at the cost of eliminating the enhanced accuracy for* $z \approx 0$. *Each pair shares a common vertical scale, with each curve showing* $n + 1$ *extrema of alternating sign.*

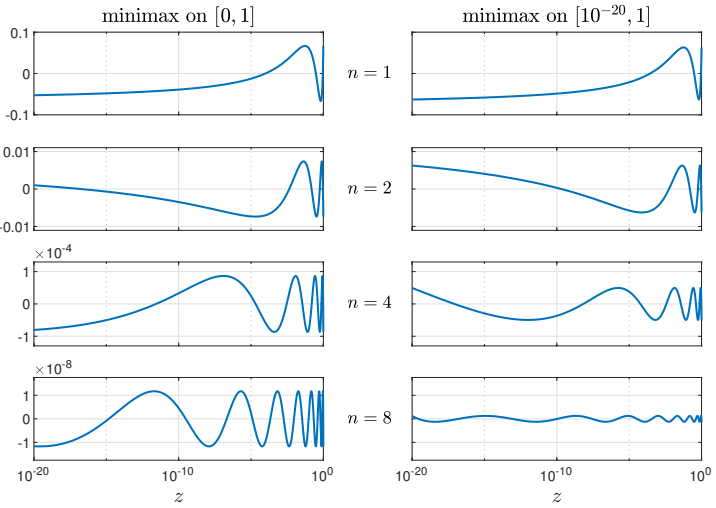

FIG. 2.3. *Error curves for true minimax reciprocal-log approximations of* $\sqrt{z}$ *on* $[0,1]$ *with* $n = 1, 2, 4, 8$ *on* $[0,1]$ *(left) and* $[10^{-20},1]$ *(right). Note that these values of* n *are half those of Figure* 2.2*, reflecting the roughly squared accuracy of minimax approximations. Again each pair shares a common vertical scale, and now there are* $2n+2$ *equioscillating extrema.*

oscillating extrema, visible especially in the Lawson column on the right. To see this, consider the leftmost minimum of each error curve in the right column of the figure. With $n = 2$, 4, and 8, these minima lie at approximately $z = 10^{-1.5}$, $10^{-4.5}$, and (as shown by computations on a longer interval) 10^{-25}. A rough model of such behavior would be $z_{\min} = \exp(-1.7^n)$, which suggests values for $n = 16$ and 32 on the order of 10^{-2000} and $10^{-10,000,000}$. In the figure, we see nothing like this, since the computations are carried out on $[10^{-20},1]$, and it would be impossible to resolve much beyond $[10^{-300},1]$ in standard IEEE floating point arithmetic. However, this probably does not matter, as the approximants on $[0,1]$ will not differ much from those associated with a truncated interval such as $[10^{-20},1]$. We discuss these matters further in section 6.

All these approximants are linear, based on poles s_k fixed a priori. What about true minimax approximations, with $\{s_k\}$ chosen optimally to minimize the error? For the current setting of $[0,1]$, everything is known about such approximations since they are transplants of minimax rational approximants on $(-\infty,0]$ [18, Chapter 24]. As illustrated in Figure 2.3, which was computed by the method outlined on pp. 217–219 of [18], the errors will be about the squares of what we have seen before, and now there are $2n+2$ equioscillatory extreme points. We shall not discuss such approximations further, because they become impractical once one moves to planar domains.

3. Approximation on a planar domain. We now turn to domains in the complex plane. In the basic case one has a closed Jordan region E with piecewise analytic boundary Γ and a function f continuous in E and analytic in the interior. We suppose that f is analytic on Γ except for branch point singularities at a finite set of points $\{z_j\}$, $1 \le j \le m$, which may, for example, be corners. As mentioned in the introduction, we consider approximations to f of the generic form

$$g(z) = \sum_{j=1}^{m} \sum_{k=1}^{n_j} \frac{c_{jk}}{\log(z - z_j) - s_{jk}} + p_0(z), \tag{3.1}$$

where p_0 is a degree n_0 polynomial, with $N = 1 + \sum n_j$. Each function $\log(z - z_j)$ denotes a fixed choice of a branch of the logarithm that is continuous in E. Since standard hardware puts the branch of the logarithm on the negative real axis, it is often convenient in a computer code to work with rotated terms of the form $\log(e^{i\theta_j}(z - z_j))$ for certain real numbers θ_j. (Experienced programmers in the complex plane know that branch cuts can be a source of confusion, but so far as we are aware they do not pose any computational obstacles to these approximations in principle, even for approximation on nonconvex domains.) We also employ the confluent variant

$$g(z) = \sum_{j=1}^{m} p_j\left(\frac{1}{\log(z - z_j) - s_j}\right) + p_0(z), \tag{3.2}$$

where each p_j is a polynomial of degree n_j. In both (3.1) and (3.2), we often take $n_0 = 0$, so that p_0 reduces to a constant.

The essential point, to be proved in the next two sections, is that exponential or near-exponential convergence is guaranteed so long as the numbers n_j (for $j \geq 1$) increase in proportion to N as $N \to \infty$, in the absence of rounding errors of course. In this section we will explore just one example defined on the domain shown in Figure 3.2, a two-thirds bite of the unit disk. The function is

$$f(z) = z\log(-\tfrac{1}{2}z) \cdot (1 - z/\omega)^{1/2} \cdot (1 - z/\bar{\omega})^{3/2} \tag{3.3}$$

with $\omega = e^{\pi i/3}$, with branch point singularities at each of the corners. The boundary Γ is discretized along each segment by 500 points exponentially clustered down to distances from the corners of about 10^{-14}.

As indicated in Figure 3.1, we computed three approximations to f for values of N up to 150. One curve shows the accuracy of AAA rational approximations of degree $N - 1$. (AAA approximations are rational approximations that are computed very quickly by an algorithm described in [12], implemented in the `aaa` command of Chebfun; we calculated them with the "cleanup" option turned off.) The AAA approximations are close to optimal when the degree is not too large, and this curve

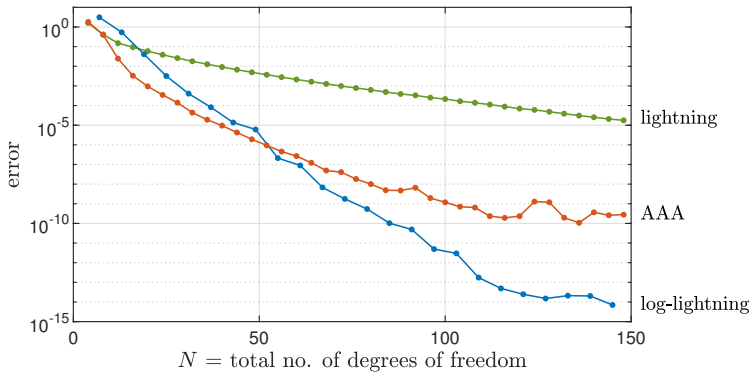

FIG. 3.1. *Convergence of three approximants to the function f of* (3.3) *on the domain E shown in Figure* 3.2. *One sees root-exponential convergence of the AAA and lightning approximations and a reasonable approximation to exponential convergence of the log-lightning approximations. The main reason the lightning curve is the smoothest is that its errors are so far from machine epsilon.*

shows root-exponential convergence down to an error on the order of 10^{-8}. Another curve shows lightning rational approximations [5] of the form

$$r(z) = \sum_{j=1}^{3} \sum_{k=1}^{n_j} \frac{c_{jk}}{z - s_{jk}} + p_0(z) \tag{3.4}$$

with $\sum_{j=0}^{3} n_j = N - 1$, where p_0 is a polynomial of degree n_0; we take $n_0 = N/4 - 1$ and $n_j = N/4$ for $j = 1, 2, 3$. (The choice $n_0 = 0$, which is often a good one for (3.1) and (3.2), tends to slow down convergence significantly for (3.4).) The poles were fixed at distances $d_{jk} = \exp(4(\sqrt{k} - \sqrt{n_j}))$, $1 \le k \le n_j$ from the corners z_j [5, equation (3.2)]. Note that the convergence is again root-exponential, at a rate five or six times slower than with AAA with its adaptively determined poles. Clearly AAA has some advantages, but it has drawbacks of sometimes greater sensitivity to rounding errors and occasional placement of poles inside a domain where they are not wanted. More fundamentally, as we shall discuss in section 7, lightning methods generalize immediately to approximation of the real part for solving Laplace problems, whereas AAA does not.

