Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

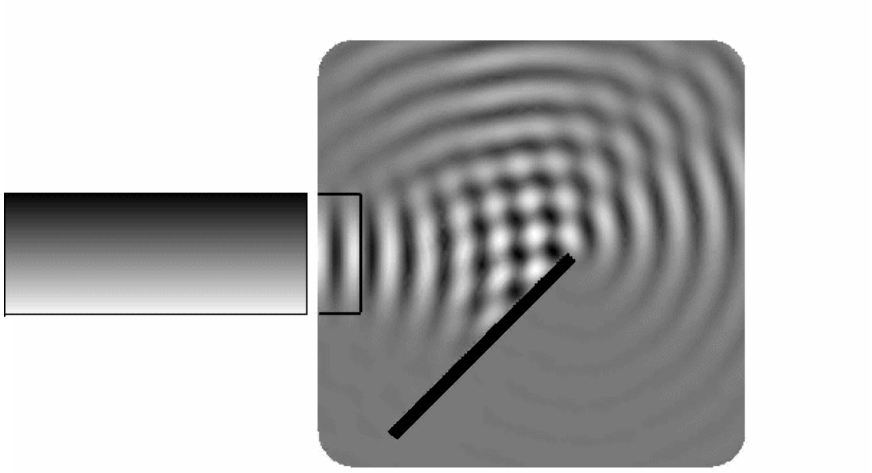
FRANK SCHMIDT

A New Approach to Coupled Interior-Exterior Helmholtz-Type Problems: Theory and Algorithms

ZIB-Report 02-45 (December 2002)

A New Approach to Coupled Interior-Exterior Helmholtz-Type Problems: Theory and Algorithms

Frank Schmidt



Key words and phrases. Pole Condition; Helmholtz equation; Wide angle one-way equations; Schrödinger equation; Unbounded domains; Transparent boundary Conditions; Finite-element method

Acknowledgement

It is a pleasure to thank all the individuals who contributed to this work. Among those I would like to thank in particular

Peter Deuflhard for making this work possible. Doubtless, without his continuous support, guidance and encouragement I would never have been in a position to reach the current state of the project. Together we overcame the first difficulties in theory and algorithms.

David Yevick for forcing the application of the theory to one-way wide angle equations.

Reinhard März for discussing the properties of periodic structures.

My colleagues Thorsten Hohage and Lin Zschiedrich for their deep interest in the subject, proofreading the manuscript, and the countless number of fruitful discussions, which intensely influenced the whole representation, especially the theoretical statements of Sections 4.2 and 4.3.

And, last but not least, my family who made it all possible in the first place.

Notations

In physical equations like Maxwell equations, we follow the usual convention and represent two- and three-dimensional vectorial quantities like the radius vector or the magnetic field strength by lower and upper case bold roman letters, in the given case for example by $\mathbf{x}$ and $\mathbf{H}$, respectively. In the algorithmic part, where the language of numerical linear algebra is used, we represent vectors by lower case roman letters, and matrices by upper case Greek or roman letters. In both cases, lower case Greek letters and lower case Italic letters represent scalars.

Scalars and Vectors

x, y, z	Cartesian coordinates
t	time
$\mathbf{x}$	radius vector; $\mathbf{x} \in \mathbb{R}^d$, $d = 1, 2, 3$
$\mathbf{x}^0$	unit vector $\mathbf{x}/\left\|\mathbf{x}\right\|$, $\mathbf{x} \in \mathbb{R}^d$
$\mathbf{n}$	normal vector pointing to the exterior
k	wavenumber; k^2- sometimes called potential
$\mathbf{k}$	wavevector
λ	wavelength
$\overline{s}$	complex conjugate number to s

Sets

$\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$	natural, real, integer, and complex numbers
$\mathbb{R}_+$	the set $\{x \in \mathbb{R} : 0 \leq x < \infty\}$
$\mathbb{N}_0$	$N \cup \{0\}$
$[a, b]$	closed interval, the set $\{x \in \mathbb{R} : a \leq x \leq b\}$
$]a, b]$	half-open interval, the set $\{x \in \mathbb{R} : a < x \leq b\}$
$]a, b[$	open interval, the set $\{x \in \mathbb{R} : a < x < b\}$
S^d	unit sphere, $S^d = \left\{\mathbf{x} \in \mathbb{R}^{d+1} : \left\|\mathbf{x}\right\| = 1\right\}$
Ω	computational domain, $\Omega \in \mathbb{R}^d$
$\partial\Omega$	boundary of Ω
$\overline{\Omega}$	closure of Ω
$\operatorname{span} S$	linear hull of the set S
$\operatorname{meas} S$	measure of the set S
ext Ω	exterior of Ω, $\mathbb{R}^2 \setminus \overline{\Omega}$
$X \times Y$	product set $\{(x, y) : x \in X \text{ and } y \in Y\}$

Derivatives

∂_x, ∂_y, ∂_z	partial derivatives, $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$, $\frac{\partial}{\partial z}$
$\frac{\partial^{\left\|\alpha\right\|}}{\partial x^\alpha}$	partial derivative in multi-index notation, $\frac{\partial^{\left\|\alpha\right\|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}$
α	multi-index, $\alpha = (\alpha_1, \ldots, \alpha_n)$ with $\alpha_i \in \mathbb{Z}$, $\alpha_i \geq 0$
$\left\|\alpha\right\|$	$\alpha_1 + \ldots + \alpha_n$
∇	nabla operator $\nabla = (\partial_x, \partial_y, \partial_z)$ in Cartesian coordinates

$\operatorname{grad} u$	gradient; in Cartesian coordinates $\nabla u = (\partial_x u, \partial_y u\, \partial_z u)^T$
$\nabla_{\xi\eta} u$	gradient with respect to general $\xi\eta$-coordinates
$\operatorname{curl} \mathbf{u}$	curl of the vector $\mathbf{u}$; in Cartesian coordinates $\nabla \times \mathbf{u}$
$\operatorname{div} \mathbf{u}$	divergence of the vector $\mathbf{u}$; in Cartesian coordinates $\nabla \cdot \mathbf{u}$
∂_n	derivative in the direction of the exterior normal
Δ	Laplacian, $\partial_x^2 + \partial_y^2 + \partial_z^2$

Operators

$\mathcal{D}(A)$	domain of definition of the operator A
$\mathcal{R}(A)$	range of the operator A, $\mathcal{R}(A) := \{v : v = Au \text{ for } u \in \mathcal{D}(A)\}$
$\mathcal{N}(A)$ or $\ker A$	null space (or kernel) of the operator A, $\mathcal{N}(A) := \{u \in \mathcal{D}(A) : Au = 0\}$
$\sigma(A)$	spectrum of A
$\mathcal{N}^{\perp}(A)$	orthogonal complement of the null space of A
$M + L$	sum of the space M and L, $M + L = \{x + y : x \in M \text{ and } y \in L\}$
$M \oplus L$	orthogonal direct sum, $M + L$, where $L = M^{\perp}$
$\mathrm{A} \otimes \mathrm{B}$	Kronecker tensor product of two matrices A and B
AB or $A \circ B$	product of the operators A and B, $(AB)(u) := A(Bu)$
Tr	Trace operator

Norms and Inner Products

$\lvert z \rvert$	absolute value of the complex number z
$\lvert \cdot \rvert$	Euclidean norm
$\lVert \cdot \rVert_{L^2(\Omega)}$	L^2 norm on Ω
(v, u)	inner product in $L^2(\Omega)$
$\langle v, u \rangle$	inner product in $L^2(\partial\Omega)$ or Euclidean inner product of two vectors
$\lVert \cdot \rVert_1$	L^1 norm
$\lVert \cdot \rVert_\infty$	L^∞ norm
$\lVert \cdot \rVert_{H^s(\Omega)}$	Sobolev norm for $s \in \mathbb{R}$; $\left\lVert (I - \Delta)^{\frac{s}{2}} \cdot \right\rVert_{L^2(\Omega)}$

Function Spaces

$\mathcal{S}$	space of rapidly decreasing functions
$\mathcal{S}'$	space of tempered distributions
C^k	space of complex, k-times continuously differentiable functions
C_0	space of complex, differentiable functions with compact support
C_0'	dual space to C_0
H^s	Sobolev space with $s \in \mathbb{R}$
L^2	space of quadratic integrable functions

Miscellaneous

$\mathcal{O}(\cdot)$, $o(\cdot)$	Landau symbols
$H_n^{(1)}$, $H_n^{(2)}$	Hankel functions of first and second kind and order n
J_n	Bessel function of order n
$\delta(\cdot)$	Dirac's delta function
$\sim$	asymptotic equivalence, $f(x) \sim g(x)$ means $f(x) = g(x)\,(1 + o\,(x))$

Contents

CHAPTER 1

Introduction

This work presents a new approach to the numerical solution of scattering problems on unbounded domains. The presentation is rather broad, so that it is hopefully of interest for an interdisciplinary community of mathematicians, physicists and engineers working on the simulation of wave propagation phenomena.

Our goal is to solve scattering problems modeled by the Helmholtz equation

$$\begin{aligned} \Delta u(\mathbf{x}) + k^2(\mathbf{x})\, u(\mathbf{x}) &= f(\mathbf{x}) && \text{in } \Omega \\ \Delta u_{\text{out}}(\mathbf{x}) + k^2(\mathbf{x})\, u(\mathbf{x})_{\text{out}} &= 0 && \text{outside } \Omega \\ u(\mathbf{x}) &= u_{\text{src}}(\mathbf{x}) + u_{\text{out}}(\mathbf{x}) && \text{on } \partial\Omega \\ \partial_n u(\mathbf{x}) &= \partial_n u_{\text{src}}(\mathbf{x}) + \partial_n u_{\text{out}}(\mathbf{x}) && \text{on } \partial\Omega \\ \partial_n u_{\text{out}}(\mathbf{x}) &= B u_{\text{out}}(\mathbf{x}) && \text{on } \partial\Omega. \end{aligned} \tag{1.0.1}$$

Here $k(\mathbf{x})$ is the position dependent wavenumber, k^2 is often also called potential, and $f(\mathbf{x})$ is a given source term with support only in some bounded computational domain $\Omega \subset \mathbb{R}^d$, where d is the space dimension. Practically we consider the cases $d = 1, 2$; the theoretical analysis is done for arbitrary space dimensions. We assume that $u(\mathbf{x})$ can be decomposed into a source field $u_{\text{src}}(\mathbf{x})$ and an outgoing part $u_{\text{out}}(\mathbf{x})$ on the boundary of the computational domain. The source field is a field excited by sources placed in the exterior of Ω and may contain both incoming and outgoing fields. The most prominent source field is a plane wave $\exp(i\mathbf{kx})$, but our approach will allow for much more general source fields. A precise definition of incoming and outgoing fields will be given in Section 3.1 (1D) and Section 3.4 (higher dimensions). The above boundary operator B is called Dirichlet-to-Neumann (DtN) operator (also Calderon operator or Poincaré-Steklov operator), because it maps Dirichlet to Neumann data on the boundary. The main part of this work will focus on the definition, construction, and analysis of discrete approximations to the boundary operator. The common way to formulate the problem, for a wavenumber k *constant* in some exterior domain, is to pose it right from the beginning on the entire, unbounded domain, and to impose the asymptotic Sommerfeld radiation condition

$$\lim_{\rho\to\infty} \rho^{\frac{d-1}{2}} (\partial_\rho u - iku) = 0, \ \rho = |\mathbf{x}|\,, \ \text{uniformely in all directions } \frac{\mathbf{x}}{|\mathbf{x}|}, \tag{1.0.2}$$

where d is the space dimension. For $k = k(\mathbf{x})$ explicitely spatially dependent, however, the Sommerfeld condition does not hold and all the methods based on it must fail. Example 1.2.2 in Section 1.2 presents a practically relevant example of a scattering problem with an inhomogenious exterior domain. Therefore our treatment is *not* based on Sommerfeld's radiation condition. We shall replace the Sommerfeld condition by a property which we call *pole condition*. This notion will be *central* in our presentation. The pole condition is applicable in a natural way to a large class of problems with position dependent wavenumbers and arbitrary space dimension. Further it offers many constructive ways to compute corresponding boundary operators B, and thus to restrict problems given on unbounded domains to such ones given on bounded domains. Moreover, the pole condition can be applied in a direct

manner to the *one-way wide angle approximations* of the Helmholtz equation which include, for example the time-dependent *Schrödinger equation*. We call the class of scattering problems consisting of the classical Helmholtz equation with possibly non-constant coefficients in the exterior domain and the one-way wide angle approximations including the time-dependent Schrödinger equation *Helmholtz-type* problems.

