MANCHESTER
1824

Near-optimal perfectly matched layers for indefinite Helmholtz problems

Druskin, Vladimir and Güttel, Stefan and Knizhnerman, Leonid

2013

Manchester Institute for Mathematical Sciences

School of Mathematics

The University of Manchester

NEAR-OPTIMAL PERFECTLY MATCHED LAYERS
FOR INDEFINITE HELMHOLTZ PROBLEMS

VLADIMIR DRUSKIN[∗], STEFAN GÜTTEL[†], AND LEONID KNIZHNERMAN[‡]

Abstract. A new construction of an absorbing boundary condition for indefinite Helmholtz problems on unbounded domains is presented. This construction is based on a near-best uniform rational interpolant of the inverse square root function on the union of a negative and positive real interval, designed with the help of a classical result by Zolotarev. Using Krein's interpretation of a Stieltjes continued fraction, this interpolant can be converted into a three-term finite difference discretization of a perfectly matched layer (PML) which converges exponentially fast in the number of grid points. The convergence rate is asymptotically optimal for both propagative and evanescent wave modes. Several numerical experiments and illustrations are included.

Key words. Helmholtz equation, Neumann-to-Dirichlet map, perfectly matched layer, rational approximation, Zolotarev problem, continued fraction

AMS subject classifications. 35J05, 65N06, 65N55, 30E10, 65D25

1. Introduction. An important task in science and engineering is the numerical solution of a partial differential equation (PDE) on an unbounded domain. To give a motivating example, let us consider the Helmholtz equation $\Delta u + \frac{k^2}{v^2} u = 0$ on an unbounded half-space Ω which may correspond to the acoustic model of wave propagation in the interior of the earth with variable wavespeed v, see Figure 1.1. In this setup, which is typical in seismic geophysical exploration methods, a pressure wave signal of a single frequency $k > 0$ is emitted by an acoustic transmitter placed on the subsurface of the earth, travels through the underground, and is then logged by receivers. Clearly, the computational domain for this problem needs to be truncated as sharply as possible and hopefully this is done so that the artificial boundaries create least possible reflections. There are various ways for achieving this, with probably the most popular approach being known as *perfectly matched layer* (PML, see [7, 9, 11]).

A perfectly matched layer can be seen as a localized modification of the spatial discretization scheme to absorb the incoming waves. In a finite difference framework such layers typically lead to variable complex-valued step sizes, which is why this approach is sometimes also referred to as *complex coordinate stretching*. The aim of an efficient PML is to achieve a strong absorption effect by adding only a few number of layers. The aim of this work is to extend a modern finite-difference construction of perfectly matched layers which are near-optimal for indefinite Helmholtz problems, that is, they achieve near-best possible absorption for a given number of layers. The number of required layers is critical in particular for large-scale simulations of three-dimensional exterior problems. A variety of such problems arise, for example, in oil and gas exploration, and near-optimal grids are part of almost all electromagnetic simulators used at Schlumberger [1, 13, 50]. However, the potential beneficiaries of effective discretization of exterior domains are by no means limited to these problems: other examples include ocean acoustics, the simulation of micro-electro-mechanical

---
[∗]Schlumberger-Doll Research, 1 Hampshire St., Cambridge, Massachusetts, 19104-2688 (druskin1@boston.oilfield.slb.com).

[†]School of Mathematics, The University of Manchester, Alan Turing Building, Manchester, M13 9PL, United Kingdom (stefan.guettel@manchester.ac.uk).

[‡]Mathematical Modelling Department, Central Geophysical Expedition, Narodnogo Opolcheniya St., house 38, building 3, 123298, Moscow, Russia (lknizhnerman@gmail.com).
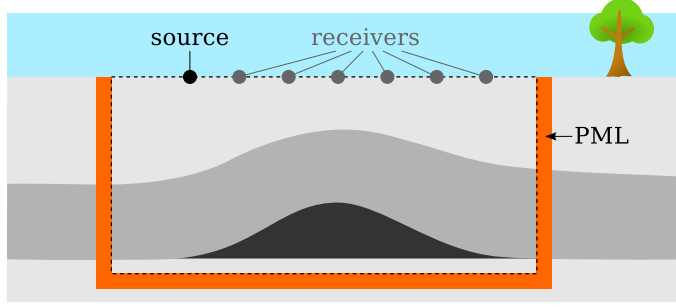
FIGURE 1.1. *A typical setup in seismic geophysical exploration where a source emits pressure waves into the earth subsurface which is then logged at (multiple) receivers. The wave propagation is modeled by the Helmholtz's equation.*

systems (MEMS), the propagation of radio waves in the atmosphere, surface and ground penetrating radars, photonic crystals, and molecular physics.

1.1. Outline of this work. We will now give a short overview of this work and explain the structure of the paper. Let us start by considering a prototype of a differential equation on an unbounded domain, the two-point boundary value problem

$$\frac{\partial^2}{\partial x^2} \boldsymbol{u} = \boldsymbol{A}\boldsymbol{u}, \quad \frac{\partial}{\partial x} \boldsymbol{u}\big|_{x=0} = -\boldsymbol{b}, \quad \boldsymbol{u}\big|_{x=+\infty} = \boldsymbol{0}, \tag{1.1}$$

where $\boldsymbol{A} \in \mathbb{C}^{N \times N}$ is nonsingular and $\{\boldsymbol{b}, \boldsymbol{u}(x)\} \subset \mathbb{C}^N$. If $\boldsymbol{A}$ is a discretization of a differential operator on some spatial domain $\Omega \subseteq \mathbb{R}^\ell$, then (1.1) is a semidiscretization of an $(\ell+1)$-dimensional partial differential equation on $[0, +\infty) \times \Omega$. Assuming that problem (1.1) is well posed (which may require some additional conditions like, e.g., the limiting absorption principle discussed below), its exact solution can be given in terms of matrix functions as $\boldsymbol{u}(x) = \exp(-x\boldsymbol{A}^{1/2})\boldsymbol{A}^{-1/2}\boldsymbol{b}$. In particular, at $x = 0$ the solution is given as

$$\boldsymbol{u}(0) = F(\boldsymbol{A})\boldsymbol{b}, \quad F(z) = z^{-1/2}. \tag{1.2}$$

The function $F(z)$ is often referred to as the *impedance function* (also known as *Weyl function*), and it completely characterizes the reaction of the unbounded domain to an external force [34]. The relation (1.2) allows for the exact conversion of the Neumann data $-\boldsymbol{b}$ at the boundary $x = 0$ into the Dirichlet data $\boldsymbol{u}(0)$, without the need for solving (1.1) on its unbounded domain. This is why $F(\boldsymbol{A})$ is often referred to as the *Neumann-to-Dirichlet (NtD) operator*.

When solving wave scattering problems one typically deals with a discretization of the negative shifted Laplacian $-\Delta - k^2$ on $\Omega \subset \mathbb{R}^\ell$, in which case problem (1.1) is a semidiscretization of the indefinite Helmholtz equation on $[0, +\infty) \times \Omega$. It is thus reasonable to consider a Hermitian indefinite matrix

$$\boldsymbol{A} = \boldsymbol{L} - k^2 \boldsymbol{I}, \tag{1.3}$$

where $\boldsymbol{L} \in \mathbb{C}^{N \times N}$ is Hermitian positive definite, $\boldsymbol{I} \in \mathbb{R}^{N \times N}$ is the identity matrix, and $k^2 > 0$ is not in the spectrum of $\boldsymbol{L}$. For a solution of (1.1) to be unique we impose the limiting absorption principle (see, e.g., [45]). This means that for a real number

k we define $\boldsymbol{u}$ as a limit of solutions $\boldsymbol{u}^{(k+i\epsilon)}$ of (1.1) with wave numbers $k+i\epsilon$ ($\epsilon > 0$) instead of k, i.e.,

$$\boldsymbol{u} = \lim_{\epsilon \searrow 0} \boldsymbol{u}^{(k+i\epsilon)}. \tag{1.4}$$

This uniquely defines the value $F(\boldsymbol{A}) = \boldsymbol{A}^{-1/2}$, notwithstanding that some eigenvalues of $\boldsymbol{A}$ may lie on the standard branch cut of $F(z)$.

We will now outline our construction in the following sections, which combines ideas of the eminent mathematicians Y. I. Zolotarev (1847–1878), T. J. Stieltjes (1856–1894), and M. G. Krein (1907–1989). The main aim in section 2 is to approximate $F(z)$ by a rational interpolant $R_n(z)$ of type $(n-1, n)$, so that $R_n(\boldsymbol{A})$ can be seen as an approximate NtD operator, mapping the Neumann data $-\boldsymbol{b}$ to the Dirichlet data $R_n(\boldsymbol{A})\boldsymbol{b}$. Clearly, the 2-norm approximation error of this map is

$$\|R_n(\boldsymbol{A})\boldsymbol{b} - F(\boldsymbol{A})\boldsymbol{b}\| = \sqrt{\sum_{j=1}^{N} |R_n(\lambda_j - k^2) - F(\lambda_j - k^2)|^2 |b_j|^2},$$

where $b_j = \boldsymbol{v}_j^* \boldsymbol{b}$ and $(\lambda_j, \boldsymbol{v}_j)$ are the eigenpairs of $\boldsymbol{L}$ with $\|\boldsymbol{v}_j\| = 1$. We have $\lambda_1 < k^2 < \lambda_N$ and thus arrive at the problem of scalar rational approximation of $F(z)$ on the union of a positive and a negative real interval. Our rational interpolant $R_n(z)$ is obtained by combining two optimal Zolotarev interpolants constructed for the two intervals separately. For illustration purposes we have graphed the relative error of such a function in Figure 1.2. In addition to the construction of such approximants, section 2 also contains a novel detailed convergence analysis, with the more technical proofs given in the appendix.

In section 3 we will show that the rational function $R_n(z)$ can be converted into an equivalent three-term finite difference scheme on a nonuniform grid with n points. This is achieved by formally rewriting $R_n(z)$ as a Stieltjes continued fraction and using Krein's interpretation of that fraction as a finite-difference scheme. However, due to the non-Stieltjes nature of $R_n(z)$ (its poles may lie on a curve in complex plane, as shown in Figure 2.1) the continued fraction coefficients can also be complex, which results in a finite difference scheme with complex-valued grid steps. This scheme allows for the simple and efficient computation of an NtD map and the construction of an absorbing boundary layer for indefinite Helmholtz problems. The near-optimality of $R_n(z)$ implies that the number of required grid points is close to smallest possible. A summary of an algorithm for computing this grid is given in section 4.

Section 5 is devoted to the adaptation of our PML construction to a second-order finite difference framework. In section 5.1, we extend our optimal rational approximation approach to the infinite lattice problem.

Finally, in section 6 we demonstrate the high accuracy and exponential convergence of our perfectly matched layer with several numerical examples.