The final curve in Figure 3.1 shows log-lightning approximations (3.2) with $s_j = n_j/6$, now with $n_0 = 0$ and $n_j = (N-1)/3$ for $j = 1, 2, 3$. The convergence appears approximately exponential down to a level of around 10^{-13}.

Figure 3.2 displays phase portraits [25] for f and its three approximations with $N = 100$. The rational approximations simulate branch cuts by strings of poles,

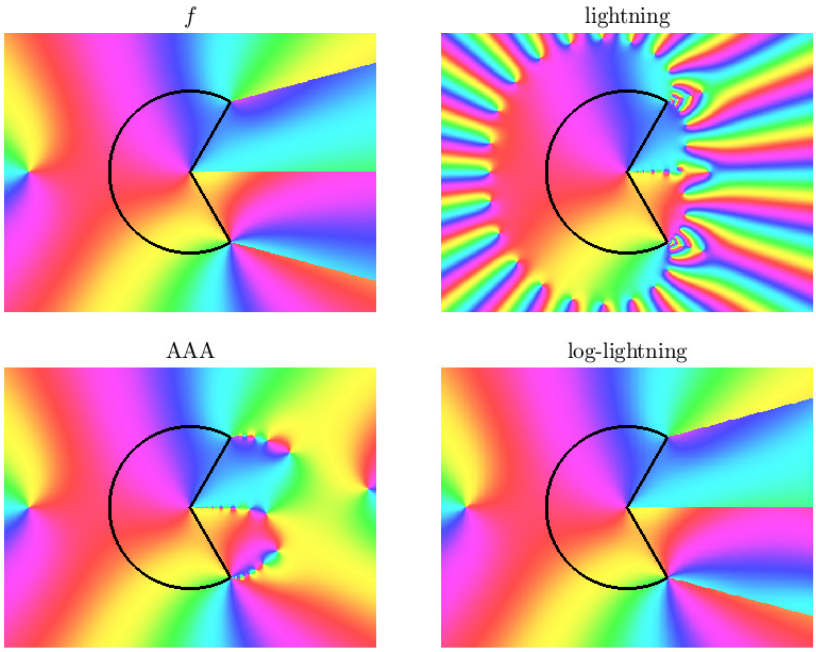

FIG. 3.2. *Phase portraits of the function f of (3.3) and three approximants with $N \approx 100$ degrees of freedom. All the approximations give many digits of accuracy in E. Outside E, the branch cuts of f are approximated by strings of poles for lightning and AAA approximants and by true branch cuts for the log-lightning approximant. The circus tent effect in the lightning plot reflects the polynomial term p_0 in (3.4).*

whereas the reciprocal log approximation incorporates true branch cuts. For f itself, the shapes of the branch cuts are arbitrary. If we had coded the function differently, their directions could have been altered, for example, but this would have had no effect on the values of f on Γ and therefore on the three approximations and the other three images. For the lightning and log-lightning approximations, the forms of the branch cuts are fixed by our computation: they have been set to be straight lines bisecting the exterior angles, the natural default choice in the absence of other information. For the AAA approximation, the branch cuts are presumably more nearly optimal, curved in a fashion analyzed by Stahl for the case of Padé approximation [14].

4. Convergence theorems: One singularity. In this section we consider the reciprocal-log approximation of a function $f(z)$ with a branch point singularity at $z = 0$. By arguments adapted from those of section 2 of [5], we show that if f can be analytically continued along any contour in the complex plane avoiding $z = 0$, then exponentially convergent approximations exist with the singularities $\{s_k\}$ extending a distance $O(n)$ into the right half-plane (Theorem 4.2). If the analyticity assumption holds just in a bounded neighborhood of 0, then near-exponentially convergent approximations exist with $\{s_k\}$ of bounded positive real part (Theorem 4.6).

To make these ideas precise the theorems of this and the next section use the following definition, which covers the familiar singularities of type $z^a(\log z)^b$ [10, 24]. The first condition, (4.1), essentially asserts that f can be analytically continued along arbitrary contours in a punctured neighborhood of each singularity z_k. But all we really need is continuation along curves that spiral in logarithmically to z_k, so that after the $s = \log(z - z_k)$ change of variables they lie in a V-shaped wedge in the s-plane, and that is why (4.1) takes the form it does. The second condition, (4.2), asserts that f is Hölder continuous at z_k, so that with $s = \log(z - z_k)$ we get exponential decay as $\mathrm{Re}\, s \to -\infty$. A function like $(z - z_k)^a \log(z - z_k)$, however, will grow slowly as z winds around z_k, and that is why (4.2) includes the factor involving $\arg(z - z_k)$.

DEFINITION 4.1. *Let E be a closed set in the complex plane, which may or may not have a nonempty interior. A function f is* analytic on E with branch point singularities at $z_1, \ldots, z_m$ *on the boundary if f is analytic in the interior of E and can be analytically continued to a neighborhood of each boundary point of E that is not among the z_k; and if moreover there is a neighborhood of each z_k within which f can be analytically continued to a multivalued function $\tilde{f}$ along any curve that avoids z_k and satisfies*

$$|\arg(z - z_k)| \le |z - z_k|^{-\tau} \tag{4.1}$$

for some constant $\tau > 0$, with $\tilde{f}$ satisfying

$$|\tilde{f}(z) - \tilde{f}(z_k)| \le c|z - z_k|^a(1 + |\arg(z - z_k)|^b) \tag{4.2}$$

for some constants $a, b, c > 0$.

For our first theorem, it is enough to use confluent singularities and consider approximations of the form (2.4) with $s_0 = \sigma n$,

$$g(z) = p\left(\frac{1}{\log(z) - \sigma n}\right), \tag{4.3}$$

where p is a polynomial of degree n and $\sigma > 0$ is a constant. This result is sharp in the sense that faster than exponential convergence is in general not possible, as we

know from the equivalence of the problem of reciprocal-log approximation of z^a on $[0,1]$ to that of rational approximation of e^{as} on $(-\infty,0]$. The condition (4.4) looks almost like a repeat of (4.2), but (4.2) is a local condition, only needing to apply for z near z_k, whereas (4.4) applies even as $z \to \infty$. With $s = \log z$, it is designed to ensure that $F(s)$ is under control in the right half of the s-plane as well as the left.