From the application point of view it is our goal to solve numerically scattering problems of Helmholtz-type

$$\text{with} \quad \begin{cases} \bullet \text{ inhomogeneities in the exterior domain,} \\ \bullet \text{ computatiol domains naturally defined by the problem,} \end{cases}$$

$$\text{without} \quad \begin{cases} \bullet \text{ an explicit use of Green}'\text{s functions,} \\ \bullet \text{ being confined to separable coordinates,} \\ \bullet \text{ explicit knowledge of an asymptotic radiation condition.} \end{cases}$$

The types of the exterior inhomogeneities for the Helmholtz-type problems, which can be handled by our method, follow from the construction principle and the kind of discretization of the exterior region. Most important here is the possibility to take into account waveguide-type perturbations of homogeneous media. In Section 1.2 we will present a typical example from integrated optics.

1.1. Physical Models of Wave Phenomena

The Helmholtz equation on unbounded domains plays an important role in many areas of science and engineering. Most prominent here are electromagnetics, acoustics and quantum mechanics. In quantum mechanics, the Helmholtz equation usually occurs as an eigenvalue problem related to the Schrödinger equation. Despite of the fact that the numerical solution of Helmholtz-type eigenvalue problems on unbounded domains has much in common with the solution of scattering problems on unbounded domains, we will not include this in our work, because the numerical treatment of eigenproblems is a problem of its own.

Maxwell's Equations. Let us discuss first the occurrence of Helmholtz equations in electromagnetics. The electromagnetic field in charge-free and non-conducting media is governed by Maxwell's equations

$$\begin{array}{rclcrcl} \operatorname{curl}\mathbf{H}(\mathbf{x},t) & = & \epsilon(\mathbf{x})\partial_t\mathbf{E}(\mathbf{x},t) & \qquad & \operatorname{div}\mu(\mathbf{x})\mathbf{H}(\mathbf{x},t) & = & 0 \\ \operatorname{curl}\mathbf{E}(\mathbf{x},t) & = & -\mu(\mathbf{x})\partial_t\mathbf{H}(\mathbf{x},t) & & \operatorname{div}\epsilon(\mathbf{x})\mathbf{E}(\mathbf{x},t) & = & 0 \end{array}$$

where $\mathbf{H}(\mathbf{x},t)$ and $\mathbf{E}(\mathbf{x},t)$ denote the magnetic and electric field strength, respectively, in space and time. The material properties are described by the electric permittivity $\epsilon(\mathbf{x})$ and the magnetic permeability $\mu(\mathbf{x})$. If we eliminate either $\mathbf{H}$ or $\mathbf{E}$ from one of the curl-equations, we obtain a vectorial wave equation. E.g. the elimination of $\mathbf{E}$ from the first curl-equation results in

$$\operatorname{curl}\frac{1}{\epsilon(\mathbf{x})}\operatorname{curl}\mathbf{H}(\mathbf{x},t) = -\mu(\mathbf{x})\partial_t^2\mathbf{H}(\mathbf{x},t).$$

One of the oldest and most successful approximations of the vectorial wave equation is obtained, when we regard the factor $1/\epsilon(\mathbf{x})$ as a constant with respect to the curl-operation. As long as the permittivity changes only weakly over a distance of a local wavelength, the approximation supplies an accepted physical model of optical wave propagation. It is the most essential approximation, which leads us from Maxwell's equations to geometrical optics, see **[13**, chapt. 3.1, pp. 111] and scalar wave optics. Following this approximation, applying the identity

$\operatorname{curl}\operatorname{curl}\mathbf{H} = \operatorname{grad}\operatorname{div}\mathbf{H} - \Delta\mathbf{H}$ and taking the divergence condition into account, we obtain the scalar wave equation

$$\Delta\mathbf{H} = \epsilon(\mathbf{x})\mu(\mathbf{x})\,\partial_t^2\mathbf{H}$$

where the Laplacian is applied to each individual component of $\mathbf{H}$ separately. Since the vectorial components are independent of each other, the equation separates into three independent scalar equations. Let us denote one of the vectorial components of $\mathbf{H}(\mathbf{x},t)$ by $u(\mathbf{x},t)$. Further let us assume that field $u(\mathbf{x},t)$ is harmonic in time, i.e. $u(\mathbf{x},t) = \Re\left(\widetilde{u}(\mathbf{x})\exp(-i\omega t)\right)$, where ω is the angular frequency of the wave. Set $k^2(\mathbf{x}) := \epsilon(\mathbf{x})\mu(\mathbf{x})$. If we introduce this complex representation into the wave equation and drop the tilde, we arrive at the scalar Helmholtz equation

$$\Delta u(\mathbf{x}) + k^2(\mathbf{x})u(\mathbf{x}) = 0.$$

Let us consider the case, where the wavenumber k is independent of the position, that is the wave propagation in homogeneous media. Then all plane waves $u(\mathbf{x}) = \text{const}\exp(i\mathbf{kx})$, with a wavevector $\mathbf{k}$ which satisfies the dispersion relation $\mathbf{k}\cdot\mathbf{k} = k^2$, are solutions of the Helmholtz equation. These plane waves correspond to time-harmonic solutions $u(\mathbf{x},t) = \text{const}\operatorname{Re}\exp\left(i\left(\mathbf{k}\cdot\mathbf{x} - \omega t\right)\right)$. We consider a fixed direction $\mathbf{x}_0$ and study the evolution of the phase $\arg\exp\left(i\left(\mathbf{k}\cdot\mathbf{x}_0 - \omega t\right)\right)$ with increasing time. Clearly, we can identify the case $\operatorname{Re}\left(\mathbf{k}\cdot\mathbf{x}_0\right) > 0$ with a plane wave, which propagates in the positive direction with respect to $\mathbf{x}_0$ and the opposite case $\operatorname{Re}\left(\mathbf{k}\cdot\mathbf{x}_0\right) < 0$ with a plane wave, which propagates in the negative direction with respect to $\mathbf{x}_0$. Note that a definition of the time-dependency based on $\exp(i\omega t)$ would cause us to define the reverse directions as positive and negative propagation direction.

Acoustics. We consider acoustic wave propagation in a compressible, ideal fluid. This involves the interaction of the following three physical effects [**31**, Vol. 1, Chapt. 47]:

(1) The fluid moves and changes its density $\rho(\mathbf{x},t)$. The corresponding physical equation is the continuity equation $\partial_t\rho + \operatorname{div}(\rho\mathbf{v}) = 0$.
(2) The change in the density results in a change of the pressure $p(\mathbf{x},t)$. This is described in linear approximation by the equation $p(\mathbf{x},t) = p_0 + c^2\left(\rho(\mathbf{x},t) - \rho_0\right)$, where ρ_0 and p_0 characterize the constant state of the fluid with all particles at rest. The constant factor c is the velocity of sound.
(3) Local differences in the pressure result in a change of the velocity $\mathbf{v}(\mathrm{x},t)$. Since $\mathrm{d}\left(\rho\mathbf{v}\right) = \mathrm{d}t\,\partial_t\left(\rho\mathbf{v}\right) + \left(\mathrm{d}\mathbf{x}\cdot\nabla\right)\left(\rho\mathbf{v}\right)$ it follows $\mathrm{d_t}\left(\rho\mathbf{v}\right) = \partial_t\left(\rho\mathbf{v}\right) + \left(\mathbf{v}\cdot\nabla\right)\left(\rho\mathbf{v}\right)$. Newton's law supplies the equation of motion $\partial_t\left(\rho\mathbf{v}\right) + \left(\mathbf{v}\cdot\nabla\right)\left(\rho\mathbf{v}\right) = -\operatorname{grad}p$. We consider only problems, where the spatial variation of $\rho\mathbf{v}$ is much smaller than its variation in time. Hence we can linearize the equation of motion to obtain Euler's equation $\partial_t\left(\rho\mathbf{v}\right) = -\operatorname{grad}p$.

Euler's law, together with the continuity condition, yields $\partial_t^2\rho = \operatorname{div}\operatorname{grad}p$. Using the linear relation between density and pressure, we obtain the wave equation for the pressure

$$\Delta p(\mathbf{x},t) = \frac{1}{c^2}\partial_t^2 p(\mathbf{x},t).$$

Proceeding as in Section 1.1, we restrict the equation to time-harmonic solutions and introduce the complex pressure $\widetilde{p}(\mathbf{x})$ by $\operatorname{Re}p(\mathbf{x},t) = \left(\widetilde{p}(\mathbf{x})\exp(-i\omega t)\right)$. We insert this expression into the wave equation, drop the tilde and obtain the Helmholtz equation for the pressure

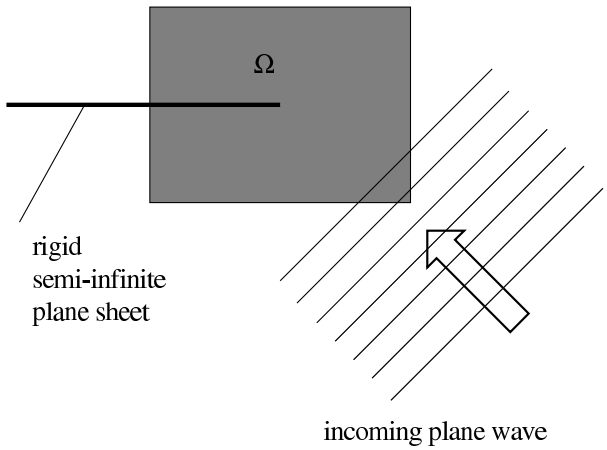


FIGURE 1.2.1. Reflection and diffraction by a semi-infinite plane sheet (Example 1.2.1)

$$\Delta p(\mathbf{x}) + k^2(\mathbf{x})p(\mathbf{x}) = 0 \quad \text{where } k^2(\mathbf{x}) = \omega^2/\,c^2.$$

Quantum Mechanics and Paraxial Beam Propagation. The non-relativistic motion of an electron of mass m is governed by Schrödinger's equation

$$i\hbar\partial_t\psi(\mathbf{x},t) = -\frac{\hbar^2}{2m}\Delta\psi(\mathbf{x},t) + V(\mathbf{x},t)\psi(\mathbf{x},t)$$

where $\psi(\mathbf{x},t)$ is the quantum mechanic wave function, $\hbar = h/(2\pi)$, with h the Planck quantum action, and $V(\mathbf{x},t)$ a potential. The lowest order wide angle one-way approximation of Helmholtz's equation is of the same type. In optics it takes the form

$$-i\partial_z u(\mathbf{x},z) = \frac{1}{2n_0k_0}\left(\Delta u(\mathbf{x},z) + k_0^2\left(n^2 - n_0^2\right)u(\mathbf{x},z)\right)$$

and is called Fresnel's equation or the paraxial wave equation. The z-axis is the main propagation direction of the beam, the transversal coordinates are $\mathbf{x} = (x,y)^T$, k_0 denotes the free-space wave number, $n(\mathbf{x},z)$ is the refractive index and n_0 is the so-called reference refractive index. The latter constant is a model parameter and must be chosen such that the deviation with respect to the Helmholtz equation becomes as small as possible, see e. g. **[84]**. The function $u(\mathbf{x},z)$ is the slowly-varying amplitude function and is defined as $u = \exp(-in_0k_0z)E$, where E stands for a component of either the electric or the magnetic field.