1.2. Review of related work. It was already shown in [16, 32] that a rational approximant $R_n(z)$ of type $(n-1, n)$ for the function $F(z)$ can be converted into an equivalent three-term finite difference scheme on a special nonuniform grid with n points, mapping the Neumann data $-\boldsymbol{b}$ to the Dirichlet data $R_n(\boldsymbol{A})\boldsymbol{b}$. In these papers the authors were mainly concerned with a special instance of (1.1) where $\boldsymbol{A}$ corresponds to a discretization of the negative Laplacian $-\Delta$, in which case $\boldsymbol{A}$ is a real symmetric positive definite matrix. The error of the approximate Neumann-to-Dirichlet (NtD) map is then bounded by the maximum of $|R_n(z) - F(z)|$ on the
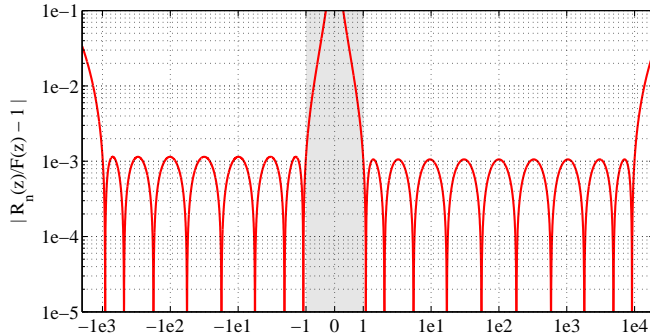
FIGURE 1.2. *Relative error $|R_n(z)/F(z)-1|$ of a rational approximant $R_n(z)$ for $F(z) = z^{-1/2}$ on $[-1e3, -1] \cup [1, 1e4]$. We have adopted a special plotting type for simultaneously visualizing large intervals on the negative and positive real semiaxes in logarithmic scales, with the gray linear region in the middle gluing the two intervals together. The rational function $R_n(z)$ is of type $(n-1, n)$, $n = 9$, and it has been constructed by combining two Zolotarev interpolants with $m_1 = 8$ and $m_2 = 10$ interpolation nodes for the negative and positive intervals, respectively. Visually the solution of our complex rational approximation problem behaves similarly to the max-norm optimal errors of the real problems, i.e., it shows "equal ripples" on the targeted intervals (although the Chebyshev alternation theory [2, Ch. II] is not applicable in the complex case [47]).*

positive spectral interval of $\boldsymbol{A}$. Approximation theory allows for the construction of exponentially convergent rational functions $R_n(z)$ with a convergence rate weakly dependent on the condition number of $\boldsymbol{A}$, thus producing a three-term finite difference scheme with a so-called *optimal grid* (also known as *finite-difference Gaussian rule* or *spectrally matched grid*). The connection of $R_n(z)$ and this grid is inspired by Krein's mechanical interpretation of a Stieltjes continued fraction [34]. It was shown in [5] that the same grids produce exponentially convergent NtD maps even for problems arising from the semidiscretization of *anisotropic* elliptic PDEs and systems with mixed second-order terms, i.e., when the second-order ODE system in (1.1) is modified by adding a first-order term.

It should be noted that the positive and negative eigenmodes of $\boldsymbol{A}$ correspond to so-called *evanescent* and *propagative* solutions $\eta(\boldsymbol{A})\boldsymbol{u}(x)$ and $\eta(-\boldsymbol{A})\boldsymbol{u}(x)$, respectively, with $\eta(s)$ denoting the Heaviside step function. The evanescent modes, i.e., the nonzero eigenmodes in the spectral decomposition of $\eta(\boldsymbol{A})\boldsymbol{u}(x)$, decay exponentially as x increases (hence the name). Therefore a simple, though possibly not the most efficient, way to absorb them is to truncate the domain at some (sometimes quite significant) distance from the targeted area of interest, and then to deal with the propagative modes alone. On the other hand, the norm $\|\eta(-\boldsymbol{A})\boldsymbol{u}(x)\|$ does actually not depend on x, so simple boundary truncation will not be effective for absorbing propagative modes.

In their seminal paper [19], Enquist and Majda computed $R_n(z)$ as a Padé approximant of $F(z)$ at some real negative point and then evaluated it via continued fraction-type recursions. This approach yielded exponential convergence on the negative real semiaxis, however, with the rate quickly deteriorating towards the origin. Another celebrated approach for absorbing propagative modes is called *complex scaling* and was originally introduced in [7] for molecular physics calculations. It is also known as *perfectly matched layer* (PML), a term coined in the influential work [9], where it was independently rediscovered and adapted for time-domain wave propa-

gation. We will use the latter term because it seems to be more established in the wave propagation literature. The well-posedness of the PML formulation was studied in [3, 8]. The essence of the PML approach is a complex coordinate transformation which changes purely imaginary exponentials of propagative modes to complex decaying ones, thus, in principle, allowing reflectionless domain truncation [7, 11]. However, coarse PML discretizations introduce undesirable numerical reflections which decay rather slowly with the grid size in case of low-order discretization schemes. This problem was partially circumvented in [4] for the solution of time-domain wave problems, where the optimal gridding approach was extended to PML discretizations. By choosing an appropriate *purely imaginary* grid this approach allowed for the construction of all possible rational interpolants $R_n(z)$ for $F(z)$ on a real negative interval, including the Padé approximants constructed in [19], and preferably the best uniform approximants targeting the spectral support of the expected solution. See also [37] and [17] for adaptations of the optimal gridding approach to the hyperbolic elasticity system and the Helmholtz equation, respectively. A non-optimal PML layer for absorbing both evanescent and propagative modes in dispersive wave equations has been proposed in [49]. However, the problem of designing discrete PMLs which are optimal for both wave modes remained open.

Our construction in section 2 is inspired by a "trick" originally used by Zolotarev and Newmann, writing the relative approximation error $R_n(z)/F(z) - 1$ in terms of $H_m(s)/H_m(-s)$, where $H_m(s)$ is a polynomial of degree $m = 2n$, $s^2 = z$. This trick was rediscovered in [26, 27], where $H_m(s)/H_m(-s)$ was identified with the numerical reflection coefficient, and a continued-fraction absorbing condition was explicitly constructed in terms of the roots of $H_m(s)$ and introduced in the PDE discretization via a so-called trapezoid finite element method. However, these important papers fell short of introducing optimal approximants. In addition to the construction of these approximants, section 2 also contains a novel detailed convergence analysis. To make our paper more pleasant to read we have decided to present the technical proofs in an appendix.

In an unfinished report [15], the authors suggested to split $H_m(s)$ into the product of polynomials with real and imaginary roots, thus decoupling the approximation problems on the positive and negative intervals. It was then suggested to apply conventional optimal rational approximants on each of the two intervals, and the resulting error was only determined by the largest error of these two approximants. A drawback of such an approach is that it requires the splitting of the PML grid into two subdomains with nonlocal finite difference stencils at the conjugation interfaces.

2. Construction of a near-optimal approximant on two intervals. The function $z^{-1/2}$ is commonly defined in the complex plane $\mathbb{C}$ with the slit $(-\infty, 0] \subset \mathbb{R}$. However, in our application we need an analytic continuation $F(z)$ of $z^{-1/2}$ from the positive real semiaxis $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x > 0\}$ to $-\mathbb{R}_+$ in accordance with the limiting absorption principle (1.4), i.e., attaining the values

$$F(z) = -i(-z)^{-1/2} \qquad \text{for} \quad z \in -\mathbb{R}_+,$$

and the principal value of the square root for $z \in \mathbb{R}_+$. We will therefore assume in the following that $F(z)$ is defined in $\mathbb{C}$ with the branch cut in the lower half-plane.

Following [15], we now construct a rational interpolant $R_n(z)$ of type $(n-1, n)$

to $F(z)$ on the union K of two real intervals

$$K = K_1 \cup K_2, \quad K_1 = [a_1, b_1], \quad K_2 = [a_2, b_2],$$
$$a_1 < b_1 < 0 < a_2 < b_2,$$

using solutions of a classical Zolotarev problem on each of the two intervals. In view of the definition (1.3), these intervals will correspond to the spectral subintervals $[\lambda_1 - k^2, \lambda_{i_0} - k^2]$ and $[\lambda_{i_0+1} - k^2, \lambda_N - k^2]$ (or their estimates), respectively, where $\lambda_1 \leq \cdots \leq \lambda_{i_0} < k^2 < \lambda_{i_0+1} \leq \cdots \leq \lambda_N$.

Separating the odd and even parts of a polynomial H_m of degree $m = 2n$, we define polynomials P_{n-1} and Q_n of degrees $\leq n-1$ and n, respectively, such that

$$H_m(s) = -s\, P_{n-1}(s^2) + Q_n(s^2). \tag{2.1}$$

The rational function

$$R_n(z) = \frac{P_{n-1}(z)}{Q_n(z)} \tag{2.2}$$

will be considered as an approximant for $F(z)$ on K. We have

$$sR_n(s^2) = \frac{s\, P_{n-1}(s^2)}{Q_n(s^2)} = \frac{H_m(-s) - H_m(s)}{H_m(-s) + H_m(s)}, \tag{2.3}$$

and thereby obtain an expression of the relative averaged approximation error as

$$2\frac{|F(s^2) - R_n(s^2)|}{|F(s^2) + R_n(s^2)|} = 2\left|\frac{H_m(s)}{H_m(-s)}\right|.$$

Following [15, Section 2], we can split the approximation problem on K into two independent problems on K_1 and K_2.

LEMMA 2.1. *Let m_1 and m_2 be positive integers such that $m = m_1 + m_2$, and let H_{m_1} and H_{m_2} be polynomials of degrees m_1 and m_2 with roots on $F(K_1)$ and $F(K_2)$, respectively. Define*

$$H_m(s) = H_{m_1}(s)H_{m_2}(s).$$

Then

$$\max_{s\in F(K_1)}\left|\frac{H_m(s)}{H_m(-s)}\right| = \max_{s\in F(K_1)}\left|\frac{H_{m_1}(s)}{H_{m_1}(-s)}\right|$$

and

$$\max_{s\in F(K_2)}\left|\frac{H_m(s)}{H_m(-s)}\right| = \max_{s\in F(K_2)}\left|\frac{H_{m_2}(s)}{H_{m_2}(-s)}\right|.$$

Proof. This lemma immediately follows from the equalities

$$\left|\frac{H_{m_1}(s)}{H_{m_1}(-s)}\right| = 1 \qquad \text{if} \quad s \in F(K_2),$$

and reciprocally

$$\left|\frac{H_{m_2}(s)}{H_{m_2}(-s)}\right| = 1 \qquad \text{if} \quad s \in F(K_1).$$

☐

Let us consider a single real interval $[c, d]$ with $0 < c < d$, and the problem of finding a real monic polynomial $Z_m^{(c,d)}$ of degree $m \geq 1$ (denoted as $Z_m^{(c,d)} \in \mathcal{P}_{m,\mathrm{real}}$) which attains the minimum in the Zolotarev problem

$$E_m^{(c,d)} = \min_{Z \in \mathcal{P}_{m,\mathrm{real}}} \max_{c \leq s \leq d} \left| \frac{Z(s)}{Z(-s)} \right|. \tag{2.4}$$

It is known from [38, 51] that this minimizer $Z_m^{(c,d)}$ exists uniquely, that its roots $s_j^{(c,d)}$ $(j = 1, \ldots, m)$ are located in (c, d), and that they are expressible in terms of elliptic integrals. More details are given in the appendix, in particular, formula (A.1).

We choose positive integers m_1 and m_2 and introduce the polynomial

$$H_m(s) = Z_{m_1}^{(\sqrt{-b_1}, \sqrt{-a_1})}(-is) \cdot Z_{m_2}^{(\sqrt{a_2}, \sqrt{b_2})}(s) \tag{2.5}$$

of degree $m = m_1 + m_2$. From Lemma 2.1 we obtain the following result.

PROPOSITION 2.2. *The polynomial $H_m(s)$ defined in (2.5) satisfies*

$$\max_{s \in F(K)} \left| \frac{H_m(s)}{H_m(-s)} \right| = \max \left\{ E_{m_1}^{(\sqrt{-b_1}, \sqrt{-a_1})}, \ E_{m_2}^{(\sqrt{a_2}, \sqrt{b_2})} \right\}.$$

It is well known that the classical Zolotarev functions in (2.4) converge exponentially. Let us denote by $\rho^{(\delta)}$ the Cauchy–Hadamard convergence rate of $Z_m^{(c,d)}$, i.e.,

$$\rho^{(\delta)} = \lim_{m \to \infty} \sqrt[m]{E_m^{(c,d)}}, \quad \delta = c/d.$$

An exact expression of $\rho^{(\delta)}$ in terms of elliptic integrals is given in (A.3). For small *interval ratios δ* one can derive a simple approximate expression

$$\rho^{(\delta)} \approx \exp\left(-\frac{\pi^2}{4 \log \frac{2}{\sqrt{\delta}}} \right)$$

in terms of elementary functions [32, Appendix A]. This expression shows the weak dependence of the Cauchy–Hadamard convergence rate on the interval ratio δ.