THEOREM 4.2. *Let E be a compact set in $\mathbf{C}\backslash(-\infty,0)$ with 0 on the boundary; more generally, E might be a multilevel surface wrapping around 0 a finite number of times. Let f be an analytic function on E with a branch point singularity at $z=0$ in the sense of Definition* 4.1*, and assume further that f can be analytically continued to a multivalued function $\tilde{f}$ along all curves in the z-plane avoiding $z=0$, satisfying in addition to* (4.2)

$$|\tilde{f}(z)| \le C(1+|z|^{a'})(1+|\arg z|^{b'}) \tag{4.4}$$

for some constants $a', b', c' > 0$. Then for all sufficiently small $\sigma > 0$, there exist functions of the form (4.3) *satisfying*

$$\|f-g\| \le \exp(-Cn) \quad \forall n \tag{4.5}$$

for some $C > 0$, where $\|\cdot\|$ is the supremum norm on E.

By the change of variables $s = \log z$ and subtraction of $f(0)$, Theorem 4.2 is equivalent to the following theorem on rational approximation. As just mentioned, (4.6) constrains $F(s)$ in the left half s-plane, and (4.7) in the right. The function $e^s + se^{2s}$, for example, satisfies the conditions with $a < 1$, $a' > 2$, and $b, b' > 1$.

THEOREM 4.3. *Let E be a subset of $\mathbf{C}$ with $|\mathrm{Im}\, E|$ bounded and $\mathrm{Re}\, E$ bounded from above. Let F be an entire function of s satisfying for all sufficiently small* $\mathrm{Re}\, s$

$$|F(s)| \le ce^{a\,\mathrm{Re}\, s}(1+|\mathrm{Im}\, s|^b) \tag{4.6}$$

and for all s

$$|F(s)| \le c'(1+e^{a'\,\mathrm{Re}\, s})(1+|\mathrm{Im}\, s|^{b'}) \tag{4.7}$$

for some constants $a, b, c, a', b', c' > 0$. Then for all sufficiently small $\sigma > 0$, there exist type $(n-1,n)$ rational functions r of the form

$$r(s) = p\left(\frac{1}{s-\sigma n}\right), \tag{4.8}$$

where p is a polynomial of degree n, with

$$\|F-r\| \le \exp(-Cn) \quad \forall n \tag{4.9}$$

for some $C > 0$, where $\|\cdot\|$ is the supremum norm on E.

Thus our job is to prove Theorem 4.3. The proof will be based on the Hermite integral formula for rational interpolation [23, Theorem 2, Chapter 8]. Neither the poles $\{s_k\}$ nor the interpolation points $\{\alpha_k\}$ in the statement below need to be distinct, and an interpolation point of multiplicity $\nu > 1$ is interpreted, as usual, to mean interpolation at that point of $f, f', \ldots, f^{(\nu-1)}$. By a Hankel contour, we mean a continuous curve winding counterclockwise around $(-\infty,0]$ from ∞ to ∞.

LEMMA 4.4 (Hermite integral formula). *Let F be an analytic function of s satisfying* (4.6) *on and inside a Hankel contour Γ in the s-plane. Let interpolation*

points $\alpha_1, \dots, \alpha_n \in (-\infty, 0]$ and poles $s_1, \dots, s_n$ anywhere in the s-plane be given; neither $\{\alpha_j\}$ nor $\{s_j\}$ need to be distinct. Let r be the unique type $(n-1, n)$ rational function with poles at $\{s_j\}$ that interpolates F at $\{\alpha_j\}$. Then for any s enclosed by Γ,

$$F(s) - r(s) = \frac{1}{2\pi i} \int_\Gamma \frac{\phi(s)}{\phi(t)} \frac{F(t)}{t-s}\, dt, \tag{4.10}$$

where

$$\phi(s) = \prod_{j=1}^{n} \frac{s-\alpha_j}{s-s_j}. \tag{4.11}$$

In proving Theorem 4.3, we will suppose $E = (-\infty, 0]$ without loss of generality; the estimates change only by constant factors for other domains since $|\mathrm{Im}\, E|$ and $\mathrm{Re}\, E$ are assumed to be bounded whereas the poles of (4.8) and the interpolation points of (4.12) below scale in proportion to n as $n \to \infty$. We will use this choice of interpolation points in $(-\infty, 0]$:

$$\alpha_j = n\sigma \left(\frac{1+\gamma_j}{1-\gamma_j} \right)^2, \quad \gamma_j = e^{i\pi j/(n+1)}, \quad 1 \le j \le n, \tag{4.12}$$

where σ is the parameter in the denominator of (4.8). These are derived from potential theory [11, 23], where poles and interpolation points are interpreted as point charges of opposite signs, and a minimal-energy charge configuration gives asymptotically optimal approximations as $n \to \infty$. The function

$$s = n\sigma \left(\frac{1+\gamma}{1-\gamma} \right)^2 \tag{4.13}$$

maps the unit disk in the γ-plane to the s-plane slit along $(-\infty, 0]$, with $\gamma = 0$ mapping to $s = n\sigma$. Thus the equilibrium distribution of interpolation points on the unit circle corresponding to poles at $\gamma = 0$, namely, the uniform distribution, maps to the distribution determined by (4.13) in the s-plane slit along $(-\infty, 0]$ with poles at $s = n\sigma$, as in the function (4.8). (The $\{\alpha_j\}$ are called *Fejér–Walsh points* [16].) The following lemma, illustrated in Figure 4.1, was provided to us by Peter Baddoo.

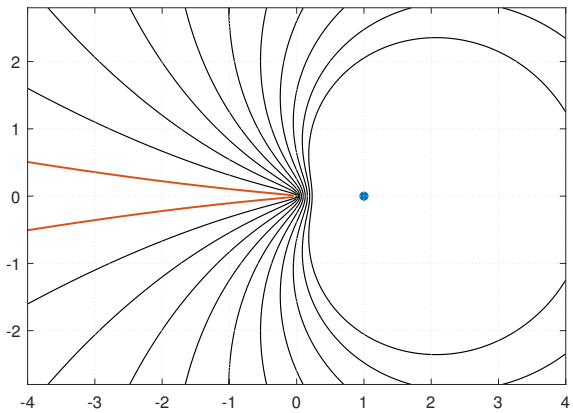

FIG. 4.1. *Level curves $n^{-1} \log |\phi(ns)| = 0, 0.1, \dots, 1$ for $n = 100$ to illustrate Lemma 4.5. As n increases, the contours converge to radial lines, with the level 0 contour (red) narrowing toward $(-\infty, 0]$. Thus $|\phi(ns)|^{1/n}$ approaches a value >1 at each point $s \in \mathbf{C} \backslash \{(-\infty, 0] \cup \sigma\}$ and is bounded by 1 for $s \in (-\infty, 0]$.*

LEMMA 4.5. *With α_j defined by* (4.12), *$s_j = n\sigma$, and s related to γ by* (4.13), *the function $\phi(s)$ of* (4.11) *can be written as*

$$\phi(s) = \frac{\gamma^{-n-1} - \gamma^{n+1}}{(n+1)(\gamma^{-1} - \gamma)} = \frac{\gamma^{-n} + \gamma^{-n+2} + \cdots + \gamma^{n}}{n+1}. \tag{4.14}$$

For $s \in (-\infty, 0]$,

$$|\phi(s)| \le 1. \tag{4.15}$$

For $s \in \mathbf{C}\backslash\{(-\infty, 0] \cup \sigma\}$,

$$|\phi(ns)|^{1/n} \to e^{u(s)} \quad \text{as } n \to \infty, \tag{4.16}$$

where

$$u(s) = |\log|\gamma|\,| > 0, \tag{4.17}$$

with uniform convergence on compact subsets of $\mathbf{C}\backslash\{(-\infty, 0] \cup \sigma\}$.