1.2. Examples of Coupled Interior-Exterior Helmholtz-Type Problems

In the following we want to characterize the problem classes under consideration by typical examples from integrated optics and acoustics. Each of the examples discussed possesses at least one typical difficulty known to be a challenge to its numerical solution.

EXAMPLE 1.2.1. *Reflection and diffraction by a semi-infinite plane sheet.* Fig. 1.2.1 shows schematically an arrangement, where an incoming plane wave strikes a semi-infinite plane sheet. In the acoustic case, the solution $u(x,y)$ of the Helmholtz equation (1.0.1) is the deviation from the hydrostatic pressure, the

gradient of whose is proportional to the velocity of the particles, see Section 1.1. If the semi-infinite plane sheet is a rigid boundary, there cannot be a normal component of the velocity across it. Therefore the boundary condition is $\partial_n u = 0$. The plane wave itself does not satisfy this boundary condition. Therefore the solution $u(x, y)$ must be a superposition of the incoming plane wave and a scattered field, composed of reflected and diffracted waves. A typical scattering problem of this type is: Given the incidence angle of the incoming plane wave, compute the field in the bounded computational domain Ω. This problem belongs to the most simple ones, and we will discuss the numerical solution in Section 5.1.2. Because the whole problem can be formulated based on our one-dimensional concept, it provides an interesting and instructive application of the one-dimensional theory. Further, since an analytic solution based on the famous Wiener-Hopf technique can be derived, it allows a step-by-step comparison between our numerical method and an integral equation approach. It will turn out that the analytic solution process consists of the following main steps:

Step 1: First, the entire field is decomposed into the incoming field, the reflected field and the scattered field, where this decomposition is motivated by geometrical optics and needs a-priori knowledge about the solution. While we take the geometrical optics ansatz for the reflected field, the scattered field has to satisfy the Helmholtz equation and Sommerfeld's asymptotic radiation condition. Note that the reflected field *does not* satisfy the radiation condition.

Step 2: Apply Green's identity to the decomposed field and derive a Fredholm integral equation of the first kind with a special structure, known as Wiener-Hopf integral equation.

Step 3: On the other hand, the density function inside the integral is partially known. This problem can be solved by the Wiener-Hopf technique, which requires, however, an additional trick. To apply the Wiener-Hopf technique, we must extend the real wavenumber k to a new, complex wavenumber $\widetilde{k} = k + ik_2$ with a small imaginary part k_2.

In contrast to the integral equation approach, the discrete method based on transparent boundary conditions needs neither a decomposition of the field other than the decomposition into incoming and outgoing waves nor an extension of the real wavenumber to a complex number. The numerical problem here is to take into account a boundary, which extends from a finite point to infinity. This however is a problem for the numerical evaluation of the analytical integral equation too, because all integrals extend to infinity.

EXAMPLE 1.2.2. *Scattering by a finite obstacle.* This problem concerns the computation of the scattered field, exited by an incoming plane wave, in the vicinity of an obstacle with a complicated boundary. The situation, schematically displayed in Fig. 1.2.2, is similar to the preceding one, except that the boundary, which causes the scattering, can be included into a bounded computational domain Ω. This simplification allows the application of other methods to compute the scattered field in or outside Ω. Especially the choice of separable coordinates in the exterior of Ω, if possible, leads to a considerable simplification of the numerical methods.

EXAMPLE 1.2.3. *Inhomogeneous exterior domain.* This example problem, see Fig. 1.2.3, presents a problem with a typical inhomogeneity in the exterior domain. The situation is: The left waveguide, which comes from infinity and ends inside the computational domain Ω, guides a mode from infinity to the right end of the waveguide, where the field leaves the waveguide. Let us assume that the wavenumber of

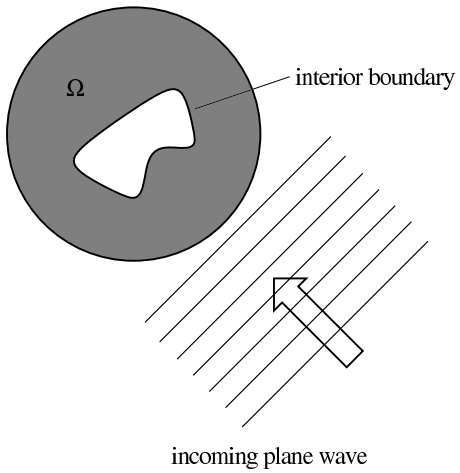


FIGURE 1.2.2. Scattering of a plane wave by a finite obstacle (Example 1.2.2)

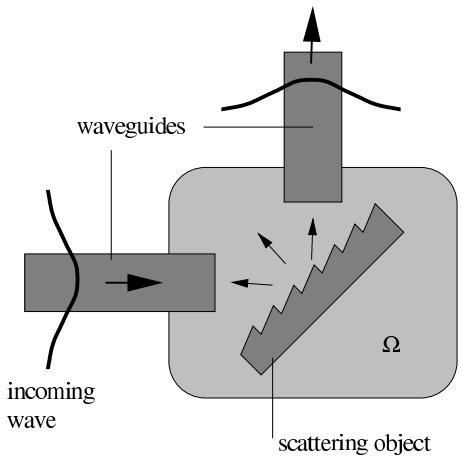


FIGURE 1.2.3. Scattering problem with an inhomogeneous exterior domain. The wavenumber corresponding to the waveguide domains (dark gray) is different from the wavenumber of the surrounding media (Example 1.2.3)

the waveguide is k_1, the wavenumber of the surrounding homogeneous medium is k_0, and that $k_1 > k_0$. Then the waveguide-mode travels without attenuation from infinity to the end-face of the waveguide with a phase factor $\exp(i\beta x)$, where x denotes the axis of the fiber, and $k_0 < \beta < k_1$. Hence the traveling mode *does not* satisfy Sommerfeld's radiation condition. The field radiated by the waveguide hits a scattering object. Schematically we used a wavelength selective mirror, where the field is scattered in all directions. Parts of the incoming field are scattered back into the waveguide, other parts are scattered into a second waveguide, and the remainder is scattered into the homogeneous part of the exterior domain. The part of the wave which is scattered into the second waveguide, can excite a guided mode traveling undamped towards infinity. Again, this part does not satisfy the Sommerfeld condition. We call the described inhomogeneities of the exterior domain waveguide-type inhomogeneities.

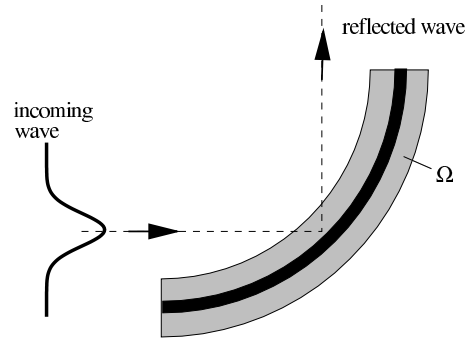


FIGURE 1.2.4. Reflection by a parabolic mirror. The computational domain is non-convex (Example 1.2.4)

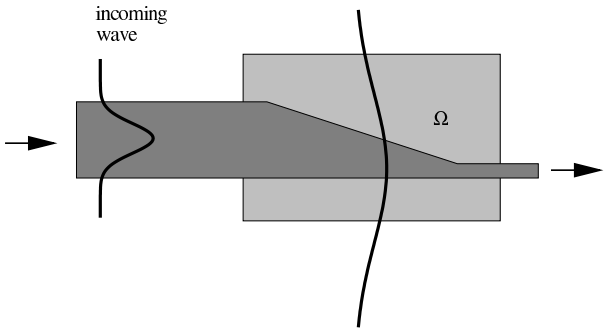


FIGURE 1.2.5. Beam propagation through a tapered waveguide (Example 1.2.5)

So far, all computational domains convex, and, regarding the physical situation, this has been the natural choice. Next, we give a problem type, where a non-convex computational domain possesses numerical advantages.

EXAMPLE 1.2.4. *Parabolic mirror.* Fig. 1.2.4 shows a typical reflection problem, where the scattering surface is of a parabolic type, embedded in a non-convex computational domain. As far as only the near-field solution of the reflection problem is needed, domains of this type might have some numerical advantages compared to domains based on the convex hull of the scattering surface, because they are much smaller than an appropriate convex domains. Thus, the numerical effort to discretize the computational domain is potentially less than in the convex case. Using one version of our Laplace domain method, the cut-function approach, we are even able to compute the far-field without additional effort.

EXAMPLE 1.2.5. *Paraxial beam propagation.* This example is concerned with the one-way wide angle approximation of the Helmholtz equation. A typical situation is shown in Fig. 1.2.5. A waveguide guides a mode form infinity to the computational domain, where the waveguide changes its shape. This change causes in general both a scattering and a deformation of the shape of the guided part of the wave. We indicated the latter schematically by a larger width of the amplitude function at the thinner end of the waveguide. Thus we have a situation similar to the previous ones, except that in this case the direction of the beam propagation is maintained which allows us to apply the one-way equation. Obviously, we could

solve this problem by the same means we will use to solve the reflection problem of Fig. 1.2.3. However, it will turn out that at the present state of the numerical methods it is much more effective to use one of the approximations of the one-way Helmholtz equation.

1.3. What Are Transparent Boundary Conditions?

Both our new method – the pole condition approach – as well as the classical methods for the solution of wave propagation problems on unbounded domains can be seen as methods to construct *transparent boundary conditions*. To prepare the very first presentation of our pole condition concept in the following section, we want to introduce this term here. A more detailed description of the classical methods follows in Section 2.

Many methods to solve the discussed wave propagation and scattering problems rely on a decomposition of the unbounded domain, on which the problem is posed, into a bounded – possibly large – interior domain Ω_{int}, which contains the scattering objects, and an exterior domain. Now the following question which quote from Grote and Keller **[42]** arises: "Does there exist a boundary condition such that the solution of the problem in Ω_{int}, with this boundary condition, coincides exactly with the restriction to Ω_{int} of the solution in the unbounded domain?" Such boundary conditions are called exact or *transparent* boundary conditions, sometimes also radiating or absorbing boundary conditions. There exist a large number of applications, where one is interested in the solution of the scattering problem only in the vicinity of the scatterer, i. e.in the solution on the interior domain. In these cases, the transparent boundary conditions plus a consistent and stable discretization of the interior problem provide the basis of a numerical solution. It is the goal of this work to supply a general approach to construct such transparent boundary conditions.