In view of Proposition 2.2, m_1 and m_2 should be chosen to balance the errors of both Zolotarev functions. One way of achieving this is by setting

$$\rho_1 = \rho^{(\sqrt{b_1/a_1})}, \qquad \rho_2 = \rho^{(\sqrt{a_2/b_2})} \tag{2.6}$$

and

$$m_1 = m \cdot \frac{\log \rho_2}{\log \rho_1 + \log \rho_2} + \theta, \quad m_2 = m - m_1, \quad |\theta| \leq 1/2, \tag{2.7}$$

where θ is chosen to round $(m \log \rho_2)/(\log \rho_1 + \log \rho_2)$ to the nearest integer. We are now in the position to formulate a near-optimality result for the obtained approximant.

THEOREM 2.3. *Let us denote*

$$\rho = \exp\left(\frac{\log \rho_1 \log \rho_2}{\log \rho_1 + \log \rho_2} \right). \tag{2.8}$$

Let the polynomial H_m be defined by (2.5), the polynomials P_{n-1} and Q_n defined by (2.1), the rational fraction R_n defined by (2.2), and $m = 2n$. Further let the conditions (2.7) and

$$2 \max \left\{ \rho_1^{-1/2}, \rho_2^{-1/2} \right\} \rho^m < 1 \tag{2.9}$$

be satisfied. Then the upper relative error bound

$$\max_{z \in K} \left| \frac{R_n(z)}{F(z)} - 1 \right| \leq \frac{4 \max \left\{ \rho_1^{-1/2}, \rho_2^{-1/2} \right\} \rho^m}{1 - 2 \max \left\{ \rho_1^{-1/2}, \rho_2^{-1/2} \right\} \rho^m} \tag{2.10}$$

holds. On the other hand, if P and $Q \not\equiv 0$ are arbitrary polynomials of degrees $\leq n-1$ and $\leq n$, respectively, then $R = P/Q$ satisfies the lower error bound

$$\max_{z \in K} \left| \frac{R(z)}{F(z)} - 1 \right| \geq \frac{2\rho^m}{1 + \rho^m}. \tag{2.11}$$

This theorem, whose proof is given in the appendix, implies that the upper error bound for our *Zolotarev approximant* $R_n(z)$ and the lower bound for the best possible approximant have the same Cauchy–Hadamard convergence rate ρ, i.e., our approximant is asymptotically optimal in the Cauchy–Hadamard sense. As is also demonstrated by the following numerical example (and the corresponding Table 2.1), the Zolotarev approximant can be worse than the best possible approximant only by a moderate factor. We should point out that, unlike their real counterparts, complex max-norm optimal rational approximation problems are generally not convex and may have non-unique solutions [47]. It therefore seems unlikely that the near-optimality result of Theorem 2.3 can be improved significantly.

EXAMPLE 2.1. *Let us, as in Figure 1.2, consider the problem of approximating $F(z) = z^{-1/2}$ by a rational function $R_n(z)$ of type $(n-1, n)$ on the union of two intervals $K = [a_1, b_1] \cup [a_2, b_2] = [-1\mathrm{e}3, -1] \cup [1, 1\mathrm{e}4]$. Using the exact formula (A.3) we calculate*

$$\rho_1 \approx 0.361, \quad \rho_2 \approx 0.439, \quad \rho \approx 0.634.$$

In Table 2.1 we list the error bounds of Theorem 2.3 for various values of $m = 2n$ together with the actual approximation error. The calculations confirm the bounds and show that they are roughly of the same order, i.e., our approximants $R_n(z)$ have relative errors of the same order as the best possible approximants.

The logarithmic surface plot in Figure 2.1 shows the relative error $|R_n(z)/F(z)-1|$ for the case $n = 9$ (the same as in Figure 1.2). Note how the poles align on a curve in the lower-left quadrant of the complex plane. We speculate that this curve asymptotically (as $n \to \infty$) approximates the shifted branch cut C of the analytic continuation of $F(z)$ into the lower half-plane, and that C possesses the so-called S-property ("symmetry property", see [24, 25, 43]) with respect to K. This would imply that the equilibrium charge of the condenser (K, C) has a logarithmic potential which is (constant and) minimal on K over all "attainable" branch cuts. Our experiments also suggested that the curve C coincides exactly with the negative imaginary semiaxis in the case of symmetric intervals $K_1 = -K_2$, and that it approaches the real positive or negative semiaxis for large or small ratios m_1/m_2, respectively.

A remarkable feature in Figure 2.1 is that the relative error $|R_n(z)/F(z) - 1|$ stays uniformly small "above" the set K, i.e., for complex numbers z with positive imaginary part and real part in K. We will return to this observation in section 6.2.

Lower and upper error bounds of Theorem 2.3 and actual errors $\max_{z\in K}|R_n(z)/F(z)-1|$, $F(z)=z^{-1/2}$, for various values of $m=2n$. The set K is chosen as $K=[-1e3,-1]\cup[1,1e4]$.

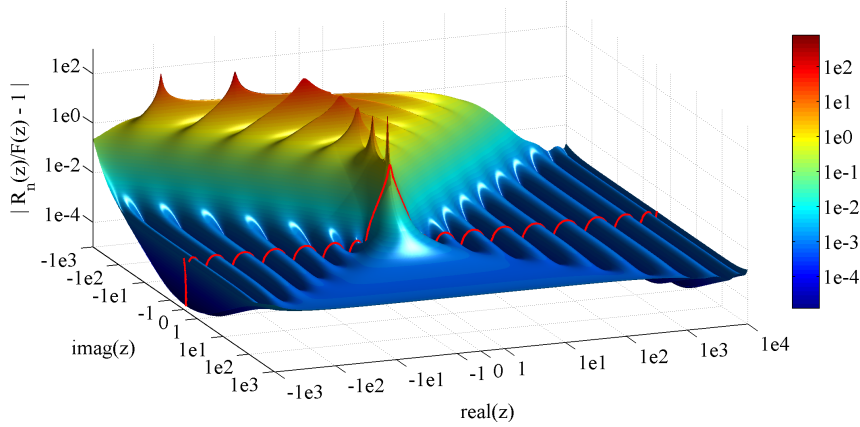| m	m_1	m_2	bound (2.11)	relative error	bound (2.10)
6	3	3	$1.22e-01$	$3.42e-01$	$5.52e-01$
12	5	7	$8.41e-03$	$2.47e-02$	$2.85e-02$
18	8	10	$5.49e-04$	$1.15e-03$	$1.83e-03$
24	11	13	$3.57e-05$	$8.95e-05$	$1.19e-04$
30	13	17	$2.32e-06$	$7.01e-06$	$7.72e-06$
36	16	20	$1.51e-07$	$3.29e-07$	$5.02e-07$
42	19	23	$9.79e-09$	$2.37e-08$	$3.26e-08$
48	21	27	$6.36e-10$	$2.01e-09$	$2.12e-09$
54	24	30	$4.13e-11$	$9.43e-11$	$1.38e-10$
60	27	33	$2.69e-12$	$6.28e-12$	$8.94e-12$

FIGURE 2.1. *Relative error $|R_n(z)/F(z)-1|$ of a Zolotarev approximant $R_n(z)$ for $K=[-1e3,-1]\cup[1,1e4]$ and $n=9$ shown as a logarithmic surface plot over a region in the complex plane. The imaginary axis is plotted in reversed direction for a better panoramic view.*

3. Finite difference grids from rational approximants. We now explain how a rational function $R_n(z)\approx F(z)$ can be transformed into an equivalent finite difference grid for (1.1). Assume that we are given primal grid points and steps

$$0=x_0,\ x_1,\ \ldots,\ x_n,\quad h_j=x_j-x_{j-1},$$

and dual grid points and steps

$$0=\widehat{x}_0,\ \widehat{x}_1,\ \ldots,\ \widehat{x}_n,\quad \widehat{h}_{j-1}=\widehat{x}_j-\widehat{x}_{j-1},$$

with $j=1,\ldots,n$ in both cases. Denote by $\boldsymbol{u}_0,\boldsymbol{u}_1,\ldots,\boldsymbol{u}_n$ approximations to the solution $\boldsymbol{u}(x)$ of (1.1) at the primal grid points $x_0,x_1,\ldots,x_n$. Let the first-order finite differences $(\boldsymbol{u}_j-\boldsymbol{u}_{j-1})/h_j$ be located at the dual points $\widehat{x}_j$ ($j=1,\ldots,n$). We

assume that the following finite difference relations

$$\frac{1}{\widehat{h}_0}\left(\frac{\boldsymbol{u}_1 - \boldsymbol{u}_0}{h_1} + \boldsymbol{b}\right) = \boldsymbol{A}\boldsymbol{u}_0, \tag{3.1}$$

$$\frac{1}{\widehat{h}_j}\left(\frac{\boldsymbol{u}_{j+1} - \boldsymbol{u}_j}{h_{j+1}} - \frac{\boldsymbol{u}_j - \boldsymbol{u}_{j-1}}{h_j}\right) = \boldsymbol{A}\boldsymbol{u}_j, \quad j = 1, \ldots, n-1, \tag{3.2}$$

are satisfied with the convention that $\boldsymbol{u}_n = \boldsymbol{0}$. It can be verified by back-substitution that the value $\boldsymbol{u}_0$ specified by these recursive relations can be written as

$$\boldsymbol{u}_0 = R_n(\boldsymbol{A})\boldsymbol{b},$$

where $R_n(z)$ is a rational function of type $(n-1, n)$. By construction, $-R_n(\boldsymbol{A})^{-1}$ is the Schur complement of the submatrix with positive indices of the system (3.1)–(3.2). Written as a finite-length Stieltjes continued fraction (S-fraction[1]) this function takes the form

$$R_n(z) = \cfrac{1}{\widehat{h}_0 z + \cfrac{1}{h_1 + \cfrac{1}{\widehat{h}_1 z + \cdots + \cfrac{1}{h_{n-1} + \cfrac{1}{\widehat{h}_{n-1} z + \cfrac{1}{h_n}}}}}}. \tag{3.3}$$

Recalling from above that the exact solution of (1.1) satisfies $\boldsymbol{u}(0) = \boldsymbol{A}^{-1/2}\boldsymbol{b}$, we are apparently left with the problem of determining $R_n(z)$ such that $R_n(\boldsymbol{A})\boldsymbol{b} \approx \boldsymbol{A}^{-1/2}\boldsymbol{b}$, optimally in some sense. The conversion of Neumann data $-\boldsymbol{b}$ to Dirichlet data $\boldsymbol{u}(0)$ can now be realized by solving a finite difference relation on a grid generated from quantities $\widehat{h}_{j-1}$ and h_j $(j = 1, \ldots, n)$ in (3.3).