Proof. The formulas (4.15)–(4.17) follow readily from (4.14), so (4.14) is what must be established. First we note that each complex number $s \ne n\sigma$ corresponds under (4.13) to two numbers γ and γ^{-1}, and the right-hand side of (4.14) gives the same value for both, so (4.14) does indeed define a function of s. This is a meromorphic function with n zeros at the points α_j and a pole of order n, it can be verified, at $s = n\sigma$, corresponding to $\gamma = 0$. But at $s = \infty$, corresponding to $\gamma = 1$, this function is analytic. Thus it is a degree n rational function with the same zeros and poles as $\phi(s)$ as defined by (4.11), and since it takes the value 1 at $s = \infty$, it must be the same function. ∎

Proof of Theorem 4.3. As mentioned above, we take $E = (-\infty, 0]$ without loss of generality, and we choose Γ for Lemma 4.4 as a V-shape passing through $n\sigma$ at a fixed acute angle, as sketched in Figure 4.2. (Any acute angle will suffice.) We divide Γ into a "head" in the right half-plane, a "tail" with $\mathrm{Re}\, t \le -n$, and a "middle" with $-n \le \mathrm{Re}\, t \le 0$, and show that their contributions I_{head}, I_{tail}, and I_{middle} to (4.10) are each exponentially small. (A quantity depending on n is *exponentially small* if it is $O(\exp(-Cn))$ as $n \to \infty$ for some $C > 0$ and *exponentially large* if its reciprocal is

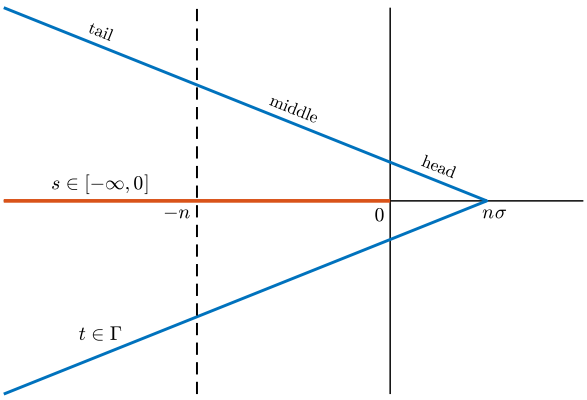

FIG. 4.2. *Hankel contour Γ for the proof of Theorem* 4.3 *with $E = (-\infty, 0]$, with all poles at $s = n\sigma$. In the tail, $F(t)$ in* (4.10) *is exponentially small. In the middle, $F(t)$ is at most algebraically large and $1/\phi(t)$ is exponentially small. In the head, $F(t)$ is exponentially large, but its exponential growth rate is limited by σ, so if σ is small enough, $F(t)/\phi(t)$ is still exponentially small.*

exponentially small.) Note that the denominator $t-s$ of (4.10) is bounded below in absolute value, $|\phi(s)|$ is bounded by 1 by (4.15), and 2π is just a constant. Thus it is enough to show that each of the three integrals

$$\int \frac{|F(t)|}{|\phi(t)|}\,dt \tag{4.18}$$

is exponentially small. In the tail, $F(t)$ is exponentially small and decreases exponentially as $\mathrm{Re}\,t \to -\infty$ by (4.6), whereas $1/|\phi(t)|$ is bounded by Lemma 4.5, since (4.16) and (4.17) ensure $|\phi(t)| \geq 1$ for all sufficiently large n (cf. Figure 4.1); thus I_{tail} is exponentially small. In the middle, $F(t)$ grows at most algebraically with n by (4.6), whereas $1/|\phi(t)|$ is exponentially small by Lemma 4.5 again, since in this range the values $u(s)$ on the right-hand side of (4.16) are bounded above 0 and thus the values $\exp(u(s))$ are bounded above 1, making their nth powers exponentially large; thus I_{middle} is exponentially small. In the head, $F(t)$ is exponentially large, but by (4.7), at most $O(\exp(A\sigma n))$. On the other hand, by Lemma 4.5 again, $1/|\phi(t)|$ is exponentially small. In particular, $|\phi(t)| > \exp(dn)$ uniformly in this region for some $d > 0$ that is independent of σ. It follows that if $\sigma < d/A$, then I_{head} too is exponentially small. ∎

We now turn to the second theorem about approximation near $z = 0$, assuming analyticity only in a bounded domain. As before, our prototypical example for discussion without loss of generality is $E = [0, 1]$. For the argument we again use a V-shaped contour Γ for the Hermite integral, as shown in Figure 4.3, but now, Γ must be fixed rather than growing with n. This necessitates a change in the choice of poles. If the poles are all at the apex, then the same argument as before ensures exponentially small contributions to the error from I_{head} and I_{middle}, but no longer so small (just root-exponential) from I_{tail}. Instead it is necessary to distribute poles along Γ, and a successful strategy is to distribute them along segments extending a distance $\sigma L = \sigma\rho n$ from the apex on the two sides of Γ, for any constant $\rho > 0$.

Specifically, one way to distribute poles is by solving a potential theory problem in the infinite V domain. As suggested by the contour lines $u = 0.1, 0.2, \ldots, 0.9$ in Figure 4.3, let u be the unique harmonic function in the infinite slit wedge taking values $u = 0$ on $(-\infty, 0]$ and $u = 1$ on the segments of length σL just mentioned, with a homogeneous Neumann condition on the remainder of the boundary. This function u can be calculated by means of a conformal map of the lower half of the slit wedge onto a rectangle $-K < \mathrm{Re}\,w < K$, $0 < \mathrm{Im}\,w < K'$ in the complex w-plane, with the Dirichlet boundary components mapping to the vertical sides and the Neumann components to the top and bottom; the ratio $K'/2K$ is known as the *conformal module* of the rectangle. If the interior angle of the wedge is π/μ for some $\mu > 1$, the map has the explicit representation

$$s(w) = \sigma - \sigma\left(\frac{b+y}{1+2b-y}\right)^{\mu}, \quad y = \mathrm{sn}(w\,|\,M^{-2}), \tag{4.19}$$

where sn is the Jacobi elliptic sine function and

$$M = 1 + 2b, \quad b = L^{\mu} + \sqrt{L^{\mu} + L^{2\mu}}. \tag{4.20}$$

The numbers K and K' are the complete elliptic integrals of the first kind with parameters M^{-2} and $1 - M^{-2}$, respectively [13]. Our Fejér–Walsh choice of poles and interpolation points for the theorem will be

$$s_k = s(w_k) \quad w_k = K + i(k - \tfrac{1}{2})K'/n, \quad 1 \leq k \leq n \tag{4.21}$$