The above discussion suggests to decompose the problem into two sub-problems: (1) The derivation of the transparent boundary conditions by some method, and, independently of it, (2) the solution of the interior problem by another method. In fact, this is an approach followed successfully by many researchers, cf. e g. the work of Keller, Givoli, and Grote **[64, 38]** for the Laplace and the Helmholtz equation, **[44, 45]** for the wave and time-dependent Maxwell equations. Nevertheless, this strategy has the disadvantage that an analytic solution representation in the exterior domain *must be known in advance*. The same holds true for the already mentioned class of boundary integral methods. Our approach is quite different. Within our methods the transparent boundary conditions and the solution of the interior problem are computed *simultaneously*.

There is another important aspect. Usually the transparent boundary conditions are derived from an exact analytic representation formula. Then they are discretized and applied to the discretized interior problem. The discretization violates the transparency of the boundary which becomes transparent only in the limiting case of a vanishing discretization step width. The pole condition approach, however, which is first given for the continuous case too, has a natural counterpart in the discrete case. Naturally enough, we require for a given discrete scheme that the pole condition has to be satisfied, thus realizing the transparency even for the discrete case. Therefore we call our boundary conditions sometimes also *discrete transparent boundary conditions*.

The next section supplies a first impression of our concept.

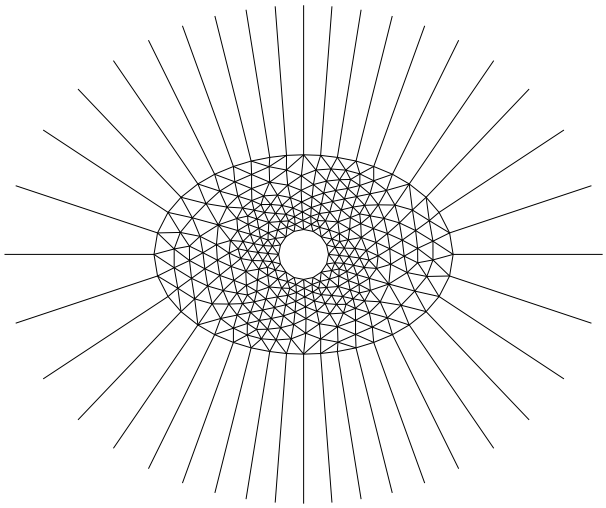


FIGURE 1.4.1. Inner discretization by triangulation and outer discretization by rays used by the methods developed in this work

1.4. Basic Idea of the Pole Condition Approach

We refer to Fig. 1.4.1. The interior domain, which is bounded by an ellipse, is discretized by a triangulation. Any other discretization technique might be applied as well. However, the numerical implementation of the algorithms, which we shall develop, is always based on finite elements and triangulations. This results both in a high flexibility with respect to the discretization of scatterers with complicated shapes and a very natural and simple coupling to the exterior domain. The exterior domain itself is discretized by linear rays with initial points on the artificial boundary. The exterior discretization is chosen such that each node on the artificial boundary is connected with infinity such that the rays do not intersect each other. The transparent boundary condition is realized automatically on the initial points of the rays, but not independently of the interior solution. We compute simultaneously

(1) A discrete interior solution on the triangulation.
(2) A discrete solution on the rays.

However, the solution on the rays is not obtained in the given spatial form, instead its Laplace transform is computed. From all possible exterior solutions we accept only those as *outgoing* solutions, whose Laplace transform has no singularity in the lower half of the complex plane. We call this selection rule *pole condition.* The restriction to outgoing solutions supplies the desired transparent boundary conditions.

We can carry out the concept without any analytic pre-treatment. The Laplace transform is computed numerically by means of an Volterra integral equation. This is done very effectively by different types of collocation methods, among them explicit and implicit Runge-Kutta methods and a classic spline approximation method. Note, that no *inverse* Laplace transform is needed. The restriction to linear rays lies not in the nature of the method, however, it supplies the easiest treatment of the exterior discretization.

Our main tool will be the Laplace transform of the fields outside of Ω. In this uniform framework the following three principles are of fundamental importance:

(1) The singularities of the Laplace transform of a complex function defined on a semi-infinite interval determine its far-field. In short: The near-field in the Laplace domain determines the far-field in the physical domain.
(2) The near field of this function is determined by its Laplace transform for large values of the Laplace variable. In short: The near-field in the physical domain is determined by the far-field in the Laplace domain.
(3) Problems in higher space dimensions can be composed from mutually interacting one-dimensional problems.

1.5. Outline of Contents

The structure of the individual chapters and their mutual relation is the following: In Chapter 2 we give a survey of important classical methods used to derive transparent boundary conditions.

In Chapter 3 we introduce the pole condition. The point of departure is a one-dimensional theory, whose basic point is a constructive definition of incoming and outgoing waves. We develop a tool, which enables us to decompose complex functions given on a semi-infinite interval into incoming and outgoing parts, and we give precise definitions of the notions "incoming" and "outgoing". We apply the one-dimensional theory to several illustrating and non-trivial examples, including e. g. periodic potentials. Next we generalize the concept to higher space dimensions. It turns out that this can be done completely in terms of the one-dimensional decomposition of waves. Finally we discuss the extension of the one-dimensional concept to the time-dependent Schrödinger equation.

In Chapter 4 we develop the general existence and uniqueness theory for a model problem: the radially symmetric Helmholtz equation. The main results are:

- The pole condition is equivalent to Sommerfeld's asymptotic radiation condition, at least in the considered finite-dimensional setting.
- The pole condition generalizes Sommerfeld's radiation condition for cases with potentials depending on the radial distance.
- A new representation formula for the exterior solutions is obtained.
- New asymptotic series representations for the exterior solutions are derived.

In Chapter 5we derive both the principal structure as well as the essential details of our algorithms to solve scattering problems based on the Helmholtz equation. The emphasis here lies on the development of a new class of algorithms, which we call *Laplace domain methods* since they approximate the essential quantities in the spectral (Laplace) domain. To the best of our knowledge, these Laplace domain methods do not have a predecessor in the numerical literature. They follow naturally from our theory. The algorithms can be divided into two sub-classes: The real axis integral approach and the cut-function approach. We study in detail the numerical properties of these algorithms based on our basic model equation, the Bessel equation. We demonstrate that accuracy in the order of 10^{-10} can be obtained even close to singular cases with a very moderate numerical effort, cf. Fig.'s 5.2.3 and 5.2.7. These studies are performed based on the 1D Bessel model, since corresponding studies in higher dimensions would cause an additional preparation effort. Along with the most general algorithmic realization we discuss briefly the asymptotic approximation of the derived general formulation. We point out its direct correlation to known analytic approximations.

In Chapter 6 we derive the algorithms for the time-dependent Schrödinger equation and the wide angle one-way equations. Since our focus here is on the time-evolution, we consider only problems with one transversal (spatial) direction. First we study the Schrödinger equation and derive the basic technique which allows

us to extend our method to time-discrete problems. In Section 6.2 we generalize this concept to the wide angle one-way approximations of the Helmholtz equation. Finally, we give some typical applications of Schrödinger-type scattering problems along with a study of their numerical properties.

CHAPTER 2

A Survey of Different Classical Methods

A number of quite different methods are available to solve scattering problems of the described types. It seems to be impossible to give a single satisfactory classification scheme which sets all the methods in relation to each other, with all their pros and cons. However, most important here is to discuss the relation between the discrete method developed in this paper and the most successful classical methods. Here use the term "classical" even for relatively new methods such as the the infinite-element method. Fig. 2.0.1 attempts to show the relation between the different methods postulating that a continuous form of a radiation condition provides the starting point for all methods. According to this classification the common strategy of all methods is to construct a representation scheme for the exterior solutions such that all solutions both satisfy the continuous Helmholtz equation and the continuous radiation condition. Here the radiation condition is of fundamental importance. It is constructed such that it allows only for "outgoing" or "incoming" solutions.

2.1. Factorization Based on Separable Coordinates

For simplicity, we consider only the solution of the Helmholtz equation exterior to a circle with radius a. To get the point, we leave out some details needed to justify the operations. Every function $u \in L^2(\partial\Omega)$, $\partial\Omega = \{x \in \mathbb{R}^2 : |x| = a\}$ can be expressed via its Fourier series expansion

$$u(a,\phi) = \sum_{\nu=-\infty}^{\infty} a_\nu e^{i\nu\phi}, \quad a_\nu = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i\nu\phi} u(a,\phi)\, d\phi\,.$$

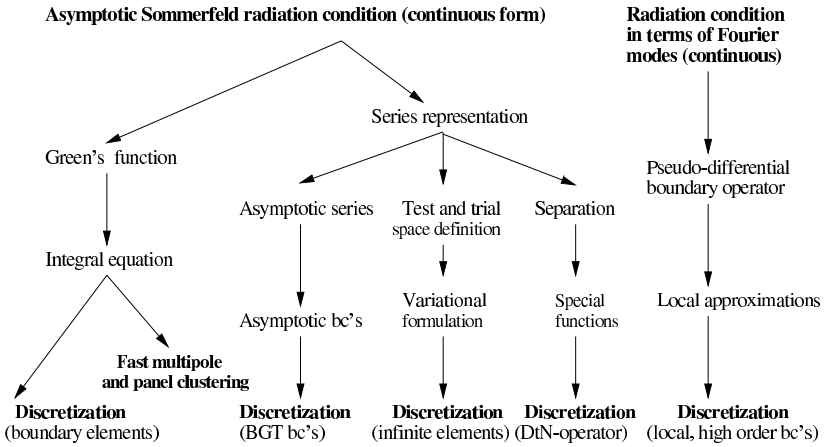


FIGURE 2.0.1. Classical approaches to derive transparent boundary conditions

A separation ansatz in cylindric coordinates $u(r,\phi) = \sum_\nu a_\nu e^{i\nu\phi} u_\nu(r)$ shows that $u_\nu(r)$ must obey Bessel's differential equation

$$\partial_r^2 u_\nu + \frac{1}{r}\partial_r u_\nu + \left(k^2 - \frac{\nu^2}{r^2}\right) u_\nu = 0\,, \quad r > a.$$

The fundamental solution and its derivative are

$$\begin{aligned} u_\nu(r) &= c_1 H_\nu^{(1)}(kr) + c_2 H_\nu^{(2)}(kr) \\ u_\nu'(r) &= c_1 H_\nu^{(1)\prime}(kr) + c_2 H_\nu^{(2)\prime}(kr) \end{aligned}$$

where $H_\nu^{(1)}, H_\nu^{(2)}$ are Hankel's functions of the first and the second kind, respectively, and the prime denotes the derivative with respect to the argument. From an asymptotic study of Hankel's functions we know that only the first kind obeys the Sommerfeld condition. In order to drop the Hankel function of the second kind, we must establish an additional condition, namely,

$$u_\nu'(r) = k\frac{H_\nu^{(1)\prime}(kr)}{H_\nu^{(1)}(kr)} u_\nu(r),$$

or, equivalently,

$$u_\nu'(r) = \left(\frac{\nu}{r} - k\frac{H_{\nu+1}^{(1)}(kr)}{H_\nu^{(1)}(kr)}\right) u_\nu(r).$$

Taking the angular-dependent part into account, we obtain the DtN-map

$$\partial_n u(r,\phi)|_{r=a} = \sum_{\nu=-\infty}^{\infty} \left(\frac{\nu}{r} - k\frac{H_{\nu+1}^{(1)}(kr)}{H_\nu^{(1)}(kr)}\right)_{r=a} \frac{e^{i\nu\phi}}{2\pi} \left\langle e^{i\nu\phi}, u(a,\phi)\right\rangle_{L^2(\partial\Omega)}.$$

The main ingredients, with respect to this classical approach, have been the expression of the fundamental solution in terms of a superposition of Hankel's functions and the decision, which of them satisfy Sommerfeld's radiation condition. For example, if we would start with another pair of fundamental solutions, say with Bessel- and Weber functions, none of them would satisfy the Sommerfeld condition and we must try to find a superposition of both, namely the Hankel functions, which obey the radiation condition.