The connection between the S-fraction (3.3) and the finite difference problem (3.1)–(3.2) is due to Mark Krein (see, e.g., [34]). He viewed the problem (3.1)–(3.2) as a so-called *Stieltjes string*, which is a string of point masses $\widehat{h}_{j-1}$ and weightless stiffnesses h_j $(j = 1, \ldots, n)$, both real positive. There is a one-to-one correspondence between the set of Stieltjes strings and Stieltjes spectral functions $R_n(z)$, which are rational functions of type $(n-1, n)$ having n non-coinciding real negative poles and real positive residues. For this case, the S-fraction parameters $\widehat{h}_{j-1}$ and h_j $(j = 1, \ldots, n)$ can be computed via $2n$ steps of the Euclidean polynomial division algorithm (see, e.g., [31]), which can be stably executed with the help of the reorthogonalized Lanczos algorithm [16]. The optimal rational approximation of $F(z)$ on a positive real interval is a Stieltjes problem [32], hence the generated grid steps are real positive. The approximation problem on a single negative interval can be solved by using $R_n(-z)$, where $R_n(z)$ is the approximation on the symmetrically reflected positive interval. This reflection rotates the grid steps $\widehat{h}_{j-1}$ and h_j $(j = 1, \ldots, n)$ by an angle of $\pi/2$ in $\mathbb{C}$, i.e., it makes the grid steps purely imaginary. Generally, the problem of optimal approximation on the union of a positive and a negative interval leads to non-Stieltjes

---

[1]We now allow for complex-valued $\widehat{h}_{j-1}, h_j$ $(j = 1, \ldots, n)$ in (3.3), which is different from the classical definition of S-fractions with real positive parameters.

rational functions $R_n(z)$ of type $(n - 1, n)$. Assuming absence of breakdowns (which are unlikely but can not be definitely excluded), the transformation to the non-Stieltjes rational function (3.3) can still be carried out via the complex $2n$-step Euclidean algorithm. We used the bi-Lanczos extension of the Lanczos-based algorithm [16] which, according to our experience, always produced meaningful results.

EXAMPLE 3.1. *We begin with reproducing a real optimal grid from (3.3) generated for a real positive interval $K = [1, 1e4]$, see Figure 3.1 (left). Similar results were reported in [32]. We can consider this example as a degenerate case of the two-interval problem with $m_1 = 0$ and $m_2 = 10$. The plot shows "alternation" of the primal and dual grid points and monotonically growing steps. The grid looks like an equidistant grid stretched by a rather smooth transform. It was shown in [32] that for large n and small interval ratios such transforms are asymptotically close to the exponential.*

In Figure 3.1 (right) we plot the complex finite difference grid points obtained from the continued fraction (3.3) in the case when $K = [-1e3, -1] \cup [1, 1e4]$ and $m_1 = 8$ and $m_2 = 10$. We notice the "alternation" of the primary and dual points on some "curve", which is an intuitive evidence of a good quality of the grid, i.e., we can speculate that the finite difference solution approximates the exact solution with second-order accuracy on that curve. This curve can be interpreted as the complex PML transform of the real positive axis in accordance with [7, 11].

In summary, we observe that the finite-difference operators on grids obtained from (3.3) approximate the second-order derivative operator on curves in the complex plane. This can be viewed as a complex extension of Krein's results on the convergence of the Stieltjes discrete string with impedance $R_n(z)$ to its continuous counterpart with impedance $F(z)$ when $R_n \to F$ on $\mathbb{R}_+$ [34]. Besides internal beauty, this phenomenon may have useful consequences. For example, it lets us hope that pseudospectral estimates and stability results for continuous PMLs and damped 1D differential operators [3, 8, 14] remain valid for (3.1)–(3.2) with the optimal grid.
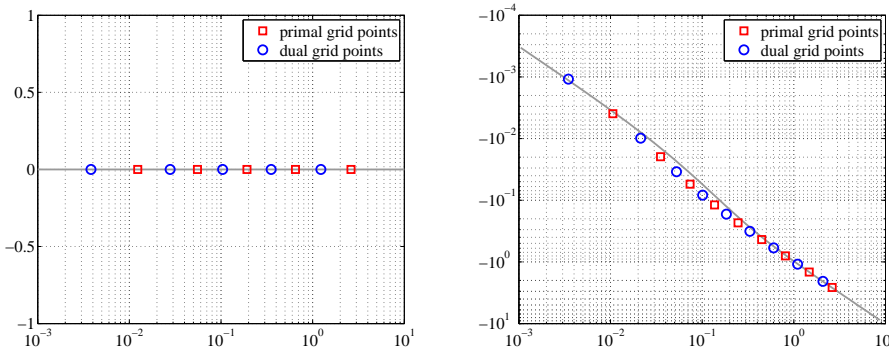

FIGURE 3.1. *Grid points generated from quantities in the continued fraction (3.3). Left: In this single-interval case the set K is chosen as $K = [1, 1e4]$ with $m_2 = 10$ (and $m_1 = 0$). Right: The set K is chosen as $K = [-1e3, -1] \cup [1, 1e4]$ with $m_1 = 8$ and $m_2 = 10$. The gray "continuous" curve has been obtained by connecting the grid points generated with the parameters $m_1 = 27$ and $m_2 = 33$, and we conjecture that the grid points align on a limit curve as $m \to \infty$.*

4. Summary of the algorithm. In the following we provide a step-by-step description for computing the grid steps $\widehat{h}_{j-1}$ and h_j ($j = 1, \ldots, n$) in (3.3).

1. It follows from (2.3) that the numbers

$$-\left(s_j^{(\sqrt{-b_1},\sqrt{-a_1})}\right)^2, \quad j = 1, \ldots, m_1, \quad \text{and} \quad \left(s_j^{(\sqrt{a_2},\sqrt{b_2})}\right)^2, \quad j = 1, \ldots, m_2,$$

   are the interpolation nodes for $R_n(z)$ as an interpolant of $F(z)$. Knowing interpolation nodes and function values, we compute the coefficients of $P_{n-1}(z)$ and $Q_n(z)$ by means of solving the corresponding system of linear algebraic equations in high-precision arithmetic.

2. The poles of the interpolant, i.e., the roots of $Q_n(z)$, can be computed as the eigenvalues of an associated companion matrix, see [22, Subsection 7.4.6]. To solve this eigenvalue problem we use the quasi-version[2] of the QR transformation method [40, § 11.6] and then, if necessary, correct the roots by means of a combination of Laguerre's [40, § 9.5] and Newton's [35] method.

3. Knowing the poles of $R_n(z)$, the corresponding residues are computed.

4. Finally, the grid steps $\widehat{h}_{j-1}$ and h_j $(j = 1, \ldots, n)$ are computed using the recursion formulas [16, (3.4)], with the underlying quasi-analog of an inverse spectral problem for a symmetric tridiagonal matrix (see [16, subsection 3.1, item 3°], [41, theorem 7.2.1]) being solved by a quasi-Lanczos process [12, Ch. 6] with quasi-reorthogonalization. Here we used the well-known connection between the Lanczos and Euclidean algorithms (see, e.g., [31]).

5. Adaptation to a second-order finite difference framework.

5.1. Approximation of the discrete impedance function. So far we have considered the function $F(z) = z^{-1/2}$, which arises when solving the boundary-value problem (1.1) for $x \in [0, +\infty)$. When this problem is seen as an infinite extension of some interior computational domain, the exponential convergence of the interpolant $R_n(z)$ is consistent with a high-order (or even spectral) discretization of the operator acting in this computational domain.

However, it is also possible to compute the NtD map of a discretized version of (1.1) on a uniform infinite grid via rational approximation of a slightly modified function $F_h(z)$ to be determined below. This function will lead to a three-term finite difference scheme which is appropriate for being combined with a standard second-order finite difference discretization in the interior computational domain, because it allows for the elimination of spurious reflections from the PML boundary due to the error of the interior discretization.

Given a fixed step size $h > 0$, let us consider the problem (3.1)–(3.2) on the infinite equidistant grid with $\widehat{h}_0 = 0.5h$ and $\widehat{h}_j = h_j = h$ for $j = 1, \ldots, \infty$. We will determine a function $F_h(z)$ such that

$$\boldsymbol{u}_0 = F_h(\boldsymbol{A})\boldsymbol{b}$$

via a well-known approach widely used in the representation of irrational numbers via continued fractions (see, e.g., [18, section 9]). This approach was already applied in [46] to the infinite lattice problem: the infinite-length S-fraction representation of

---

[2] I.e., we formally use in the complex case the formulas intended for the real case.

F_h analogous to (3.3) is

$$R_h(z) = \cfrac{1}{0.5hz + \cfrac{1}{h + \cfrac{1}{hz + \ddots}}}$$

(for a proof of convergence we refer to [44] or [33, theorem 4.58]). The remainder continued fraction

$$S(z) = \cfrac{1}{h + \cfrac{1}{hz + \cfrac{1}{h + \cfrac{1}{hz + \ddots}}}}$$

evidently satisfies the equation

$$S(z) = \cfrac{1}{h + \cfrac{1}{hz + S(z)}},$$

or equivalently $S(z)^2 + hzS(z) - z = 0$. Since $0.5hz + S(z) = R_n(z)^{-1} = F_h(z)^{-1}$, we have arrived at the quadratic equation

$$F_h(z)^2 = \frac{1}{z + (0.5hz)^2}.$$

We choose the root which converges to the exact impedance $F(z)$ as $h \to 0$, i.e.,

$$F_h(z) = \frac{1}{\sqrt{z + (0.5hz)^2}}. \tag{5.1}$$

This function, which we will refer to as the *discrete impedance function*, approximates with second-order accuracy the exact impedance at the boundary, so being centered, the resulting finite difference scheme is of second order globally.

Analogously to what we had achieved with (1.2) for continuous x, the relation (5.1) allows us to convert the Neumann data $-\boldsymbol{b}$ at $x = 0$ into the Dirichlet data $\boldsymbol{u}_0$ without actually solving the infinite lattice problem.

For a given $h > 0$ let us define $\sigma = \frac{h^2}{4}$. The invertible linear fractional change of variables

$$w = \frac{z}{\sigma z + 1} \tag{5.2}$$

translates the union of a negative and a positive segment $K = [a_1, b_1] \cup [a_2, b_2]$ again into the union of a negative and a positive segment. Let us assume[3] that $-\sigma^{-1} < a_1$.

---

[3] As discussed earlier, the parameter a_1 should be set to a lower bound of $\boldsymbol{A}$'s spectral interval, in which case the condition $-\sigma^{-1} < a_1$ corresponds to the Nyquist sampling criterion of two grid points per wave length. This assumption should be met by any reasonable discretization scheme.

Let P_{n-1}/Q_n denote the rational approximant of theorem 2.3 for the image of K under transformation (5.2). Then

$$\left| \sqrt{w} \cdot \frac{P_{n-1}(w)}{Q_n(w)} - 1 \right| = \left| \sqrt{\frac{z}{\sigma z + 1}} \cdot \frac{P_{n-1}(\frac{z}{\sigma z+1})}{Q_n(\frac{z}{\sigma z+1})} - 1 \right|$$

$$= \left| \sqrt{\frac{z}{\sigma z + 1}} \cdot \frac{P_{n-1}(\frac{z}{\sigma z+1})(\sigma z + 1)^n}{Q_n(\frac{z}{\sigma z+1})(\sigma z + 1)^n} - 1 \right| = \left| F_h(z) \cdot \frac{P_{n-1}(\frac{z}{\sigma z+1})(\sigma z + 1)^{n-1}}{Q_n(\frac{z}{\sigma z+1})(\sigma z + 1)^n} - 1 \right|$$

is small on K, the numerator and the denominator

$$P_{n-1}\left(\frac{z}{\sigma z + 1} \right)(\sigma z + 1)^{n-1}, \qquad Q_n\left(\frac{z}{\sigma z + 1} \right)(\sigma z + 1)^n$$

being polynomials of degrees $\leq n - 1$ and $\leq n$, respectively. We have thereby established a direct relation between the errors of the rational interpolants for $F(z)$ and $F_h(z)$ on transformed compact sets, respectively, with the interpolation nodes being transformed accordingly. This allows us to conclude that we obtain identical convergence rates for both interpolation processes. In particular, Theorem 2.3 holds with $F(z)$ being replaced by $F_h(z)$.