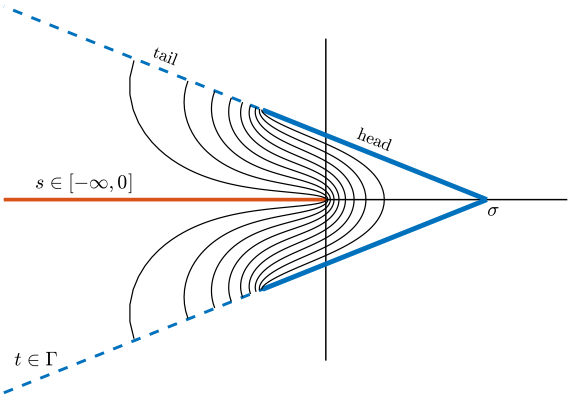

FIG. 4.3. *Hankel contour* Γ *for the proof of Theorem* 4.6 *with* $E = [0,1]$*, together with level curves* $u = 0.1, \dots, 0.9$ *for the associated potential theory problem obtained from the conformal map* (4.19)–(4.20)*. Poles are distributed according to* (4.21) *along segments of length* $\sigma L = O(n)$ *on each side of* Γ*. In the tail,* $F(t)$ *in* (4.10) *is exponentially small. In the head,* $\phi(s)/\phi(t)$ *is near-exponentially small.*

and

$$\alpha_k = s(\tilde{w}_k) \quad \tilde{w}_k = -K + i(k - \tfrac{1}{2})K'/n, \quad 1 \le k \le n. \tag{4.22}$$

THEOREM 4.6. *Let* E *be a simply-connected compact set in* $\mathbf{C}\backslash(-\infty, 0)$*; more generally,* E *might be a multilevel surface wrapping around* 0 *a finite number of times. Let* f *be an analytic function in* E *with a branch point singularity at* $z = 0$ *in the sense of Definition* 4.1 *with an exponent* τ*. Then there exist functions*

$$g(z) = c_0 + \sum_{k=1}^{n} \frac{c_k}{\log(z) - s_k} \tag{4.23}$$

satisfying

$$\|f - g\| \le \exp(-Cn/\log n) \quad \forall n \tag{4.24}$$

for some $C > 0$*, where* $\|\cdot\|$ *is the supremum norm on* E*. If* $E = [0,1]$*, a suitable choice of* $\{s_k\}$ *is* (4.21) *for any* μ *with* $\tan(\pi/2\mu) < \tau$ *and sufficiently small* σ*.*

Proof. As with Theorem 4.3, the proof is carried out via the Hermite integral formula for the equivalent problem of rational approximation of a function $F(s)$ on the domain $\log(E)$ in the $s = \log z$ variable; again we first remove $f(0)$ from the problem by setting $c_0 = f(0)$. If E is a compact set in $\mathbf{C}\backslash(-\infty, 0)$, or a multilevel surface wrapping around 0 a finite number of times, then $\log(E)$ is a closed subset of $\mathbf{C} \cup \{\infty\}$ with upper-bounded $\operatorname{Re} E$ and $|\operatorname{Im} E|$. We take $E = [0,1]$ and $\log(E) = [-\infty, 0]$ for simplicity, as sketched in Figure 4.3; the more general case is treated by an adjustment of the contour Γ and the conformal map. By the assumption (4.1), F can be extended to a single-valued analytic function throughout a V-shaped region in the s-plane, as in the figure, for any inner angle π/μ with $\tan(\pi/2\mu) < \tau$ and sufficiently small σ. As in the proof of Theorem 4.2, we now estimate the Hermite integral (4.10) over the "head" region of Γ, consisting of the two segments of length $\sigma L = \sigma\rho n$ touching the apex, and the remaining "tail." The contribution I_{tail} is exponentially small because of (4.2), which in the s variable becomes (4.6). The contribution from I_{head} is near-exponentially small, which we establish as follows. The factor $F(t)/(t - s)$ in (4.10)

is bounded as a consequence of (4.6) and the contour Γ being fixed, so we need to show that $\phi(s)/\phi(t)$ is near-exponentially small. From potential theory related to Theorem 19 of Chapter 9 of [23], it is known that this selection of $\{s_k\}$ and $\{\alpha_k\}$ as Fejér–Walsh points ensures $|\phi(s)/\phi(t)| \approx \exp(-n(K/K'))$ as $n \to \infty$. The result follows from the estimate $K'/K \sim (4\mu/\pi)\log n$ as $n \to \infty$, which can be derived from [13, equation (19.9.5)]. ∎

5. Convergence theorem: Multiple singularities. Theorems 4.2 and 4.6 are local assertions, establishing exponential and near-exponential resolution of isolated singularities. It remains to show that global approximations can be constructed by adding together these local pieces. We do this in the style of Theorem 4.6, requiring analyticity just in a neighborhood of E. The argument is adapted from the discussion around Theorem 2.3 in [5] and illustrated schematically in Figure 5.1.

THEOREM 5.1. *Let E be a simply connected compact set in $\mathbf{C}$, and let f be an analytic function in E with branch point singularities in the sense of Definition* 4.1 *at boundary points $z_1, \ldots, z_m$; more generally, E might be a multilevel surface wrapping around each branch point a finite number of times. Then there exist functions*

$$g(z) = c_0 + \sum_{j=1}^{m} \sum_{k=1}^{n_j} \frac{c_{jk}}{\log(z - z_j) - s_{jk}} \tag{5.1}$$

satisfying

$$\|f - g\| \le \exp(-Cn/\log n) \quad \forall n \tag{5.2}$$

for some $C > 0$, where $\|\cdot\|$ is the supremum norm on E.

Proof. The function f can be represented on E as a Cauchy integral

$$f(z) = \frac{1}{2\pi i} \int_\Gamma \frac{f(t)}{t - z}\, dt, \tag{5.3}$$

where Γ is any fixed contour that lies outside E and within the region of analyticity of f, except that Γ touches E at each of the points $\{z_k\}$. We split up Γ into m pieces Γ_k, with Γ_k touching just z_k, giving

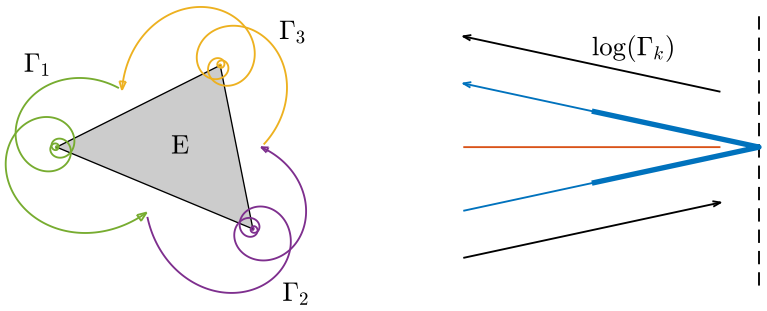

FIG. 5.1. *Illustrations for the proof of Theorem* 5.1. *A function f on E is decomposed into m pieces, with f_k defined by a Cauchy integral of f over an arc Γ_k touching the branch point z_k; Γ_k is a pair of logarithmic spirals on a Riemann surface to which f is extended by analytic continution. The right image shows the configuration near a branch point after a change of variable $s = \log(e^{i\alpha}(z - z_k))$. The arc defining f_k becomes a disjoint pair of rays $\log(\Gamma_k)$ in the s-plane, and poles $\{s_k\}$ of the rational approximation will be placed along segments of length $O(n)$ on a V-shaped contour lying between $\log(\Gamma_k)$ and $(-\infty, 0]$.*

$$f(z) = f_1(z) + \cdots + f_m(z), \tag{5.4}$$

where f_k is the Cauchy integral evaluated over the arc Γ_k,

$$f_k(z) = \frac{1}{2\pi i} \int_{\Gamma_k} \frac{f(t)}{t - z}\, dt. \tag{5.5}$$

For the mathematics of Cauchy integrals over open arcs, see Chapter 14 of [8].