This technique has been introduced by Givoli and Keller for the Laplace equation, the Helmholtz equation and the elastic wave equation in in a series of papers, cf. e. g. **[38, 37, 35, 64]**, and has been extended later by Givoli, Grote and Keller to time-dependent problems (e. g. wave equation, Maxwell's equations), see the corresponding articles **[36, 43, 42, 44, 45]**.

Some of the features of the method are:

- The DtN-number $u_\nu'(a)/u_\nu(a)$ corresponding to a given mode number ν and a distance a is derived in straightforward way.
- In standard cases like the one above the implementation does not cause a problem, since the necessary special functions are available from standard libraries.
- Efficient tools to couple the non-local Fourier modes to local interior discretizations are available.

The obvious drawbacks are:

- The problem must have a separable structure. In general, this is not the case for variable coefficient problems.
- The generalization to other then cylindric and elliptic coordinate systems and their 3D generalizations is difficult.

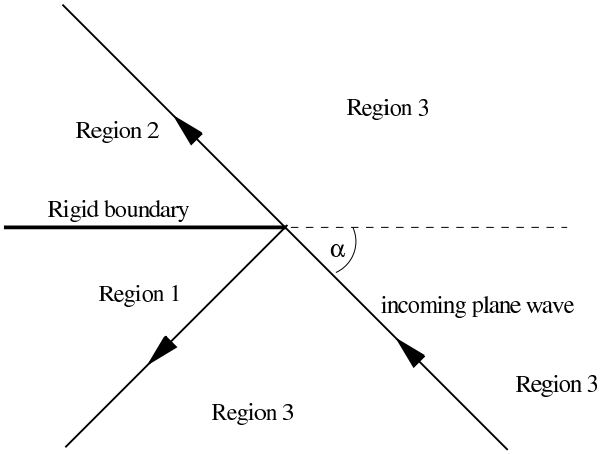


FIGURE 2.2.1. Diffraction problem

2.2. Boundary Integral/Element Methods

We describe the concept of boundary integral methods (boundary element methods) by means of a particular example: the above mentioned classical diffraction problem of a plane wave by a rigid semi-infinite plane sheet, Example 1.2.1, p. 12. A representation of boundary integral methods for the Helmholtz equation for problems with bounded scatterers may be found in **[3]**, for the coupling between finite elements and boundary elements see, e. g. **[2]**. Our presentation follows the ones given in **[24**, Sec. 8.7] and **[23**, Sec. 5.6, Chap. 8]. An alternative operator theoretical approach of the same subject can be found in **[73]**.

We decompose the whole procedure into the tree main steps discussed briefly in connection with Example 1.2.1:

(1) Decomposition of the full field into three parts: The incoming plane wave $u_{\text{in}} = \exp(i\mathbf{kx})$, the reflected wave u_{refl} and the diffracted part $u_{\text{diff.}}$
(2) Solution representation by a properly constructed Green's function.
(3) Solution of the integral equation.

Without going too much into the details, we describe each of the steps.

Field decomposition. The situation is illustrated in Fig. 2.2.1. The incoming plane wave strikes the semi-infinite plane. The incoming field already satisfies the Helmholtz equation. It excites a reflected field, which is assumed to live mainly in Region 1, and a diffracted field, which lives in all the regions 1,2, and 3. The entire field, the incoming plus the fields excited by the rigid boundary (the scattered field), must satisfy the Helmholtz equation. The problem to formulate a Green's function representation of the scattered field now is that it obviously does not meet Sommerfeld's radiation condition (1.0.2), since the reflected field satisfies it only in one but not in all directions. In order to circumvent this difficulty, that is to extract the part of the field which in fact satisfies the Sommerfeld condition, one considers the solutions predicted by geometrical optics first, see Fig. 2.2.1:

$$\begin{array}{lcl} \text{Region 1} : u & = & u_{\text{in}} + u_{\text{ref}} \\ \text{Region 2} : u & = & 0 \\ \text{Region 3} : u & = & u_{\text{in}}. \end{array}$$

From geometrical optics we know the explicit formula for the reflected field. If the incoming plane wave is $u_{\text{in}} = \exp(ik_x x + ik_y y)$ then $u_{\text{ref}} = \exp(ik_x x - ik_y y)$ in Region 1. We use these fields as initial guess of the solution of the Helmholtz

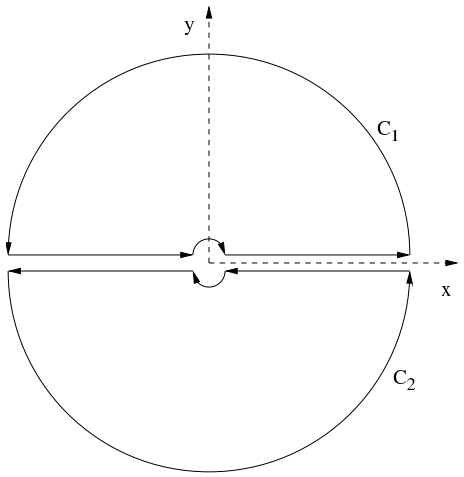


FIGURE 2.2.2. Semi-circles used for Green's identity

equation, which is corrected by the diffracted part u_{diff} according to

$$\begin{array}{lcll} \text{Region 1} & : & u & = & u_{\text{in}} + u_{\text{ref}} + u_{\text{diff}} \\ \text{Region 2} & : & u & = & u_{\text{diff}} \\ \text{Region 3} & : & u & = & u_{\text{in}} + u_{\text{diff}}. \end{array}$$

Now, we require that u_{diff} satisfies Sommerfeld's radiation condition. Additionally we require $\partial_y u = 0$ on the rigid boundary, since the normal derivative of the acoustic velocity potential supplies the normal velocity of the particles, which must be zero on a sound-hard boundary. If we compare this decomposition step with the corresponding discrete solution based on the pole condition, which we will present in 5.1.2, pp. 102, we find that no comparable procedure is necessary. The only point, where in our approach some a-priori knowledge about the solution behavior plays a role, is in the definition of the fields along artificial vertical boundaries far away from the origin. Here, we will use again informations from geometrical optics.

Integral representation of the solution. We construct a suitable Green's function from the reflection principle, based on Green's function for the free-space problem:

$$\begin{aligned} G(\xi,\eta,x,y) &= \frac{i}{4}H_0^{(1)}\left(k\sqrt{(x-\xi)^2+(y-\eta)^2}\right) \\ &\quad + \frac{i}{4}H_0^{(1)}\left(k\sqrt{(x-\xi)^2+(y+\eta)^2}\right). \end{aligned}$$

This Green's function satisfies the Helmholtz equation except at $(x,y) = (\xi, \pm\eta)$, due to the Hankel functions Sommerfeld's radiation condition and $\partial_y G(\xi,\eta,x,0) = 0$, which is the imposed boundary condition for $x \leq 0$. Next, we apply Green's identity to the whole field in the upper half of the plane $u^{(1)}$ using the contour C_1, see Fig. 2.2.2. This yields

$$\begin{aligned} u^{(1)}(\xi,\eta) &= \oint_{c_1} ds \left(G\frac{du^{(1)}}{dn} - u^{(1)}\frac{dG}{dn}\right) \\ &= -\int_{-\infty}^{\infty} dx \; G\frac{du^{(1)}}{dy}\bigg|_{y=+0} \end{aligned}$$

The latter equation follows because the integrals on the small and the large semi-circles tend to zero if their radii go to zero and infinity, respectively, see **[23**, Sec. 5.6, pp. 264-265]. Next, we derive a second integral equation, applying Green's identity to the whole field in the lower half of the plane minus the incoming plane wave and the reflected wave. We obtain

$$u^{(2)}(\xi,\eta) - u_{\text{in}}(\xi,\eta) - u_{\text{refl}}(\xi,\eta) \quad = \quad \int_{-\infty}^{\infty} dx \; G \frac{du^{(2)}}{dy}\bigg|_{y=-0}.$$

We let η approach ± 0, take $u^{(1)}(\xi,0) = u^{(2)}(\xi,0)$ for $\xi > 0$ into account, and take the difference between the two integral equations into account to obtain

$$f(\xi) = i \int_{-\infty}^{\infty} dx \, H_0^{(1)}\left(k\,|x-\xi|\right) \partial_y u(x,0),$$

where

$$f(\xi) = \begin{cases} -2e^{ik_x\xi}, & \xi > 0 \\ u\,(\xi,-0) - u\,(\xi,+0) - 2e^{ik_x\xi}, & \xi < 0. \end{cases}$$

Integral equations of this type are called Wiener-Hopf integral equations. Their difficulty lies in the fact that the function f is not completely known (namely for $\xi < 0$). However, the function $\partial_y u(\xi,0)$ under the integral sign is already known in parts (namely for $\xi > 0$).

If we compare this step with our alternative approach from Section 5.1.2, we find that we obtain an algebraic system of exactly the same structure, a linear system of the type Ax = b, where the right-hand side b is not completely known, whereas x is already partially known. However, we obtain this system without applying the techniques described above.

Solution of the integral equation. There are several ways to solve the Wiener-Hopf equation. The first one is to reformulate this integral equation by analytical means such that a classical Fredholm integral equation of the first kind results. The technique to do this is known as Wiener-Hopf technique. It is based on a Fourier transform of the convolution integral and certain factorizations of the Fourier-transformed kernel and the known part of the right-hand side of the integral equation. Applying analytic continuation techniques and Liouville's theorem supplies the desired representation. However, finally we are left with an integral equation, which must be solved, preferably by numerical methods. Therefore we can, as a second way to solve the problem, directly apply numerical discretization techniques to the Wiener-Hopf equation.

To formalize this approach, we restrict the infinite interval, on which the integral equation is defined to the finite interval $[-a,a]$ and set the approximate boundary conditions known from the geometrical optics model to $u^{(1)}(-a) = 0$, $u^{(1)}(a) = u^{(2)}(a) = u_{\text{in}}(a)$, $u^{(2)}(-a) = u_{\text{in}}(-a) + u_{\text{refl}}(-a)$. Setting $v(x) := \partial_y u(x,0)$ we have to solve the linear Fredholm integral equation

$$\int_{-a}^{a} dx \, K(x,\xi) v(x) = f(\xi),$$

where the kernel $K(x,\xi) = H_0^1\left(k\,|x-\xi|\right)$ is in $L^2[-a,a]^2$ due to the logarithmic singularity $H_0^1\left(k\,|x-\xi|\right) \to \text{const} \ln|x-\xi|$ as $(x-\xi) \to 0$. The integral operator is given by the formula

$$(\mathcal{K}u)(\xi) := \int_{-a}^{a} dx \, K(x,\xi) u(x)$$

for all $u \in L^2[-a,a]$ and all $\xi \in [-a,a]$. The Nyström, or quadrature, method replaces this integral operator by a finite-dimensional approximation

$$(\mathcal{K}_h u)(\xi) := \sum_{j=1}^{n} w_j K(x_j,\xi) u(x_j)$$

for all $u \in L^2[-a,a]$ and all $\xi \in [-a,a]$ and given points $x_1,\ldots,x_n \in [-a,a]$. Ideally, the real weights $w_1,\ldots,w_n$ are chosen such that integration along the desired solution is approximated very well.