We would like to mention that a rational approximation-based absorbing boundary condition for the infinite lattice was suggested in [46] and combined with a trapezoidal finite element approach in [27]. However, that approach required a modification of the Helmholtz equation by a higher-order term. On the contrary, in our framework the discreteness can be incorporated simply by adjusting the PML grids. Visually these grids look very similar to the ones shown in Figure 3.1, i.e., we can speculate again that they approximate the exact solution $\boldsymbol{u}(x)$ of (1.1) with second-order accuracy on some modified x-curve in the complex plane.

5.2. Matching interior and exterior discretizations via a single grid. Let us consider the second-order infinite equidistant finite difference problem

$$\frac{1}{h}\left(\frac{\boldsymbol{u}_{j+1} - \boldsymbol{u}_j}{h} - \frac{\boldsymbol{u}_j - \boldsymbol{u}_{j-1}}{h} \right) - \boldsymbol{A}\boldsymbol{u}_j = \boldsymbol{q}_j, \qquad j = -\ell, \ldots, -1, 0, 1, \ldots, \infty \quad (5.3)$$

with boundary conditions

$$\boldsymbol{u}_{-\ell-1} = \boldsymbol{0}, \quad \lim_{j \to \infty} \boldsymbol{u}_j = \boldsymbol{0}, \tag{5.4}$$

assuming $\boldsymbol{q}_j = \boldsymbol{0}$ for $j \geq 0$. Problem (5.3) can be split equivalently into an interior finite-dimensional system

$$\frac{1}{h}\left(\frac{\boldsymbol{u}_{j+1} - \boldsymbol{u}_j}{h} - \frac{\boldsymbol{u}_j - \boldsymbol{u}_{j-1}}{h} \right) - \boldsymbol{A}\boldsymbol{u}_j = \boldsymbol{q}_j, \qquad j = -\ell, \ldots, -1, \tag{5.5}$$

$$\frac{1}{0.5h}\left(-\boldsymbol{b} - \frac{\boldsymbol{u}_0 - \boldsymbol{u}_{-1}}{h} \right) - \boldsymbol{A}\boldsymbol{u}_0 = \boldsymbol{0},$$

and an exterior infinite system

$$\frac{1}{0.5h}\left(\frac{\boldsymbol{u}_1 - \boldsymbol{u}_0}{h} + \boldsymbol{b} \right) - \boldsymbol{A}\boldsymbol{u}_0 = \boldsymbol{0}, \tag{5.6}$$

$$\frac{1}{h}\left(\frac{\boldsymbol{u}_{j+1} - \boldsymbol{u}_j}{h} - \frac{\boldsymbol{u}_j - \boldsymbol{u}_{j-1}}{h} \right) - \boldsymbol{A}\boldsymbol{u}_j = \boldsymbol{0}, \qquad j = 1, \ldots, \infty,$$

both systems being coupled via a vector variable $\boldsymbol{b}$.[4]

Problem (5.6) (with the condition at infinity) was already considered in section 5.1, and can be exactly eliminated using the discrete impedance function (5.1),

$$\frac{1}{h}\left(\frac{\boldsymbol{u}_{j+1}-\boldsymbol{u}_j}{h}-\frac{\boldsymbol{u}_j-\boldsymbol{u}_{j-1}}{h}\right)-\boldsymbol{A}\boldsymbol{u}_j=\boldsymbol{q}_j,\qquad j=-\ell,\ldots,-1,$$

$$\frac{1}{0.5h}\left(-F_h(\boldsymbol{A})^{-1}\boldsymbol{u}_0-\frac{\boldsymbol{u}_0-\boldsymbol{u}_{-1}}{h}\right)-\boldsymbol{A}\boldsymbol{u}_0=\boldsymbol{0},\qquad \boldsymbol{u}_{-\ell-1}=\boldsymbol{0}.$$

This formally corresponds to a Schur complement. Upon substitution $R_n(\boldsymbol{A})\approx F_h(\boldsymbol{A})$ we arrive at the approximate problem

$$\frac{1}{h}\left(\frac{\boldsymbol{u}_{j+1}^n-\boldsymbol{u}_j^n}{h}-\frac{\boldsymbol{u}_j^n-\boldsymbol{u}_{j-1}^n}{h}\right)-\boldsymbol{A}\boldsymbol{u}_j^n=\boldsymbol{q}_j,\qquad j=-\ell,\ldots,-1,$$

$$\frac{1}{0.5h}\left(-R_n(\boldsymbol{A})^{-1}\boldsymbol{u}_0^n-\frac{\boldsymbol{u}_0^n-\boldsymbol{u}_{-1}^n}{h}\right)-\boldsymbol{A}\boldsymbol{u}_0^n=\boldsymbol{0},\qquad \boldsymbol{u}_{-\ell-1}^n=\boldsymbol{0}.$$

Hence

$$\|\boldsymbol{u}_j^n-\boldsymbol{u}_j\|=O(\|R_n(\boldsymbol{A})-F_h(\boldsymbol{A})\|),$$

since all the involved linear systems are well posed uniformly in n.

Performing similar manipulations with the approximate problem in reverse order, we obtain the equivalent system (5.7)–(5.8)

$$\frac{1}{h}\left(\frac{\boldsymbol{u}_{j+1}^n-\boldsymbol{u}_j^n}{h}-\frac{\boldsymbol{u}_j^n-\boldsymbol{u}_{j-1}^n}{h}\right)-\boldsymbol{A}\boldsymbol{u}_j^n=\boldsymbol{q}_j,\qquad j=-\ell,\ldots,-1,\qquad(5.7)$$

$$\frac{1}{0.5h}\left(-\boldsymbol{b}-\frac{\boldsymbol{u}_0^n-\boldsymbol{u}_{-1}^n}{h}\right)-\boldsymbol{A}\boldsymbol{u}_0^n=\boldsymbol{0},$$

$$\frac{1}{\widehat{h}_0}\left(\frac{\boldsymbol{u}_1^n-\boldsymbol{u}_0^n}{h_1}+\boldsymbol{b}\right)-\boldsymbol{A}\boldsymbol{u}_0^n=\boldsymbol{0},\qquad(5.8)$$

$$\frac{1}{\widehat{h}_j}\left(\frac{\boldsymbol{u}_{j+1}^n-\boldsymbol{u}_j^n}{h_{j+1}}-\frac{\boldsymbol{u}_j^n-\boldsymbol{u}_{j-1}^n}{h_j}\right)-\boldsymbol{A}\boldsymbol{u}_j^n=\boldsymbol{0},\qquad j=1,\ldots,n-1,$$

by introducing $\boldsymbol{b}$ and fictitious variables $\boldsymbol{u}_j^n$ with positive subindices which, unlike their negative counterparts, do not approximate corresponding components of $\boldsymbol{u}(x)$. Finally, eliminating $\boldsymbol{b}$ we can merge the systems (5.7)–(5.8) into a single recursion

$$\frac{1}{\widehat{h}_j}\left(\frac{\boldsymbol{u}_{j+1}^n-\boldsymbol{u}_j^n}{h_{j+1}}-\frac{\boldsymbol{u}_j^n-\boldsymbol{u}_{j-1}^n}{h_j}\right)-\boldsymbol{A}\boldsymbol{u}_j^n=\boldsymbol{q}_j,\qquad j=-\ell,\ldots,n-1,$$

$$\boldsymbol{u}_{-\ell}^n=\boldsymbol{0},\ \boldsymbol{u}_n^n=\boldsymbol{0},$$

---

[4]Problem (5.3)–(5.4) can be viewed as the second-order discretization of $\frac{\partial^2}{\partial x^2}\boldsymbol{u}-\boldsymbol{A}\boldsymbol{u}=\boldsymbol{q}$, $\boldsymbol{u}\big|_{x=-h(\ell+1)}=\boldsymbol{0}$, $\boldsymbol{u}\big|_{x=+\infty}=\boldsymbol{0}$ for some regular enough $\boldsymbol{q}$ supported on $[-h(\ell+1),0]$. As the infinite exterior problem (5.6) approximates with second-order accuracy the same equation on $[0,+\infty)$ with conditions $\boldsymbol{u}\big|_{x=0}=-\boldsymbol{b}$ and $\boldsymbol{u}\big|_{x=+\infty}=\boldsymbol{0}$, the relation (5.5) approximates with second order the same equation restricted to $[-h(\ell+1),0]$ with conditions $\boldsymbol{u}\big|_{x=-h(\ell+1)}=\boldsymbol{0}$ and $\boldsymbol{u}\big|_{x=0}=-\boldsymbol{b}$.

with the convention that $\widehat{h}_j := h$ for $j < 0$, $h_j := h$ for $j \leq 0$, $\widehat{h}_j := \widehat{h}_j$ for $j > 0$, $h_j := h_j$ for $j > 0$, and $\widehat{h}_0 := \widehat{h}_0 + h/2$ (see also Figure 5.1). This finite difference scheme is easy to implement by simply modifying the n trailing primal and dual grid steps in a given finite difference scheme with step size h. We reiterate that this scheme converges exponentially with error $O(\|R_n(\boldsymbol{A}) - F_h(\boldsymbol{A})\|)$ to the solution of (5.3)–(5.4) in the interior domain, i.e., for the nonpositive subindices.

The above derivation can easily be extended to variable operators $\boldsymbol{A} = \boldsymbol{A}_j$ in the interior domain and tensor-product PML discretizations. This will be illustrated by a numerical example in section 6.2.
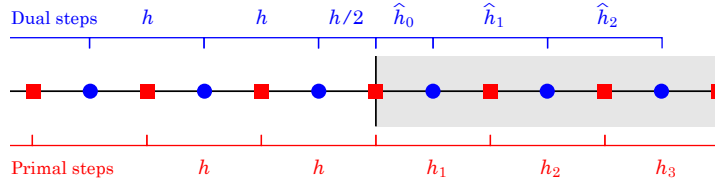

FIGURE 5.1. *Schematic view of a finite difference grid appended with an absorbing boundary layer generated from quantities in the continued fraction (3.3). The example shown here is for the case $n = 3$. The gray-shaded region corresponds to the appended absorbing boundary layer, and the grid steps $\widehat{h}_0, \widehat{h}_1, \ldots, \widehat{h}_{n-1}$ and $h_1, h_2, \ldots, h_n$ in this layer are generally complex.*

6. Numerical experiments.

6.1. Waveguide example. To test the accuracy of our absorbing boundary layer, we consider the inhomogeneous Helmholtz equation

$$\Delta u(x,y) + k^2 u(x,y) = f(x,y)$$

on a rectangular domain $\Omega = [0, L] \times [0, H]$ of length L and height H. We prescribe homogeneous Dirichlet conditions at the upper and lower boundaries in y. The source term is set to

$$f(x,y) = 10 \cdot \delta(x - 511\pi/512) \cdot \delta(y - 50\pi/512),$$

with the Dirac delta function $\delta(\cdot)$.