We are done if we can show that each function f_k can be approximated on E with near-exponential accuracy by a function of the form (4.23), with $\log(z)$ replaced by $\log(z - z_k)$ (that is, a continuous branch of $\log(z - z_k)$ in E). This will be ensured by Theorem 4.6 if f_k satisfies the conditions (4.1) and (4.2) in our definition of a branch point singularity. At this stage, the choice of Γ becomes crucial. A choice of Γ as a closed, nonintersecting contour in the z-plane will not do; the corresponding functions f_k will not be defined along curves winding around z_k as required by (4.1). Instead, Γ must be taken to be a spiral on a Riemann surface for f whose projection on the z-plane is a self-intersecting contour with logarithmic spiral behavior at each z_k, as sketched in Figure 5.1. The integrals defining each f_k now make sense since each z_k is a branch point around which f can be analytically continued. Condition (4.2) holds since it holds for f by assumption and f_k differs from f by a function that is analytic near z_k, namely, the sum of the contributions f_j with $j \neq k$. ∎

6. Behavior near singularities. As mentioned in the penultimate paragraph of section 2 and illustrated in Figure 2.3, reciprocal-log approximations have surprising properties near the singularities. Whereas error curves for rational approximations vary over a scale exponentially close to the singularity, for reciprocal-log approximations this becomes doubly exponential. Fortunately, so far as we are aware, these effects need not cause difficulties in using these approximations, and in particular, it appears that they do not necessitate the use of extended-precision arithmetic. We shall now explain our understanding of these matters with reference to Figure 6.1, which presents four approximations of $\sqrt{z}$ for $z \in [0, 1]$. In each case the approximation is computed over $[10^{-20}, 1]$, and then the absolute value of the error is plotted over $[10^{-80}, 1]$. For comparison, fine dots show the corresponding results for computations over $[10^{-80}, 1]$. Our continued use of $\sqrt{z}$ as the basic example of a function to be approximated is justified by both numerical experience and theory showing that different branch point singularities dominated by fractional powers do not behave too differently; the starting point of this phenomenon appeared in the second sentence of this paper with the function e^{as} depending only inessentially on a.

Image (a) shows the best (minimax) reciprocal-log approximation with $n = 4$. We see here that the best approximation over $[10^{-20}, 1]$ is by no means optimal over $[10^{-80}, 1]$. The error is 10 times larger there (6.2e−4 versus 5.0e−5), and this ratio will grow rapidly with increasing n. One might think this implies that reciprocal-log approximations must be ineffective at approximating all the way up to the singularity, but the small dots in the figure, corresponding to the best approximation over $[10^{-80}, 1]$, contradict this expectation, showing an error of just 7.6e−5. As can be seen in Figure 2.3, the error in best approximation over all of $[0, 1]$ is also not much larger, just 8.7e−5.

Image (b) shows the error for reciprocal-log approximation (2.2) with poles positioned on a Hankel contour according to (2.3), now with $n = 8$. Again the error in the left-hand part of $[10^{-80}, 1]$ is larger than it needs to be, but this may be less important for these approximations since the error is dominated by that at larger values of z.

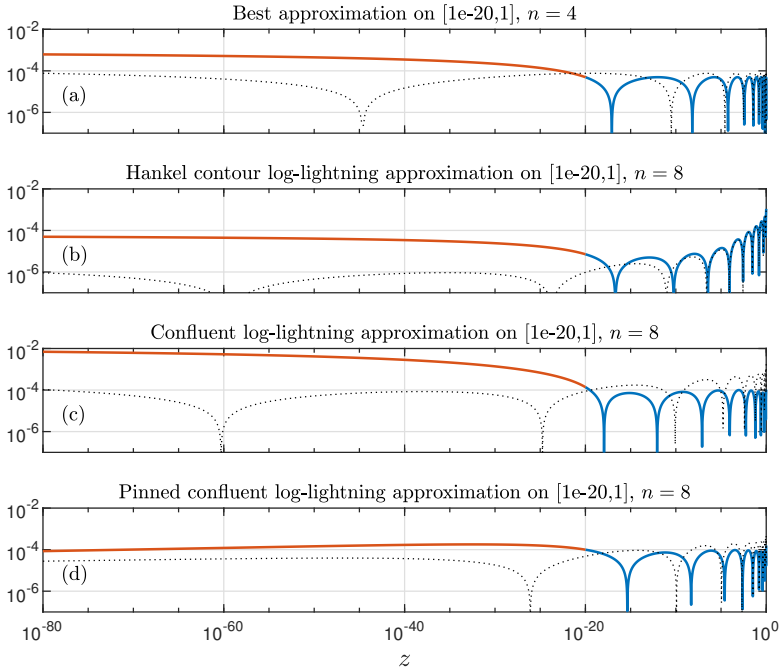

FIG. 6.1. *Four approximations of* $\sqrt{z}$ *computed on* $[10^{-20}, 1]$ *and then plotted over* $[10^{-80}, 1]$ *(the absolute value of the error). The fine dots show corresponding results for an approximation computed over* $[10^{-80}, 1]$. *See the text for discussion.*

Image (c) shows a very different situation for reciprocal-log approximation (2.4) with confluent poles at $s_0 = n/2$. Now the approximation over $[10^{-20}, 1]$ gives an error over $[10^{-80}, 1]$ that is orders of magnitude too large.

It is at this point that we wish to pause and ask, what might be the practical implications of such behavior? It certainly seems disturbing if an approximation is much less accurate on $[10^{-80}, 1]$ than on $[10^{-20}, 1]$. There are three rather disparate observations to be made, which combine to an encouraging picture.

First we note that in IEEE floating point arithmetic, whereas one can represent numbers near 0 smaller than 10^{-300}, the resolution near other points z_k is only on the order of 16 digits. Thus even if we wanted to compute approximations by least-squares fitting on a grid in $z_k + (0, 10^{-20}]$, we could not do so. This might seem worrying if one takes Figure 6.1 to suggest that such domains will be required.

The second observation is that as a purely practical matter, they should not be required. On a planar domain E such as that of Figure 3.2, for example, if a function f is approximated to a satisfactory precision everywhere except at distances $< 10^{-20}$ from the vertices $\{z_k\}$, what difference does it make? What is wrong with an approximation that is mathematically inaccurate in theory but only at points where it will never be evaluated?