We require that the approximation be exact at all the points x_j (collocation method), which supplies the algebraic system

$$(\mathcal{K}_h u)(\xi_j) = f(\xi_j), \quad j = 1,\ldots,n.$$

If we again compare this last step with with the formulation of our discrete method, we find a perfect coincidence with respect to the structure of the algebraic system. However, the derivation is quite different. Instead of using a discretization of an integral equation, we will use a direct discretization of the Helmholtz operator to derive the discrete system.

REMARK. In general, integral equations of the first kind are ill-posed problems. To solve such problems numerically, regularization techniques has to be applied. Since this has no effect on our elaboration in the sequel, we will not discuss this point further. The special Wiener-Hopf structure is not characteristic for the class of integral equation methods. It is a property which results from our diffraction problem. On the other hand, the main steps of the above procedure are characteristic: (1) The construction of a representation formula, (2) The restriction of the representation formula to the boundary of an obstacle or to an artificial boundary, (3) Numerical solution of an integral equation.

2.3. Infinite Element Methods

The infinite element method belongs to the most attractive methods for the solution of exterior Helmholtz scattering problems.

To keep the setting simple, we consider only the scattering exterior to the unit sphere, i. e., we set $\Omega = \{\mathbf{x} \in \mathbb{R}^3 : |\mathbf{x}| < a\}$ and assume given Neumann data on the boundary:

$$\begin{aligned} \Delta u + k^2 u &= 0 && \text{in } \operatorname{ext}\Omega \\ \partial_n u &= g && \text{on } \partial\Omega \\ \partial_r u - iku &= o\left(\frac{1}{r}\right) && (r = |\mathbf{x}| \to \infty) \text{ uniformly in all directions.} \end{aligned} \tag{2.3.1}$$

In our representation here we follow the infinite element method proposed by Astley et. al., following the presentation given by Ihlenburg in **[61]**.

First, the scattering problem is rewritten in variational form: Find $u \in \mathrm{H}^1_{\mathrm{w}}(\operatorname{ext}\Omega)$ such that

$$b(v,u) = \langle v, g\rangle_{\partial\Omega} \qquad \text{for all } v \in \mathrm{H}^1_{\mathrm{w}^*}(\operatorname{ext}\Omega)$$

Here, the sesquilinear forms $b(\cdot,\cdot)$ and $\langle\cdot,\cdot\rangle_{\partial\Omega}$are defined by

$$\begin{aligned} b(v,u) &:= \int_{\operatorname{ext}\Omega} \left(\overline{\nabla v}\cdot\nabla u - k^2 \overline{v} u\right) d\mathrm{x} \\ \langle v, g\rangle_{\partial\Omega} &:= \int_{\partial\Omega} \overline{v} g \, d\mathrm{x} \end{aligned}$$

and

$$\begin{aligned} \mathrm{H}^1_{\mathrm{w}}(\mathrm{ext}\,\Omega) &= \left\{ u : \int_{\mathrm{ext}\,\Omega} \frac{1}{r^2} \left(|\nabla u|^2 + |u|^2 \right) d\mathrm{x} + \int_{\mathrm{ext}\,\Omega} |\partial_r u - iku|^2 \, d\mathrm{x} < 0 \right\} \\ \mathrm{H}^1_{\mathrm{w}*}(\mathrm{ext}\,\Omega) &= \left\{ u : \int_{\mathrm{ext}\,\Omega} r^2 \left(|\nabla u|^2 + |u|^2 \right) d\mathrm{x} < 0 \right\} \end{aligned}$$

are the trial and the test space, respectively. The subscript w indicates that the test and trial space defined this way are *weighted* Sobolev spaces. Based on the theoretical framework of Leis **[68]** this variational form was proposed by Demkovicz and Gerdes **[22]**. The existence and uniqueness of a solution of the variational problem is known. To construct a convergent discrete method, sequences of finite dimensional subspaces $V^1_{\mathrm{w}} \subset V^2_{\mathrm{w}} \subset \ldots \subset V^N_{\mathrm{w}} \subset \ldots \subset \mathrm{H}^1_{\mathrm{w}}(\mathrm{ext}\,\Omega)$ and $V^1_{\mathrm{w}*} \subset V^2_{\mathrm{w}*} \subset \ldots \subset V^N_{\mathrm{w}*} \subset \ldots \subset \mathrm{H}^1_{\mathrm{w}*}(\mathrm{ext}\,\Omega)$ must be constructed. The trial space must be such that the approximation property $\inf_{v^N \in V^N_{\mathrm{w}}} \left\| u - v^N \right\|_{\mathrm{H}^1_{\mathrm{w}}} \to 0$ as $N \to \infty$ for each $u \in \mathrm{H}^1_{\mathrm{w}}(\mathrm{ext}\,\Omega)$ holds true. The test space, together with the trial space and the variational formulation must guaranty that the discrete solution process is stable, i.e., the condition number of the discrete problem must be bounded. The practical construction of the trial and test spaces is based on the Wilcox expansion theorem **[90]**, which also supplies the basis for the Bayliss-Gunzburger-Turkel boundary conditions. Since we will come back several times to the Wilcox theorem and its 2D counterpart, the Karp theorem, we state them here explicitly.

THEOREM 2.3.1. *(The Wilcox Expansion Theorem* **[90]***.) Let $u \in C^2$ be a solution of (2.3.1) for the region exterior to a sphere $|\mathbf{x}| = a$, and let (r, θ, ϕ) be the spherical coordinates for $\mathbf{x}$. Then*

$$u(r, \theta, \phi) = \frac{\mathrm{e}^{ikr}}{kr} \sum_{j=0}^{\infty} \frac{F_j(\theta, \phi)}{(kr)^j}$$

where the series converges for $r > a$ and converges absolutely and uniformly in r, θ and ϕ in any region $|\mathbf{x}| \geq a + \epsilon > a$. The series may be differentiated term by term with respect to r, θ and ϕ any number of times and the resulting series all converge absolutely and uniformly.

The corresponding theorem for two dimensions has been proved by Karp in 1961:

THEOREM 2.3.2. *(The Karp Expansion Theorem* **[63]***.) Let $u \in C^2$ be a solution of (2.3.1) for the region exterior to a circle $|\mathbf{x}| = a$, and let (r, θ) be the cylindrical coordinates for $\mathbf{x}$. The there exists a convergent expansion, valid for $r > a$, in the form*

$$u = H_0^{(1)}(kr) \sum_{j=0}^{\infty} \frac{F_j(\theta)}{(kr)^j} + H_1^{(1)}(kr) \sum_{j=0}^{\infty} \frac{G_j(\theta)}{(kr)^j}$$

where the series converges uniformly and absolutely for $|\mathbf{x}| \geq a + \epsilon > a$, and may be differentiated term-wise with respect to r as often as desired.

Thus, the key-point in infinite element methods is to have an expansion theorem valid on the unbounded domain, together with a variational formulation which supplies stable discrete problems. In contrast, our Laplace domain methods derived from the pole condition are not based on series representations like the ones above (even not theoretically). Instead, the field representation on the unbounded domain is computed simultaneously with the interior solution.

2.4. Asymptotic Methods

In practice asymptotic approximations are used very often. The two main categories of asymptotic methods are the one where the distance between the obstacle and the artificial boundary tends to infinity and the one where the wavenumber tends to infinity. Especially the latter class allows to include problems with variable coefficients. We emphasize this here, since this is a key feature of the pole condition approach, developed in the main part of this work, too. The other key feature of the asymptotic methods is that the derived asymptotic boundary condition are local conditions.

Far-Field Approximation (Bayliss-Gunzburger-Turkel). In [**10**] Bayliss and Turkel and in [**9**] Bayliss, Gunzburger and Turkel developed a concept of asymptotic boundary conditions (BGT boundary conditions) characterized by the central property that these boundary conditions can be associated with local boundary conditions on the artificial boundary. Especially when used with direct sparse solvers, the BGT-type boundary conditions are have the advantage as opposed to boundary conditions based on standard integral formulations because that they do not destroy the sparsity (at least in their lowest order form) of the FEM-discretization of the interior problem.

The starting point of the derivation of the BGT boundary conditions is, exactly as in the case of infinite elements, the existence of a convergent Wilcox expansion in 3D [**90**], cf. Theorem 2.3.1 and the Karp expansion in 2D [**63**], cf. Theorem 2.3.2. We consider the derivation of the BGT boundary conditions for the 2D case. Since the series expansion from Theorem 2.3.2 is difficult to work with, it is replaced by a series of the simpler 3D type of Theorem 2.3.1 using the well-known asymptotic expansion of the Hankel functions (cf. Corollary 4.4.5 in Section 4.2, p. 84)

$$u(r,\phi) \sim \sqrt{\frac{2}{\pi kr}} \mathrm{e}^{i(kr-\pi/2)} \sum_{j=0}^{\infty} \frac{f_j(\phi)}{r^j}. \tag{2.4.1}$$

Note, that this manipulation results in a series which is no longer convergent. Now the central idea of the BGT boundary conditions is the construction of a family of boundary operators B_m, $m = 1, 2, \ldots$, such that the action of the operator $B_m(r, \partial_r)$ applied to the series (2.4.1) results in

$$B_m(r,\partial_r)\, u(r,\phi) \sim \sqrt{\frac{2}{\pi kr}} \mathrm{e}^{i(kr-\pi/2)} \sum_{j=2m}^{\infty} \frac{f_j(\phi)}{r^j}. \tag{2.4.2}$$

Hence the operator is designed such that it annihilates the leading $2m$ terms (the terms with $j = 0, \ldots, 2m-1$) of the series expansion. It is not immediately clear that the operator B_m defines an asymptotic boundary condition. To discuss this point we consider a parameterization

$$\begin{aligned} \gamma : [0,1] &\to \mathbb{R}_+ \times [-\pi, \pi] \\ s &\mapsto (r,\phi) \in \partial\Omega \end{aligned}$$

of the artificial boundary $\partial\Omega$ in polar coordinates. We set $B_m(r,\partial_r)\, u(r,\phi)|_{r=r(s)} = 0$ for each parameter $s \in [0,1]$ which supplies a relation between the function $u(r(s),\phi(s))$ at each point on the boundary (the Dirichlet data) and radial derivatives at this point. Thus, asymptotic boundary conditions of the form $B_m u = 0$ yield local boundary conditions containing, in general, higher order derivatives in radial direction.