Our aim is to verify that our absorbing boundary layer models the correct physical behavior. To this end we solve the above Helmholtz equation on two rectangular domains with fixed height $H = \pi$ and different lengths $L = \pi$ and $L = 2\pi$, respectively. See also Figure 6.1 (left and right, respectively). The wave number is chosen as $k = 50$. The problem is discretized by central finite differences with step size $h = \pi/512$ in both coordinate directions. The eigenvalues of the resulting tridiagonal matrix $\boldsymbol{A}$, corresponding to the operator $-\partial^2/\partial y^2 - k^2$ on $[0, \pi]$ with homogeneous Dirichlet boundary conditions, are explicitly known and eigenvalue inclusion intervals are

$$[a_1, b_1] \cup [a_2, b_2] = [-2.50\mathrm{e}3, -1.95\mathrm{e}1] \cup [7.98\mathrm{e}1, 1.04\mathrm{e}5].$$

We extend the interior finite difference grid by our absorbing boundary layer with $n = m/2$ additional grid points to the left of $x = 0$ and to the right of $x = L$, with the near-optimal grid steps computed from a rational interpolant $R_n(z)$ of $F_h(z)$ as explained in section 5.1. The physical domain can hence be thought of as an infinite strip parallel to the x-axis. We therefore expect the solutions of both problems (with

$L = \pi$ and $L = 2\pi$) to coincide when they are restricted to $[0, \pi] \times [0, \pi]$. Visually, this is indeed the case, as one can see in Figure 6.1 (where $n = 10$). Note how the amplitude of the solution is damped very quickly inside the absorbing boundary layer.

To quantify the accuracy of our absorbing boundary layer numerically, we plot in Figure 6.2 the relative uniform norm of the difference of the two numerical solutions $u_1(x, y)$ and $u_2(x, y)$ restricted to $[0, \pi] \times [0, \pi]$, i.e.,

$$\text{err} = \max_{0 \leq x, y \leq \pi} |u_1(x, y) - u_2(x, y)| \Big/ \max_{0 \leq x, y \leq \pi} |u_1(x, y)|. \tag{6.1}$$

Indeed, this figure reveals exponential convergence with the rate ρ given in Theorem 2.3. In this example, the expected rate is $\rho \approx 0.57$ and this is indicated by the slope of the dashed line in Figure 6.2.

We would like to mention that absorbing boundary layers usually require some physical separation from the support of the right-hand side (the source term) [30]. However, thanks to the efficient absorption of evanescent and propagative modes even on spectral subintervals with extreme interval ratios, we are able to place our Dirac source extremely close to the PML boundary (only one grid point away, see the right of Figure 6.1) without deteriorating convergence (see Figure 6.2).

6.2. PML in multiple coordinate directions. In this experiment we demonstrate how our perfectly matched layer can be used to mimic domains which are unbounded in several coordinate directions. Consider the Helmholtz equation

$$c(x, y)\Delta u(x, y) + k^2 u(x, y) = f(x, y)$$

on a square domain $\Omega_1 = [0, 1]^2$. The coefficient $c(x, y)$ is set to 1 everywhere except for a U-shaped region inside Ω_1, where $c(x, y) = 0$ (see Figure 6.3). The wave number is chosen as $k = 120$ and the source term is set to

$$f(x, y) = \delta(x - 140/400) \cdot \delta(y - 260/400).$$

We discretize Ω_1 by central finite differences with step size $h = 1/400$ in both coordinate directions. We aim to append absorbing boundary layers with $n \in \{7, 9, 11, 13\}$ grid points at each of the four edges of Ω_1. For constructing the absorbing layers in the x-direction (to the left of $x = 0$ and to the right of $x = 1$) we need inclusion intervals for the negative and positive eigenvalues of $\boldsymbol{L}_y - k^2 \boldsymbol{I}$, where $\boldsymbol{L}_y$ is the discretization of $-\partial^2/\partial y^2$ on $[0, 1]$ with homogeneous Dirichlet boundary. Possible intervals are

$$[a_1, b_1] \cup [a_2, b_2] = [-1.45\text{e}4, -2.58\text{e}2] \cup [2.37\text{e}2, 5.99\text{e}5]. \tag{6.2}$$

Likewise, for constructing the absorbing layers in the y-direction (below $y = 0$ and above $y = 1$) we can use the same inclusion intervals (6.2) for the negative and positive eigenvalues of $\boldsymbol{L}_x - k^2 \boldsymbol{I}$, where $\boldsymbol{L}_x$ is the finite difference discretization of $-\partial^2/\partial x^2$ on $[0, 1]$ with homogeneous Dirichlet boundary (the matrices $\boldsymbol{L}_x$ and $\boldsymbol{L}_y$ are identical in this example, although this is not a requirement). From the intervals in (6.2) we can now calculate the grid steps of absorbing boundary layers in the x- and y-directions, and then modify the finite difference matrices to $\widehat{\boldsymbol{L}}_x$ and $\widehat{\boldsymbol{L}}_y$, respectively. As in the previous example, this is done by computing a rational interpolant $R_n(z)$ of $F_h(z)$ defined in section 5.1.

However, there is a small subtlety one has to be aware of with the approach just described: effectively, the NtD operators are now given as $F_h(\widehat{\boldsymbol{L}}_x - k^2 \boldsymbol{I})$ and
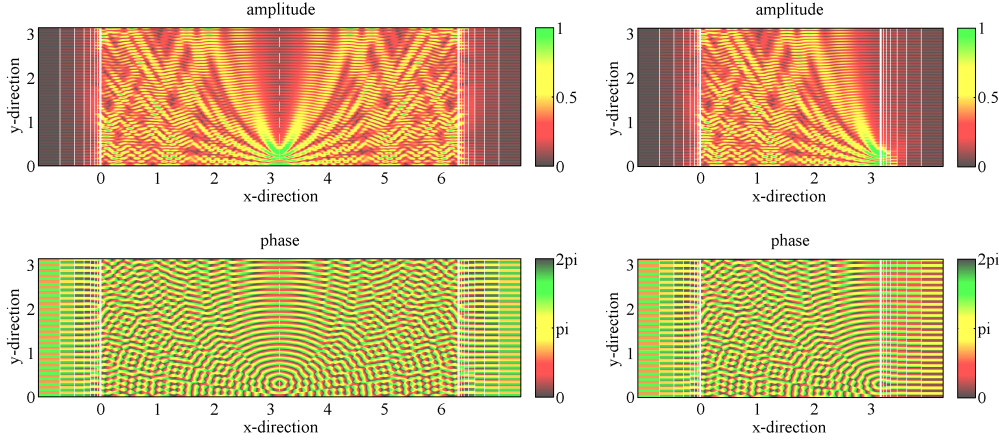
FIGURE 6.1. *Amplitude (top) and phase (bottom) of the solution to the waveguide problem in section 6.1 on two rectangular domains (left/right) which differ in their length. The left domain is of length $L = 2\pi$ in the x-direction, whereas the right domain is of length $L = \pi$. Both domains have been appended with absorbing layers at the left and right boundaries. As the absorbing boundary layers serve the purpose of extending the physical domain towards infinity, both solutions are expected to coincide on the restriction to $x \in [0, \pi]$. In these pictures we have chosen $m = 20$, so there are $n = 10$ points appended to the left and right boundaries. The step size in the interior domain is $h = \pi/512$ in both coordinate directions.*
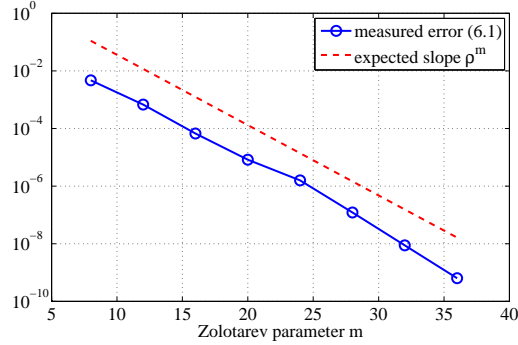

FIGURE 6.2. *Exponential convergence of the accuracy of the absorbing boundary layers for the waveguide problem in section 6.1 with varying Zolotarev parameter $m \in \{8, 12, \ldots, 36\}$ (twice the number of grid points in each absorbing boundary layer). The expected convergence rate $\rho \approx 0.57$ by Theorem 2.3 is indicated by the dashed line.*

$F_h(\widehat{\boldsymbol{L}}_y - k^2 \boldsymbol{I})$, respectively, and the involved matrices are no longer Hermitian. In Figure 6.4 (left) we show the eigenvalues of $\widehat{\boldsymbol{L}}_x$ (or $\widehat{\boldsymbol{L}}_y$, which is the same). We observe that the eigenvalues have "lifted off" the real axis into the upper half of the complex plane, in agreement with the analysis of [14] for continuous one-dimensional damped operators. From Figure 2.1 we find at least visually that the Zolotarev approximant is of a good quality in this region as well, and the accuracy of the resulting absorbing boundary layer should be satisfactory.

To quantify the accuracy numerically, we solve the same Helmholtz problem on a smaller domain $\Omega_2 = [0.25, 0.75]^2$, again appended with absorbing boundary layers of n grid points at each of the four edges of Ω_2. As the source term $f(x, y)$ is supported

inside Ω_2, we expect coinciding solutions $u_1(x, y)$ and $u_2(x, y)$ on their restrictions to Ω_2. In Figure 6.4 (right) we have plotted the relative uniform norm of the difference of both solutions, i.e.,

$$\text{err} = \max_{0.25 \le x,y \le 0.75} |u_1(x, y) - u_2(x, y)| \Big/ \max_{0.25 \le x,y \le 0.75} |u_1(x, y)|. \tag{6.3}$$

Again we observe exponential convergence, and the reduction of the measured error is in good agreement with (even slightly better than) the rate $\rho = 0.55$ expected from Theorem 2.3.
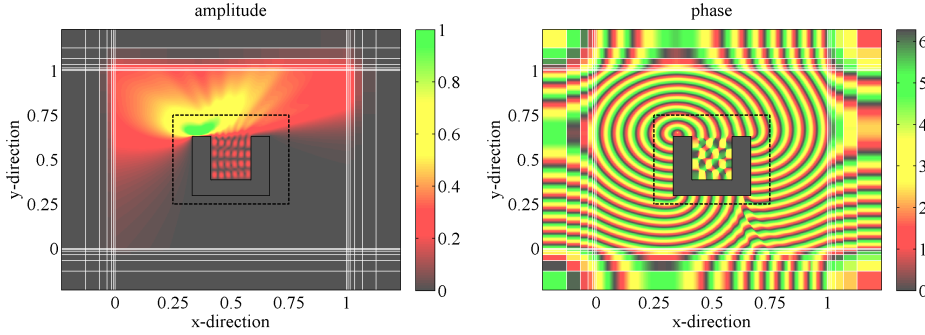

FIGURE 6.3. *Amplitude (left) and phase (right) of the solution to Helmholtz problem in section 6.2 on a square domain $\Omega_1 = [0, 1]^2$ appended with absorbing boundary layers at all boundary edges. In these pictures we have chosen the Zolotarev parameter $m = 14$, so there are $n = 7$ grid points appended to the boundaries. The step size in the interior domain is $h = 1/400$ in both coordinate directions. The dashed square in the interior indicates the smaller domain $\Omega_2 = [0.25, 0.75]^2$, on which we solve the same Helmholtz problem for assessing the numerical accuracy of our absorbing boundary layers.*
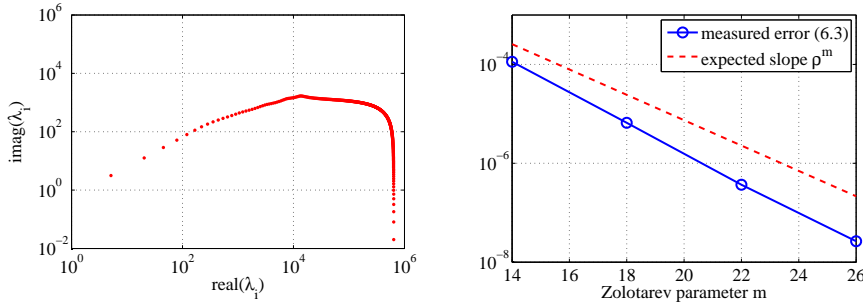

FIGURE 6.4. *Left: Eigenvalues of the matrix $\widehat{L}_x$ (same as those of $\widehat{L}_y$) associated with the Helmholtz problem in section 6.2, appended with $n = 7$ grid points at the boundaries. Right: Exponential convergence of the accuracy of the absorbing boundary layers with varying Zolotarev parameter $m \in \{14, 18, \ldots, 26\}$ (twice the number of grid points in each absorbing boundary layer). The expected convergence rate $\rho \approx 0.55$ by Theorem 2.3 is indicated by the dashed line.*

7. Summary and open problems. We have presented a new approach for the construction of discrete absorbing boundary layers for indefinite Helmholtz problems via complex coordinate transforms. This approach is based on the use of near-optimal relative rational interpolants of the inverse square root (or a modification thereof) on a negative and a positive real interval. Bounds for the approximation error have been

derived, and the exponential convergence of the approximants has been established theoretically and demonstrated at numerical examples. Although our focus in this paper was on absorbing boundary conditions for indefinite Helmholtz problems, it was recently understood that these conditions also constitute good approximations to Schur complements of certain PDE discretization matrices, and they became a crucial component of modern Helmholtz preconditioners, such as Schwarz domain decomposition [10,21] and the sweeping preconditioner in [20]. Preliminary results have shown successful application to a multilevel domain decomposition preconditioner, and a related Schlumberger patent application is pending.