And yet the use of such approximations would be troubling. This brings us to the third observation, which is the basis of image (d) of Figure 6.1: it may be possible to make an approximation accurate on $z_k + (0, 10^{-20}]$ without doing any arithmetic there. The trick is to forget least-squares gridding in $z_k + (0, 10^{-20}]$ and instead impose additional conditions at z_k, forcing the singular part of the approximation to decay there at a rate $O(1/(\log|z - z_k|)^J)$ for some power $J > 1$. This is natural

since in floating point arithmetic, all of $z_k + (0, 10^{-20}]$ rounds to z_k anyway. Image (d) corresponds to an approximation of this kind with $J = 2$: we approximate in the n-dimensional space spanned by $1/(\log(z) - n/2)^j$ with $2 \le j \le n + 2$. The work and the number of degrees of freedom are unchanged, and now least-squares fitting on $[10^{-20}, 1]$ gives an approximation that is accurate on all of $[0, 1]$. We call this a *pinned* approximation, since certain terms of the approximation are constrained to be zero at the singularity. In this example $J = 2$ is effective, but for larger n a choice closer to $J = n/2$ will be called for. To estimate this number, one can use analysis such as that of (4.12) or (4.22) to monitor how many interpolation points might be expected to fall closer to z_k than about machine precision. For the problem of image (d) with $n = 8$, (4.12) gives distances $\exp(\alpha_j) \approx 0.8, 0.6, 0.3, 0.06, 0.003, 10^{-5}, 10^{-13}, 10^{-56}$, and our pinned approximation might be thought of as effectively replacing the last two of these numbers by zero. There are many mathematical questions here, and it would be interesting to investigate them.

7. Log-lightning PDE solvers. The motivation for this work has been the development of numerical methods for the solution of PDEs on planar domains with corners, beginning with the Laplace equation with Dirichlet boundary conditions. For these problems we do not know an analytic function on E, just its real part on the boundary. This is no difficulty for linear least-squares fitting, however, where all that matters is having an efficient approximation space. The idea of solving Laplace problems in this fashion was presented in [5], where root-exponential convergence of approximations based on rational functions is established theoretically and experimentally. Here we move to reciprocal-log approximations, and the theorems of sections 4 and 5 assert that the convergence should improve to exponential or near-exponential.

Figure 7.1 presents a pair of curves confirming this expectation. The solutions computed here are for the problem posed on NA Digest in December 2018, defined by the L-shaped region with vertices $0, 2, 2 + i, 1 + i, 1 + 2i, 2i$ and boundary data $u(z) = \mathrm{Re}(z^2)$ [17]. As usual, N denotes the total number of degrees of freedom, which in this plot is a number of the form $14n+1$ with $n = 2, 4, 6, \ldots$. For each n, the approximating function consists of the real part of a polynomial of degree n together with n singularities $\{s_k\}$ at each of the six corners as defined by (2.3), though with a factor $n/3$ instead of $n/4$ in front. This gives $7n + 1$ complex degrees of freedom, hence $14n + 2$ real ones, and the number reduces to $14n + 1$ since the imaginary part of the constant term of the polynomial does not affect the fit.

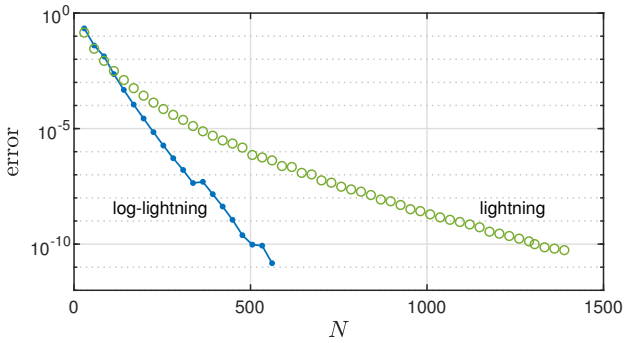

FIG. 7.1. *Convergence of lightning and log-lightning solutions to the NA Digest problem of the Laplace equation on an L-shaped region* [17]. *The difference between root-exponential and exponential convergence is evident.*

This experiment is a long way from the software that has been developed for lightning approximation [19]. The boundary has been resolved simply by 500 exponentially clustered points on each side, whereas adaptive software can make the gridding both more efficient and more careful. More importantly, each corner has been allocated the same number n of singularities, whereas adaptive software will put more singularities at vertices with stronger singularities. So although the log-lightning convergence curve of Figure 7.1 is promising, the reliable fast way to solve Laplace problems at this date is still by means of the `laplace` code at [19]. This code also addresses a number of variations such as Neumann boundary conditions, discontinuous data, and piecewise smooth curved boundaries.

The number of degrees of freedom N is not a direct measure of computer time. For reasons of linear algebra, the work actually increases in proportion to N^2 or N^3, depending on whether the grid is fixed or refined with N, which tends to make the difference between the lightning and log-lightning computations greater than it may appear in Figure 7.1. A consideration pushing in the other direction is that computer evaluation of a complex logarithm $\log z$ is typically slower than that of a reciprocal $1/z$ by a factor such as 3 or 4.

In developing software for solving PDEs by the log-lightning method, an issue deserving of attention will be the appropriate weighting of sample points in the least-squares problems. There are two reasons why uniform weighting may not be optimal. First, since the sampling grids are exponentially clustered, uniform pointwise weighting introduces a nonuniform weighting in an integrated sense, which one may wish to compensate. Second, the singularities in the solution may require nonuniform weighting; they will certainly do so when the solution is discontinuous. As a practical matter, we find that weighting seems less important than one might expect, essentially because it is covered by the great speed of the exponential or near-exponential convergence, but this matter will certainly need attention as the PDE side of this subject is developed.

It was shown in [6] that the original lightning Laplace solver can be generalized to a lightning Helmholtz solver by replacing poles by Hankel functions. Possibly there is a Helmholtz generalization of the log-lightning method too, but we have not investigated this.

8. Conclusion. Exponential or near-exponential convergence of approximations to branch point singularities is a newly discovered phenomenon. Reciprocal-log approximations are interesting in theory and may have consequences in practice if good software can be developed.

Philosophical questions are attached to this kind of approximation. It is an old idea that polynomials and rational functions have a special status since "the only operations that can really be carried out numerically are the four elementary operations of addition, subtraction, multiplication and division."[1] This point of view has contributed to the dominance of polynomials and rational functions in approximation theory, with other classes of approximating functions seeming less fundamental. We are not ourselves immune to the impression that there is something contrived about approximations of the forms (1.1) and (1.2). But their power is undeniable, and the philosophical distinction that may have seemed sharp before the arrival of computers seems less sharp today.

[1] This quote comes from the PhD thesis of Hilbert's student Kirchberger in 1902 [18, pp. 196–197].

A fascinating aspect of reciprocal-log approximations is their exploitation of the behavior of a function on a Riemann surface. As one of the smallest consequences of this feature, these approximations have no difficulty at all in treating domains with slits, and the prospects for further adventures with multivalued functions seem enticing.