To construct the family of operators B_m we use the identities

$$\left(\partial_r - ik + \frac{j+\frac{1}{2}}{r}\right)\frac{\mathrm{e}^{ikr}}{r^{j+1/2}} = 0 \tag{2.4.3}$$

$$\left(\partial_r - ik + \frac{j+\frac{1}{2}}{r}\right)\frac{\mathrm{e}^{ikr}}{r^{j+1+1/2}} = -\frac{\mathrm{e}^{ikr}}{r^{j+2+1/2}}\,. \tag{2.4.4}$$

We set

$$B_1 = \partial_r - ik + \frac{1}{2r}$$

which annihilates the zeroth and the first order term of the asymptotic expansion (2.4.1), i.e., it holds $B_1 u = O\left(1/r^{2+1/2}\right)$ for $r \to \infty$. Based on (2.4.3) and (2.4.4) we find that the application of $(\partial_r - ik + (j+1/2)/r)$ annihilates both the jth and the $j+1$th terms in the series (2.4.1) yielding the recurrence

$$B_j = \left(\partial_r - ik + \frac{2j-\frac{3}{2}}{r}\right) B_{j-1},\ j \geq 1. \tag{2.4.5}$$

with an asymptotic behavior

$$B_m u = \mathcal{O}\left(\frac{1}{r^{2m+1/2}}\right),\ m \geq 1 \quad (r \to \infty).$$

We add the results for three dimensions. The same procedure applied to the Atkinson-Wilcox expansion supplies the recurrence

$$B_j = \left(\partial_r - ik + \frac{2j-1}{r}\right) B_{j-1},\ j \geq 1$$

with $B_0 := 1$ and an asymptotic behavior

$$B_m u = \mathcal{O}\left(\frac{1}{r^{2m+1}}\right),\ m \geq 1 \quad (r \to \infty).$$

Operator Factorization. In two important papers, Engquist and Majda **[28, 29]** constructed a family of transparent boundary conditions for the wave equation. An error analysis of this scheme specialized to the constant coefficient scalar wave equation, carried out by Halpern and Rauch **[54]**, shows that the error consists of two terms. One is proportional to the largest reflection coefficient for the artificial boundary condition, and can be made arbitrary small in case of a planar boundary. The second term is inverse proportional to the average frequency present in the solution. Roughly speaking, the Engquist-Majda construction becomes the better the smoother the artificial boundary and the higher the frequency of the wave. Since their method is very general and can take into account both variable coefficients and general curved boundaries, we repeat the main steps of the construction here. Instead of summarizing the full theory in the following, we will introduce only the very basic quantities needed to motivate and develop the desired technique of operator factorization. The applicability and limitations of the method are studied using the Helmholtz equation with constant coefficients and a fixed angular frequency in the exterior of a circle as a model problem. We introduce the method here for two reasons:

(1) The method is applicable to very general situations. It allows arbitrary curved, smooth boundaries and variable coefficients in the exterior domain.
(2) The method is directly applicable to the semi-discrete formulation of the scattering problem to be derived in Section 5.3. Hence it might be considered as one of the several possible approximate realizations of our exact solution of the basic semi-discrete formulation.

Since we are interested in the solution of the Helmholtz equation, we specialize the general consideration of **[28, 29]** to the time-harmonic case. Further we use the Helmholtz operator L in cylindric coordinates in 2D by

$$L := \partial_r^2 + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_\phi^2 + n^2(\phi)k^2.$$

In the special variable coefficient case which we want to consider here, we allow a variation of the wavenumber with respect to the angular coordinate ϕ, denoted by the scalar, real refractive index function $n(\phi)$. We want to compute a family of boundary conditions at $r = r_0$. The key tool to derive the boundary conditions is a factorization of the differential operator L by methods arising on the theory of pseudodifferential operators. Before we study the method, we recall some basic facts about pseudodifferential operators which are used by Engquist and Majda, see the book of Taylor **[87]**. Our presentation here follows the one given in **[29]** for the variable coefficient case. We need the definition of a symbol class, the construction of the associated pseudodifferential operators and the asymptotic expansion of the symbol of composite pseudodifferential operators.

Let Ω be an open subset of $\mathbb{R}^n$. The complex, infinitely differentiable function $b(\mathbf{x}, \boldsymbol{\xi}) \in C^\infty(\Omega \times \mathbb{R}^n)$ is an element of the symbol class $S^m(\Omega)$ provided there exists constants $C_{\alpha,\beta}$ such that

$$\left| \frac{\partial^{|\alpha|}}{\partial \mathbf{x}^\alpha} \frac{\partial^{|\beta|}}{\partial \boldsymbol{\xi}^\beta} b(\mathbf{x}, \boldsymbol{\xi}) \right| \leq C_{\alpha,\beta} (1 + |\boldsymbol{\xi}|)^{m-|\beta|} \tag{2.4.6}$$

for any multi-indices α, β, ([**87**, Chapt. II, Def. 1.1]). Associated with any symbol $b(\mathbf{x}, \boldsymbol{\xi})$ is a pseudodifferential operator $b(\mathbf{x}, (1/i)\,\partial/\partial\mathbf{x}) \in OP(S^m)$ defined via the Fourier transform by

$$b\left(\mathbf{x}, \frac{1}{i}\frac{\partial}{\partial \mathbf{x}}\right) v(\mathbf{x}) = \int \exp(i\mathbf{x}\cdot\boldsymbol{\xi})\, b(\mathbf{x}, \boldsymbol{\xi})\, \widehat{v}(\boldsymbol{\xi})\, d\boldsymbol{\xi}\,. \tag{2.4.7}$$

Conversely, each pseudodifferential operator $b(\mathbf{x}, (1/i)\,\partial/\partial\mathbf{x})$ has a symbol $\sigma(b(\mathbf{x}, (1/i)\,\partial/\partial\mathbf{x})) = b(\mathbf{x}, \boldsymbol{\xi})$. Note that the symbol of a pseudodifferential operator does not contain any differential operator. Given two pseudodifferential operators $a(\mathbf{x}, (1/i)\,\partial/\partial\mathbf{x}) \in OP(S^{m_1})$ and $b(\mathbf{x}, (1/i)\,\partial/\partial\mathbf{x}) \in OP(S^{m_2})$ it holds $a \circ b \in OP(S^{m_1+m_2})$ and the symbol of the composite pseudodifferential operator $a \circ b$ possesses the asymptotic expansion ([**87**, Chapt. II, Theorem 4.4])

$$\sigma(a \circ b) \sim \sum_{|\alpha|\geq 0} \frac{(-i)^{|\alpha|}}{\alpha!} \frac{\partial^{|\alpha|}}{\partial \boldsymbol{\xi}^\alpha} a(\mathbf{x}, \boldsymbol{\xi}) \frac{\partial^{|\alpha|}}{\partial \mathbf{x}^\alpha} b(\mathbf{x}, \boldsymbol{\xi}) \qquad (|\boldsymbol{\xi}| \to \infty)\,. \tag{2.4.8}$$

In our special case we define $\mathbf{x}$ by 2D cylindric coordinates $\mathbf{x} = (r, \phi)$ and consider the exterior of a circle with radius r_0, that is $\Omega = \{\mathbf{x} \in \mathbb{R}^2 : r > r_0, -\pi < \phi \leq \pi\}$. Let us denote the Fourier coordinates dual to the original coordinates by $\boldsymbol{\xi} = (\rho, \nu)$.

The key of the Engquist-Majda approach is the construction of pseudodifferential operators $\lambda_+(\mathbf{x}, (1/i)\,\partial/\partial\mathbf{x})$ and $\lambda_-(\mathbf{x}, (1/i)\,\partial/\partial\mathbf{x})$ such that the factorization

$$L = (\partial_r - \lambda_+)(\partial_r - \lambda_-) + (\text{smooth error}) \tag{2.4.9}$$

holds true. This factorization requires the computation of the symbol of the composite pseudodifferential operators $\sigma(\lambda(\mathbf{x}, (1/i)\,\partial/\partial\mathbf{x})\,\partial_r)$ and $\sigma(\partial_r\,\lambda(\mathbf{x}, (1/i)\,\partial/\partial\mathbf{x}))$,

where λ is either λ_+ or λ_-. It follows directly from (2.4.7) that

$$\begin{aligned}\left[\lambda\left(\mathbf{x},\frac{1}{i}\frac{\partial}{\partial\mathbf{x}}\right)\partial_r\right]u(\mathbf{x}) &= \int\exp(i\mathbf{x}\cdot\boldsymbol{\xi})\,\lambda(\mathbf{x},\boldsymbol{\xi})\,\widehat{\partial_r u}\,d\boldsymbol{\xi}\\ &= \int\exp(i\mathbf{x}\cdot\boldsymbol{\xi})\,[\lambda(\mathbf{x},\boldsymbol{\xi})\,i\xi]\,\widehat{u}(\boldsymbol{\xi})\,d\boldsymbol{\xi}.\end{aligned}$$

Similarly, we have

$$\begin{aligned}\partial_r\left[\lambda\left(\mathbf{x},\frac{1}{i}\frac{\partial}{\partial\mathbf{x}}\right)u(\mathbf{x})\right] &= \left[\partial_r\lambda\left(\mathbf{x},\frac{1}{i}\frac{\partial}{\partial\mathbf{x}}\right)\right]u(\mathbf{x})+\left[\lambda\left(\mathbf{x},\frac{1}{i}\frac{\partial}{\partial\mathbf{x}}\right)\partial_r\right]u(\mathbf{x})\\ &= \int\exp(i\mathbf{x}\cdot\boldsymbol{\xi})\,[\partial_r\lambda(\mathbf{x},\boldsymbol{\xi})+\lambda(\mathbf{x},\boldsymbol{\xi})\,i\xi]\,\widehat{u}(\boldsymbol{\xi})\,d\boldsymbol{\xi}.\end{aligned}$$

This shows the identities

$$\sigma\left(\lambda\left(\mathbf{x},\frac{1}{i}\frac{\partial}{\partial\mathbf{x}}\right)\partial_r\right) = \lambda(\mathbf{x},\boldsymbol{\xi})\,i\xi \tag{2.4.10}$$

$$\sigma\left(\partial_r\lambda\left(\mathbf{x},\frac{1}{i}\frac{\partial}{\partial\mathbf{x}}\right)\right) = \lambda(\mathbf{x},\boldsymbol{\xi})\,i\xi+\partial_r\lambda(\mathbf{x},\boldsymbol{\xi}). \tag{2.4.11}$$

Hence we have recovered a very special form of the asymptotic expansion (2.4.8), where we can replace the asymptotic equality by the equality itself.

We shall show experimentally that we can construct sequences of pseudodifferential operators λ_+^1, $\lambda_+^0,\ldots,\lambda_+^{-j}$ and λ_-^1, $\lambda_-^0,\ldots,\lambda_-^{-j}$, $1\geq j\geq -1$ such that the following crucial asymptotic properties hold true, with $u(r)=\mathrm{H}_\nu^{(1)}(knr)$ the exact solution of Bessel's equation:

i) *Far field asymptotics.*

$$\frac{\left(\partial_r-\sum_{k=-1}^{j}\lambda_-^{-k}\left(\mathbf{x},\frac{1}{i}\frac{\partial}{\partial\mathbf{x}}\right)\right)u(r)}{u(r)}=\mathcal{O}\left(\frac{1}{r^{j+3}}\right)\qquad\text{for } r\to\infty \text{ and fixed } \nu.$$

ii) *High frequency asymptotics in ν.*

$$\frac{\left(\partial_r-\sum_{k=-1}^{j}\lambda_-^{-k}\left(\mathbf{x},\frac{1}{i}\frac{\partial}{\partial\mathbf{x}}\right)\right)u(r)}{u(r)}=\mathcal{O}\left(\frac{1}{|\nu|^{2^{j+1}}}\right)\qquad\text{for } |\nu|\to\infty \text{ and fixed } r,$$

and $\mathrm{Im}\left\{\sigma\left(\lambda_-^j\right)\right\}>0$ as $r\to\infty$ and ν fixed. Once a function $u(\mathbf{x})$ in Ω is known which satisfies $(\partial_r-\lambda_-)\,u=0$ it follows $Lu\cong 0$, where $\cong$ means equal up to a smooth error. Viewing λ_- as a constant with positive imaginary part, the equation $(\partial_r-\lambda_-)\,u=0$, derived from (2.4.9), possesses a solution $\exp(\lambda_- r)$ showing that $u(\mathbf{x})$ is an outgoing solution with respect to r. Hence, if a function $u(\mathbf{x})$ satisfies $(\partial_r-\lambda_-)\,u=0$, it satisfies $Lu\cong 0$ and consists only of outgoing functions. Consequently, the restriction of $\mathbf{x}$ to the boundary supplies the desired boundary condition (up to the factorization error)

$$\left(\partial_r-\lambda_-\left(\mathbf{x},\frac{1}{i}\frac{\partial}{\partial\mathbf{x}}\right)\right)u(\mathbf{x})=0\qquad\mathbf{x}\in\partial\Omega.$$