The uniform approximation approach requires bounds for the smallest/largest negative and positive eigenvalues, which can be rather loose due to the weak dependence of the convergence rate of the Zolotarev approximants on the interval ratios. The external bounds of the intervals can thus be estimated roughly. Still, the numerical estimation of the internal bounds can be rather difficult, and accidentally at least one eigenvalue may be very close to the origin, in which case even an optimal approximant may require significant order for a satisfactory accuracy. To circumvent this problem, it would be interesting to derive a parameter-free near-optimal rational approximant of $\boldsymbol{A}^{-1/2}\boldsymbol{b}$, which takes into account the discrete nature of the spectrum of $\boldsymbol{A}$ and the spectral weights of the vector $\boldsymbol{b}$. Promising first steps have been made by using adaptive rational Krylov algorithms [28, 29] for this purpose.

In its present form the developed PML can not be applied directly in combination with time-stepping methods; however, it can be used efficiently for time-domain simulations with model order reduction [17].

Finally, we would like to point out that cloaking problems (which are popular in the inverse problems community) are closely related with the construction of PMLs, because the latter can be viewed as cloaking of the point at infinity. Cloaking problems can also be formulated via complex coordinate transforms [36] and lead to the approximation of NtD maps. Although the involved Stieltjes impedance function $F(z)$ is typically different in these applications, techniques similar to those presented in this paper may still be applicable.

Appendix A. Zolotarev approximation and proof of Theorem 2.3.
The solution of the Zolotarev problem (2.4) can be computed as

$$Z_m^{(c,d)}(z) = \prod_{j=1}^{m}(z - s_j^{(c,d)}), \qquad s_j^{(c,d)} = d \cdot \mathrm{dn}\left(\frac{(2m - 2j + 1)\mathrm{K}(\delta')}{2m}, \delta'\right), \qquad \text{(A.1)}$$

where

$$\delta = c/d, \quad \delta' = \sqrt{1 - \delta^2},$$

$$\mathrm{K}(\delta) = \int_0^1 \frac{1}{\sqrt{(1 - t^2)(1 - \delta^2 t^2)}}\, \mathrm{d}t$$

is the complete elliptic integral of the first kind[5] and where the Jacobian elliptic

---

[5]The definition of $K(\delta)$ is not consistent in the literature. We stick to the definition used in [39, Ch. VI]. In MATLAB one would type `ellipke(delta^2)` to obtain the value $K(\delta)$.

function dn is defined via another such function, sn, by the relations

$$\mathrm{dn}(u, \kappa) = \sqrt{1 - \kappa^2 \, \mathrm{sn}(u, \kappa)}, \qquad \xi = \mathrm{sn}(u; \kappa), \qquad u = \int_0^\xi \frac{dt}{\sqrt{(1 - t^2)(1 - \kappa^2 t^2)}}.$$

In order to prove near-optimality results, we first need to study the quantity $E_m^{(c,d)}$ in (2.4) carefully. Evidently, $E_m^{(c,d)} < 1$. Upper and lower bounds for (2.4) were given in [38] as

$$\frac{2 \exp\left(-\frac{\pi \mathrm{K}(\mu')}{4\mathrm{K}(\mu)} m\right)}{1 + \left[\exp\left(-\frac{\pi \mathrm{K}(\mu')}{4\mathrm{K}(\mu)} m\right)\right]^2} \leq E_m^{(c,d)} \leq 2 \exp\left(-\frac{\pi \mathrm{K}(\mu')}{4\mathrm{K}(\mu)} m\right) \tag{A.2}$$

with

$$\mu = \left(\frac{1 - \sqrt{\delta}}{1 + \sqrt{\delta}}\right)^2 \quad \text{and} \quad \mu' = \sqrt{1 - \mu^2}.$$

Hence the Cauchy–Hadamard convergence rate can be computed as

$$\rho^{(\delta)} = \exp\left(-\frac{\pi \mathrm{K}(\mu')}{4\mathrm{K}(\mu)}\right). \tag{A.3}$$

Recalling the equalities (2.6) and (2.8), let us define the sets

$$\widetilde{K} = F(K), \quad \widetilde{K}_1 = F(K_1), \quad \widetilde{K}_2 = F(K_2),$$

and consider the following auxiliary problem: find a (complex) monic polynomial H_m of degree m being the minimizer of

$$\min_{H \in \mathcal{P}_m} \max_{s \in \widetilde{K}} \left| \frac{H(s)}{H(-s)} \right|. \tag{A.4}$$

We now construct an approximate solution of this problem and show that the approximate solution gives the maximum in (A.4) which yields the best possible functional value up to a moderate multiplier.

Accounting, as it was done in [15, Section 2], that

$$\left| \frac{Z_{m_1}^{(\sqrt{-b_1}, \sqrt{-a_1})}(-is)}{Z_{m_1}^{(\sqrt{-b_1}, \sqrt{-a_1})}(is)} \right| = 1 \qquad \text{if} \quad s \in \widetilde{K}_2$$

and

$$\left| \frac{Z_{m_2}^{(\sqrt{a_2}, \sqrt{b_2})}(s)}{Z_{m_2}^{(\sqrt{a_2}, \sqrt{b_2})}(-s)} \right| = 1 \qquad \text{if} \quad s \in \widetilde{K}_1$$

because these polynomials have real coefficients, the polynomial H_m defined in (2.5) satisfies

$$\max_{s \in \widetilde{K}_1} \left| \frac{H_m(s)}{H_m(-s)} \right| = E_{m_2}^{(\sqrt{a_2}, \sqrt{b_2})} \tag{A.5}$$

and

$$\max_{s\in\widetilde{K}_2}\left|\frac{H_m(s)}{H_m(-s)}\right| = E_{m_1}^{(\sqrt{-b_1},\sqrt{-a_1})}. \tag{A.6}$$

LEMMA A.1. *The polynomial H_m defined in (2.5) satisfies the inequality*

$$\max_{s\in\widetilde{K}}\left|\frac{H_m(s)}{H_m(-s)}\right| \le 2\max\left\{\rho_1^{-1/2},\rho_2^{-1/2}\right\}\rho^m \tag{A.7}$$

with the numbers ρ_1,ρ_2 and ρ defined in (2.6) and (2.8), provided that m_1,m_2 are chosen according to (2.7).

On the other hand, for any complex polynomial $H\in\mathcal{P}_m$ we have

$$\max_{s\in\widetilde{K}}\left|\frac{H(s)}{H(-s)}\right| \ge \rho^m. \tag{A.8}$$

Proof. Let H_m be defined as in (2.5) and conditions (2.7) be satisfied. Accounting for (A.5), (A.6) and (A.2), we obtain

$$\max_{s\in\widetilde{K}}\left|\frac{H_m(s)}{H_m(-s)}\right| \le 2\max\left\{\rho_1^{m_1},\rho_2^{m_2}\right\} = 2\rho^m\max\left\{\rho_1^{\theta},\rho_2^{-\theta}\right\},$$

which gives assertion (A.7).

To prove assertion (A.8), we consider the third Zolotarev problem in the complex plane for the condenser $\left(\widetilde{K},-\widetilde{K}\right)$ (see [23], [48, § 8.7] or [42, § VIII.3]). Due to the symmetry of the condenser, the two measures forming the (unique) equilibrium pair for $\left(\widetilde{K},-\widetilde{K}\right)$ are symmetric to each other in the evident sense. Thus, one can choose an (in the Cauchy–Hadamard sense) optimal sequence of type (m,m) rational functions of the form $H(s)/H(-s)$, $\deg(H)=m\ge 1$, such that the roots s_j $(1\le j\le m)$ of each polynomial H belong to $\widetilde{K}$. Define

$$H^{(1)}(s)=\prod_{\substack{1\le j\le m\\ s_j\in\widetilde{K}_1}}(s-s_j),\qquad \deg(H^{(1)})=m_1,$$

$$H^{(2)}(s)=\prod_{\substack{1\le j\le m\\ s_j\in\widetilde{K}_2}}(s-s_j),\qquad \deg(H^{(2)})=m_2,$$

$$m_1+m_2=m.$$

By virtue of (A.2) and the location of the roots we have

$$\max_{s\in\widetilde{K}_1}\left|\frac{H(s)}{H(-s)}\right| = \max_{s\in\widetilde{K}_1}\left|\frac{H^{(1)}(s)}{H^{(1)}(-s)}\right| \ge \frac{2\rho_1^{m_1}}{1+\rho_1^{2m_1}}$$

and

$$\max_{s\in\widetilde{K}_2}\left|\frac{H(s)}{H(-s)}\right| = \max_{s\in\widetilde{K}_2}\left|\frac{H^{(2)}(s)}{H^{(2)}(-s)}\right| \ge \frac{2\rho_2^{m_2}}{1+\rho_2^{2m_2}},$$

whence

$$\max_{s\in\widetilde{K}}\left|\frac{H(s)}{H(-s)}\right|\cdot\max_{s\in-\widetilde{K}}\left|\left[\frac{H(s)}{H(-s)}\right]^{-1}\right|=\max_{s\in\widetilde{K}}\left|\frac{H(s)}{H(-s)}\right|^2\geq\max\left\{\frac{2\rho_1^{m_1}}{1+\rho_1^{2m_1}},\frac{2\rho_2^{m_2}}{1+\rho_2^{2m_2}}\right\}^2$$

$$\geq\max\left\{\rho_1^{2m_1},\rho_2^{2m_2}\right\}.$$

Since the quantity $\max\left\{\rho_1^{2x_1},\rho_2^{2x_2}\right\}$ under the conditions $x_1\geq 0,\ x_2\geq 0,\ x_1+x_2=m$ is minimal at

$$x_1=m\cdot\frac{\log\rho_2}{\log\rho_1+\log\rho_2},\qquad x_2=m\cdot\frac{\log\rho_1}{\log\rho_1+\log\rho_2},$$

we obtain

$$\max_{s\in\widetilde{K}}\left|\frac{H(s)}{H(-s)}\right|\cdot\max_{s\in-\widetilde{K}}\left|\left[\frac{H(s)}{H(-s)}\right]^{-1}\right|\geq\rho^{2m}\qquad\text{as}\quad m\to\infty,$$

so

$$\liminf_{m\to\infty}\left(\max_{s\in\widetilde{K}}\left|\frac{H(s)}{H(-s)}\right|\cdot\max_{s\in-\widetilde{K}}\left|\left[\frac{H(s)}{H(-s)}\right]^{-1}\right|\right)^{1/m}\geq\rho^2.$$

It follows in view of [23, Theorem 1, Formula (12)] that the logarithmic capacity of our condenser satisfies

$$\exp\left(-1/\operatorname{cap}\left(\widetilde{K},-\widetilde{K}\right)\right)\geq\rho^2.$$

Moreover, [23, Theorem 1, Formula (11)] yields for all $H,G\in\mathcal{P}_m$

$$\max_{s\in\widetilde{K}}\left|\frac{H(s)}{G(s)}\right|\cdot\max_{s\in-\widetilde{K}}\left|\frac{G(s)}{H(s)}\right|\geq\rho^{2m},$$

from which (A.8) follows. □

We are now prepared to conclude the proof of Theorem 2.3.