Appendix: Vandermonde with Arnoldi codes. Figure 8.1 lists the codes `VAorthog` and `VAeval` from [2] that we use for numerical stability. The mathematics

```
function [Hes,R] = VAorthog(Z,n,varargin)  % Vand.+Arnoldi orthogonalization
%  Input:   Z = column vector of sample points
%           n = degree of polynomial (>= 0)
%         Pol = cell array of vectors of poles (optional)
% Output: Hes = cell array of Hessenberg matrices (length 1+length(Pol))
%           R = matrix of basis vectors
M = length(Z); Pol = []; if nargin == 3, Pol = varargin{1}; end
% First orthogonalize the polynomial part
Q = ones(M,1); H = zeros(n+1,n);
for k = 1:n
   q = Z.*Q(:,k);
   for j = 1:k, H(j,k) = Q(:,j)'*q/M; q = q - H(j,k)*Q(:,j); end
   H(k+1,k) = norm(q)/sqrt(M); Q(:,k+1) = q/H(k+1,k);
end
Hes{1} = H; R = Q;
% Next orthogonalize the pole parts, if any
while ~isempty(Pol)
   pol = Pol{1}; Pol(1) = [];
   np = length(pol); H = zeros(np,np-1); Q = ones(M,1);
   for k = 1:np
      q = Q(:,k)./(Z-pol(k));
      for j = 1:k, H(j,k) = Q(:,j)'*q/M; q = q - H(j,k)*Q(:,j); end
      H(k+1,k) = norm(q)/sqrt(M); Q(:,k+1) = q/H(k+1,k);
   end
   Hes{length(Hes)+1} = H; R = [R Q(:,2:end)];
end


function [R0,R1] = VAeval(Z,Hes,varargin)  % Vand.+Arnoldi basis construction
%  Input:   Z = column vector of sample points
%         Hes = cell array of Hessenberg matrices
%         Pol = cell array of vectors of poles, if any
% Output:  R0 = matrix of basis vectors for functions
%          R1 = matrix of basis vectors for derivatives
M = length(Z); Pol = []; if nargin == 3, Pol = varargin{1}; end
% First construct the polynomial part of the basis
H = Hes{1}; Hes(1) = []; n = size(H,2);
Q = ones(M,1); D = zeros(M,1);
for k = 1:n
   hkk = H(k+1,k);
   Q(:,k+1) = ( Z.*Q(:,k) - Q(:,1:k)*H(1:k,k)          )/hkk;
   D(:,k+1) = ( Z.*D(:,k) - D(:,1:k)*H(1:k,k) + Q(:,k) )/hkk;
end
R0 = Q; R1 = D;
% Next construct the pole parts of the basis, if any
while ~isempty(Pol)
   pol = Pol{1}; Pol(1) = [];
   H = Hes{1}; Hes(1) = []; np = length(pol); Q = ones(M,1); D = zeros(M,1);
   for k = 1:np
      Zpki = 1./(Z-pol(k)); hkk = H(k+1,k);
      Q(:,k+1) = ( Q(:,k).*Zpki - Q(:,1:k)*H(1:k,k)                  )/hkk;
      D(:,k+1) = ( D(:,k).*Zpki - D(:,1:k)*H(1:k,k) - Q(:,k).*Zpki.^2 )/hkk;
   end
   R0 = [R0 Q(:,2:end)]; R1 = [R1 D(:,2:end)];
end
```

FIG. 8.1. *MATLAB codes from* [2] *for Vandermonde with Arnoldi orthogonalization and evaluation. The polynomial part and each string of poles are orthogonalized separately, with the resulting Hessenberg matrices of coefficients stored in the cell array* `Hes`.

is described in [1]. These codes were designed for rational functions, not reciprocal logs, but as shown in the example of section 2, they can be applied to the latter case by supplying `log(z)` as one of the arguments.

Acknowledgments. The research for this article was carried out in regular communication with André Weideman, whose many suggestions we are happy to acknowledge. Another key contribution was Lemma 4.5, from Peter Baddoo. In addition we are grateful for helpful comments from Marco Fasondini, Abi Gopal, Alex Townsend, and Elias Wegert.

REFERENCES

[1] P. D. Brubeck, Y. Nakatsukasa, and L. N. Trefethen, *Vandermonde with Arnoldi,* SIAM Rev., 63 (2021), pp. 405–415.
[2] P. D. Brubeck and L. N. Trefethen, *Lightning Stokes solver,* SIAM J. Sci. Comp., submitted.
[3] W. J. Cody, G. Meinardus, and R. S. Varga, *Chebyshev rational approximations to e^{-x} in $[0, +\infty)$ and applications to heat-conduction problems,* J. Approx. Theory, 2 (1969), pp. 50–65.
[4] T. A. Driscoll and B. Fornberg, *A Padé-based algorithm for overcoming the Gibbs phenomenon,* Numer. Algorithms, 26 (2001), pp. 77–92.
[5] A. Gopal and L. N. Trefethen, *Solving Laplace problems with corner singularities via rational functions,* SIAM J. Numer. Anal., 57 (2019), pp. 2074–2094.
[6] A. Gopal and L. N. Trefethen, *New Laplace and Helmholtz solvers,* Proc. Nat. Acad. Sci. USA, 116 (2019), 10223.
[7] P. C. Hansen, *Discrete Inverse Problems: Insight and Algorithms,* SIAM, Philadelphia, 2010.
[8] P. Henrici, *Applied and Computational Complex Analysis,* vol. 3, Wiley, New York, 1986.
[9] C. L. Lawson, *Contributions to the Theory of Linear Least Maximum Approximations,* Ph.D thesis, UCLA, 1961.
[10] R. S. Lehman, *Development of the mapping function at an analytic corner,* Pacific J. Math., 7 (1957), pp. 1437–1449.
[11] E. Levin and E. Saff, *Potential theoretic tools in polynomial and rational approximation,* in Harmonic Analysis and Rational Approximation, Springer, New York, 2006, pp. 71–94.
[12] Y. Nakatsukasa, O. Sète, and L. N. Trefethen, *The AAA algorithm for rational approximation,* SIAM J. Sci. Comput., 40 (2018), pp. A1494–A1522.
[13] F. W. Olver, et al., *NIST Handbook of Mathematical Functions,* Cambridge University Press, Cambridge, UK, 2010.
[14] H. Stahl, *The convergence of Padé approximants to functions with branch points,* J. Approx. Theory, 91 (1997), pp. 139–204.
[15] H. Stahl and T. Schmelzer, *An extension of the '1/9' problem,* J. Comp. Appl. Math., 233 (2009), pp. 821–834.
[16] G. Starke, *Fejér–Walsh points for rational functions and their use in the ADI iterative method,* J. Comp. Appl. Math., 46 (1993), pp. 129–141.
[17] L. N. Trefethen, 8-*digit Laplace solutions on polygons?,* NA Digest, December 3, 2018, http://www.netlib.org/na-digest-html.
[18] L. N. Trefethen, *Approximation Theory and Approximation Practice, Extended Edition,* SIAM, Philadelphia, 2019.
[19] L. N. Trefethen, *Lightning Laplace software*, 2020, https://people.maths.ox.ac.uk/trefethen/lightning.
[20] L. N. Trefethen, Y. Nakatsukasa, and J. A. C. Weideman, *Exponential node clustering at singularities for rational approximation, quadrature, and PDEs,* Numer. Math., 147 (2021), pp. 227–254.
[21] L. N. Trefethen and J. A. C. Weideman, *The exponentially convergent trapezoidal rule,* SIAM Rev., 56 (2014), pp. 385–458.
[22] L. N. Trefethen, J. A. C. Weideman, and T. Schmelzer, *Talbot quadratures and rational approximations,* BIT Numer. Math., 46 (2006), pp. 653–670.
[23] J. L. Walsh, *Interpolation and Approximation by Rational Functions in the Complex Domain,* Amer. Math. Soc., Providence, RI, 1935.
[24] W. Wasow, *Asymptotic development of the solution of Dirichlet's problem at analytic corners,* Duke Math. J., 24 (1957), pp. 47–56.

[25] E. Wegert, *Visual Complex Functions: An Introduction with Phase Portraits,* Birkhäuser, Basel, 2012.
[26] J. A. C. Weideman and L. N. Trefethen, *Parabolic and hyperbolic contours for computing the Bromwich integral,* Math. Comput., 76 (2007), pp. 1341–1356.