Proof. To establish (2.10), it suffices to note that

$$\max_{z\in K}\left|\frac{R_n(z)}{F(z)}-1\right|=\max_{s\in\widetilde{K}}\left|\frac{s\,P_{n-1}(s^2)}{Q_n(s^2)}-1\right|,$$

and to apply (A.7) from Lemma A.1, condition (2.9), and a consequence of (2.3) for finding

$$\left|\frac{sP_{n-1}(s^2)}{Q_n(s^2)}-1\right|=\frac{2\left|\frac{H_m(s)}{H_m(-s)}\right|}{\left|1+\frac{H_m(s)}{H_m(-s)}\right|}.$$

To justify (2.11), set $z=s^2$ and assume that for some pair (P,Q) and $R=P/Q$ we have the inequality

$$\max_{z\in K}\left|\frac{R(z)}{F(z)}-1\right|<\frac{2\rho^m}{1+\rho^m}.$$

Define H by means of (2.1) and rewrite the equality (2.3) in the form

$$\frac{H(s)}{H(-s)} = -\frac{\frac{R(z)}{F(z)} - 1}{\left(\frac{R(z)}{F(z)} - 1\right) + 2}.$$

We readily derive

$$\max_{s \in \widetilde{K}} \left| \frac{H(s)}{H(-s)} \right| < \frac{\frac{2\rho^m}{1+\rho^m}}{2 - \frac{2\rho^m}{1+\rho^m}} = \rho^m,$$

which contradicts (A.8) and thereby proves the assertion (2.11). □

REFERENCES

[1] A. ABUBAKAR, T. M. HABASHY, V. L. DRUSKIN, L. KNIZHNERMAN, AND D. ALUMBAUGH, *2.5D forward and inverse modeling for interpreting low-frequency electromagnetic measurements*, Geophysics 73:4 (2008), pp. F165–F177.

[2] N. I. AKHIEZER, *Theory of Approximation*, Dover, 1992.

[3] D. APPELÖ, T. HAGSTROM, AND G. KREISS, *Perfectly matched layers for hyperbolic systems: General formulation, well-posedness, and stability*, SIAM J. Appl. Math., 67 (2006), pp. 1–23.

[4] S. ASVADUROV, V. DRUSKIN, M. GUDDATI, AND L. KNIZHNERMAN, *On optimal finite-difference approximation of PML*, SIAM J. Numer. Anal., 41 (2003), pp. 287–305.

[5] S. ASVADUROV, V. DRUSKIN, AND S. MOSKOW, *Optimal grids for anisotropic problems*, Electron. Trans. Numer. Anal., 56 (2007), pp. 55–81.

[6] D. N. BAILEY, *A Fortran-90 based multiprecision system*, RNR Technical Report RNR-94-013, NAS Scientific Computation Branch, NASA, Ames Research Center, CA, USA (1994).

[7] E. BALSLEV AND J. COMBES, *Spectral properties of many body Schrödinger operators with dilation analytic interactions*, Commun. Math. Phys., 22 (1971), pp. 280–294.

[8] E. BECACHE AND P. JOLY, *On the analysis of Berenger's perfectly matched layers for Maxwell's equations*, Math. Model. Num. Anal., 36 (2002), pp. 87–119.

[9] J. P. BERENGER, *A perfectly matched layer for the absorption of electromagnetic waves*, J. Comp. Phys., 114 (1994), pp. 185–200.

[10] Y. BOUBENDIR, X. ANTOINE, AND C. GEUZAINE, *A quasi-optimal non-overlapping domain decomposition algorithm for the Helmholtz equation*, J. Comp. Phys., 231 (2012), pp. 262–280.

[11] W. CHEW AND B. WEEDON, *A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates*, Microwave Opt. Technol. Lett., 7 (1994), pp. 599–604.

[12] J. CULLUM AND R. A. WILLOUGHBY, *Lanczos Algorithms for Large Symmetric Eigenvalue Computations. Volume 1, Theory*, SIAM, Philadelphia, 2002.

[13] S. DAVYDYCHEVA, V. DRUSKIN, AND T. HABASHY, *An efficient finite-difference scheme for electromagnetic logging in 3D anisotropic inhomogeneous media*, Geophysics, 68 (2003), pp. 1525–1536.

[14] T. A. DRISCOLL AND L. N. TREFETHEN, *Pseudospectra for the wave equation with an absorbing boundary*, J. Comput. Appl. Math., 69 (1996), pp. 125-142.

[15] V. DRUSKIN, M. GUDDATI, AND T. HAGSTROM, *On generalized discrete PML optimized for propagative and evanescent waves*, arXiv:1210.7862v1 [math.NA], 29 Oct 2012.

[16] V. DRUSKIN AND L. KNIZHNERMAN, *Gaussian spectral rules for the three-point second differences: I. A two-point positive definite problem in a semiinfinite domain*, SIAM J. Numer. Anal., 37 (1999), pp. 403–422.

[17] V. DRUSKIN AND R. REMIS, *A Krylov stability-corrected coordinate-stretching method to simulate wave propagation in unbounded domains*, SIAM J. Sci. Comput., 35 (2013), pp. 313–357.

[18] S. ELAYDI, *An Introduction to Difference Equations*, Springer, New York, 2005.

[19] B. ENGQUIST AND A. MAJDA, *Radiation boundary conditions for acoustic and elastic wave calculations*, Comm. Pure Appl. Math., 32 (1979), pp. 313–357.

[20] B. ENGQUIST AND L. YING, *Sweeping preconditioner for the Helmholtz equation: hierarchical matrix representation*, Comm. Pure Appl. Math., 64:5 (2011), pp. 697–735.

[21] M. Gander, *Optimized Schwarz methods*, SIAM J. Numer. Anal., 44:2 (2006), pp. 699–731.

[22] G. H. Golub and C. F. van Loan, *Matrix computations*, The Johns Hopkins Univ. Press, Baltimore and London, 1989.

[23] A. A. Gonchar, *Zolotarev problems connected with rational functions*, Math. Digest (Matem. Sbornik), 7 (1969), pp. 623–635.

[24] A. A. Gonchar, *The rate of rational approximation of analytic functions*, Proc. Steklov Inst. Math., 166 (1984), pp. 52–60.

[25] A. A. Gonchar and E. A. Rakhmanov, *Equilibrium distributions and degree of rational approximation of analytic functions*, Sb. Math., 176 (1987), pp. 306–352.

[26] M. N. Guddati and K.-W. Lim, *Continued fraction absorbing boundary conditions for convex polygonal domains*, Int. J. Numer. Meth. Engng., 66 (2006), pp. 949–977.

[27] M. Guddati and J. Tassoulas, *Continued-fraction absorbing boundary conditions for the wave equation*, J. Comp. Acoust., 8 (2000), pp. 139–156.

[28] S. Güttel, *Rational Krylov approximation of matrix functions: Numerical methods and optimal pole selection*, GAMM-Mitt., 36:1 (2013), pp. 8–31.

[29] S. Güttel and L. Knizhnerman, *A black-box rational Arnoldi variant for Cauchy-Stieltjes matrix functions*, BIT Numer. Math., 53:3 (2013), pp. 595–616.

[30] T.Hagstrom and T. Warburton, *Complete radiation boundary conditions: minimizing the long time error growth of local methods*, SIAM J. Numer. Anal., 47:5 (2009), pp. 3678–3704.

[31] O. Holtz and M. Tyaglov, *Structured matrices, continued fractions, and root localization of polynomials*, SIAM Rev., 54:3 (2012), pp. 421–509.

[32] D. Ingerman, V. Druskin, and L. Knizhnerman, *Optimal finite difference grids and rational approximations of the square root, I. Elliptic problems*, Comm. Pure Appl. Math., 53 (2000), pp. 1039–1066.

[33] W. B. Jones and W. J. Thron, *Continued Fractions. Analytic Theory and Applications*, Enc. of Math. and its Applic., v. 11, Addison-Wesley P. C., London et al., 1980.

[34] I. S. Kac and M. G. Krein, *On the spectral functions of the string*, Amer. Math. Soc. Transl., 103 (1974), pp. 19–102.

[35] C. T. Kelley, *Solving nonlinear equations with Newton's method*, Fundamentals of Algorithms, No. 1, SIAM, Philadelphia, 2003.

[36] R. V. Kohn, D. Onofrei, M. S. Vogelius and M. Weinstein, *Cloaking via change of variables for the Helmholtz equation*, Comm. Pure Appl. Math., 63 (2010), pp. 973–1016.

[37] V. Lisitsa, *Optimal discretization of PML for elasticity problems*, Electron. Trans. Numer. Anal., 30 (2008), pp. 258–277.

[38] A. A. Medovikov and V. I. Lebedev, *Optimization of L_ω stable Crank–Nicolson method*, Russian J. Numer. Anal. Math. Modelling, 20 (2005), pp. 283–304.

[39] Z. Nehari, *Conformal mapping*, Dover, New York, 1975.

[40] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, Cambridge Univ. Press, Cambridge, 1992.

[41] B. N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, SIAM, Philadelphia, 1998.

[42] E. B. Saff and V. Totik, *Logarithmic Potentials with External Fields*, Springer, Berlin, 1997.

[43] H. Stahl, *Orthogonal polynomials with complex valued weight function, I*, Constr. Approx., 2 (1986), pp. 225–240.

[44] T. J. Stieltjes, *Recherches sur les fractions continues*, Annales de la Faculté des Sciences de Toulouse, 8 (1984), p. 1–122, 9 (1894), p. 1–47.

[45] M. E. Taylor, *Partial Differential Equations II. Qualitative Studies of Linear Equations*, Springer, New York, 1996.

[46] S. Thirunavukkarasu and M. N Guddati, *Absorbing boundary conditions for time harmonic wave propagation in discretized domains*, Comput. Methods Appl. Mech. Engrg., 200:33 (2011), pp. 2483–2497.

[47] R. Varga, *Scientific computation on mathematical problems and conjectures*, CBMS-NSF, SIAM, 1990.

[48] J. L. Walsh, *Interpolation and Approximation by Rational functions in the Complex Domain*, AMS, Rhode Island, 1960.

[49] M. A. Zahid and M. N. Guddati *Padded continued fraction absorbing boundary conditions for dispersive waves*, Comput. Methods Appl. Mech. Engrg., 195 (2006), pp. 3797–3819.

[50] M. Zaslavsky, V. Druskin, S. Davydycheva, L. Knizhnerman, A. Abubakar, and T. Habashy, *Hybrid finite-difference integral equation solver for 3D frequency domain anisotropic electromagnetic problems*, Geophysics 76:2 (2011), pp. F123-F137.

[51] Y. I. Zolotarev, *Collection of Works*, Saint Petersburg Academy of Sciences, 30:5, 1877.