To obtain a family of factorizations of ascending degree of accuracy, we construct sequences of pseudodifferential operators λ_+^1, $\lambda_+^0,\ldots,\lambda_+^{-j}$ and λ_-^1, $\lambda_-^0,\ldots,\lambda_-^{-j}$,

$j \geq -1$, such that $\sigma\left(\lambda_{\pm}^{-j}\right) \in S^{-j}$ and the symbol of the residual satisfies

$$\sigma\left(L - \left(\partial_r - \sum_{k=-1}^{j} \lambda_{+}^{-k}\right)\left(\partial_r - \sum_{k=-1}^{j} \lambda_{-}^{-k}\right)\right) \in S^{-j}. \tag{2.4.12}$$

The idea behind this is to approximate λ_{-} by $\sum_{k=-1}^{j} \lambda_{-}^{-k}$ such that the terms $\geq j$ are all smoother than the initial terms since $\sigma\left(\lambda_{-}^{-j-k}\right) = \mathcal{O}\left(|\boldsymbol{\xi}|^{-j-k}\right)$ for $k \geq 1$, by definition of the symbol class (2.4.6). Consequently we construct a method adapted to situations where high frequencies play the dominant role. First we compute the symbol of the Helmholtz operator

$$\sigma(L) = -\rho^2 + \frac{1}{r} i\rho - \frac{1}{r^2}\nu^2 + n^2 k^2 . \tag{2.4.13}$$

Its principal part consists of the terms quadratic in ρ, ν, k (since we view it as time-harmonic wave equation) and possesses a factorization

$$-\rho^2 - \left[(1/r^2)\,\nu^2 - n^2 k^2\right] = \left(i\rho + \sqrt{(1/r^2)\,\nu^2 - n^2 k^2}\right)\left(i\rho - \sqrt{(1/r^2)\,\nu^2 - n^2 k^2}\right),$$

where we take the branch of the square root such that its real part is positive and $\sqrt{-1} = +i$. Motivated by this factorization we choose

$$\lambda_{\pm}^{1}\left(\mathbf{x}, \frac{1}{i}\frac{\partial}{\partial \mathbf{x}}\right) = \mp\sqrt{-\frac{1}{r^2}\partial_\phi^2 - n^2 k^2}\,.$$

Its symbol is

$$\lambda_{\pm}^{1}(\mathbf{x}, \boldsymbol{\xi}) = \mp\sqrt{\frac{1}{r^2}\nu^2 - n^2 k^2} \tag{2.4.14}$$

and is bounded by

$$\left|\lambda_{\pm}^{1}(\mathbf{x}, \boldsymbol{\xi})\right| \leq C\left(1 + |\xi|\right), \quad C = \max\left(nk, \frac{1}{r}\right).$$

However, already the first derivative $\partial_\nu \lambda_{\pm}^{1}(\mathbf{x}, \boldsymbol{\xi})$ generates a factor $1/\sqrt{\nu^2/r^2 - n^2 k^2}$ which prevents to compute a constant $C_{\alpha,1}$ as required by (2.4.6) in the definition of the symbol class. Hence the concept to realize smoother and smoother approximations can be realized only in the cones

$$\frac{|\nu^2|}{r^2} < n^2 k^2 \tag{2.4.15}$$

$$\frac{|\nu^2|}{r^2} > n^2 k^2. \tag{2.4.16}$$

We refer to the first one as *far-field regime* and to the latter one as *high-frequency regime* in ν. Note that the high-frequency regime in k is identical with the far-field regime. We show that the choice (2.4.14) satisfies condition (2.4.12) with $j = -1$. Based on (2.4.10) and (2.4.11) we compute first the symbol of the composite operators

$$\begin{aligned} \sigma\left(\partial_r \lambda_{\pm}^{1}\right) &= i\rho\lambda_{\pm}^{1}(\mathbf{x}, \boldsymbol{\xi}) + \tfrac{\partial}{\partial r}\lambda_{\pm}^{1}(\mathbf{x}, \boldsymbol{\xi}) \\ \sigma\left(\lambda_{\pm}^{1}\,\partial_r\right) &= i\rho\lambda_{\pm}^{1}(\mathbf{x}, \boldsymbol{\xi}). \end{aligned}$$

This shows that the highest order terms of the residual

$$\begin{aligned} \sigma\left(L - \left(\partial_r - \lambda_{+}^{1}\right)\left(\partial_r - \lambda_{-}^{1}\right)\right) &= \sigma\left(L - \partial_r^2 + \partial_r \lambda_{-}^{1} + \lambda_{+}^{1}\partial_r - \lambda_{+}^{1}\lambda_{-}^{1}\right) \\ &= \tfrac{1}{r} i\rho + \tfrac{\partial}{\partial r}\lambda_{-}^{1}(\mathbf{x}, \boldsymbol{\xi}) \end{aligned}$$

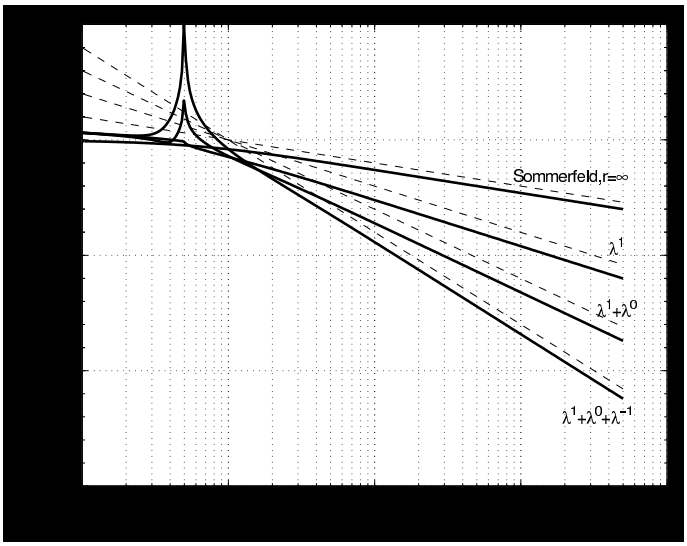


FIGURE 2.4.1. Error of the asymptotic boundary conditions ($\nu = 0$). The dashed lines represent the functions r^{-1}, r^{-2}, r^{-3}, r^{-4}.

are in fact first order terms, hence the residual is in S^1, if we modify the definition of S^1, see (2.4.6), such that it contains elements defined on the cones (2.4.15) or (2.4.16). Engquist and Majda extend the procedure up to pseudodifferential operators $\lambda_{\pm}^{-1}$. That is, we approximate $\lambda_{\pm} \simeq \lambda_{\pm}^{1}+\lambda_{\pm}^{0}+\lambda_{\pm}^{-1}$ and set the appropriate terms in the residual equation to zero. This yields

$$\begin{array}{rcrcll} 0 & = & \frac{1}{r} i\rho + i\rho\left(\lambda_{+}^{0}+\lambda_{-}^{0}\right) & + & \partial_r \lambda_{-}^{1} - \left(\lambda_{+}^{1}\lambda_{-}^{0}+\lambda_{+}^{0}\lambda_{-}^{1}\right) & \text{for } \mathcal{O}\left(|\boldsymbol{\xi}|\right) \\ 0 & = & i\rho\left(\lambda_{+}^{-1}+\lambda_{-}^{-1}\right) & + & \partial_r \lambda_{-}^{0} - \left(\lambda_{+}^{1}\lambda_{-}^{-1}+\lambda_{+}^{0}\lambda_{-}^{0}+\lambda_{+}^{-1}\lambda_{-}^{1}\right) & \text{for } \mathcal{O}\left(1\right). \end{array}$$

We have the freedom to prescribe a relation between λ_{+}^{j} and λ_{-}^{j} as long as the order condition is not violated, and following **[29]** we use $\lambda_{+}^{-j} = -\lambda_{-}^{-j}$ for $j \geq 0$ to obtain two equations for the symbols $\lambda_{-}^{0}(\mathbf{x}, \boldsymbol{\xi})$ and $\lambda_{-}^{-1}(\mathbf{x}, \boldsymbol{\xi})$

$$\begin{array}{rcrcl} 0 & = & \frac{1}{r} i\rho & + & \partial_r \lambda_{-}^{1} - \lambda_{-}^{0}\left(\lambda_{+}^{1}-\lambda_{-}^{1}\right) \\ 0 & = & & + & \partial_r \lambda_{-}^{0} - \lambda_{-}^{-1}\left(\lambda_{+}^{1}-\lambda_{-}^{1}\right) + \lambda_{-}^{0}\lambda_{-}^{0}\,. \end{array}$$

The first of these equations supplies λ_{-}^{0} using $\lambda_{\pm}^{1}$ and the second supplies λ_{-}^{-1} using $\lambda_{\pm}^{1}$and λ_{-}^{0}.

Example. We study Bessel's equations with $nk = 1$. First we investigate the far-field asymptotic for the two fixed frequencies $\nu = 0$ and $\nu = 20$. Fig. 2.4.1 shows the relative error of the normal derivative as function of the radius r. Clearly, the result becomes better if the order of the approximation is increased. Additionally we observe that this behavior is *not* uniform. Close to the resonance case $nk = 1 = |\nu/r|$, which separates the cones (2.4.15) and (2.4.16), the approximation fails.

The same observation can be made in Fig. 2.4.2 for $\nu = 20$. Again we find that the approximation fails close to $nk = 1 = |\nu/r|$. To justify this, Fig. 2.4.3 displays a zoom into Fig. 2.4.2 around the position $\nu/r = 20$.

Next, we investigate the the high-frequency asymptotic. Fig. 2.4.4 shows the error for a fixed radius $r = 0.1$ as function of the frequency ν. Again, we find that

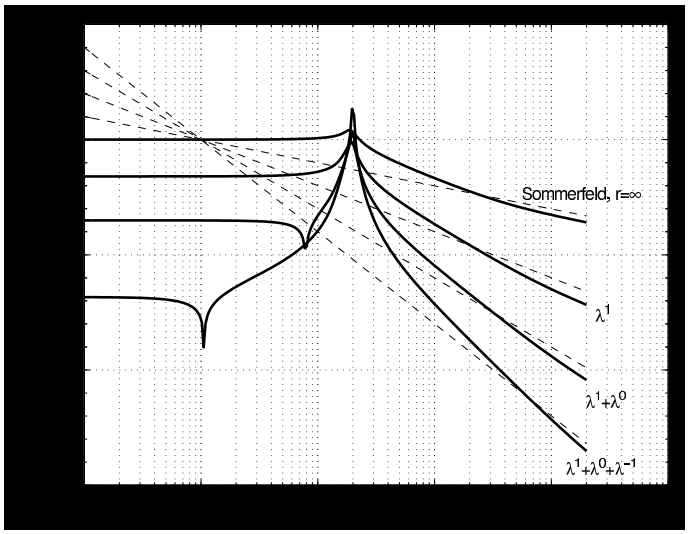


FIGURE 2.4.2. Error of the asymptotic boundary conditions ($\nu = 20$). The dashed lines represent the functions r^{-1}, r^{-2}, r^{-3}, r^{-4}.

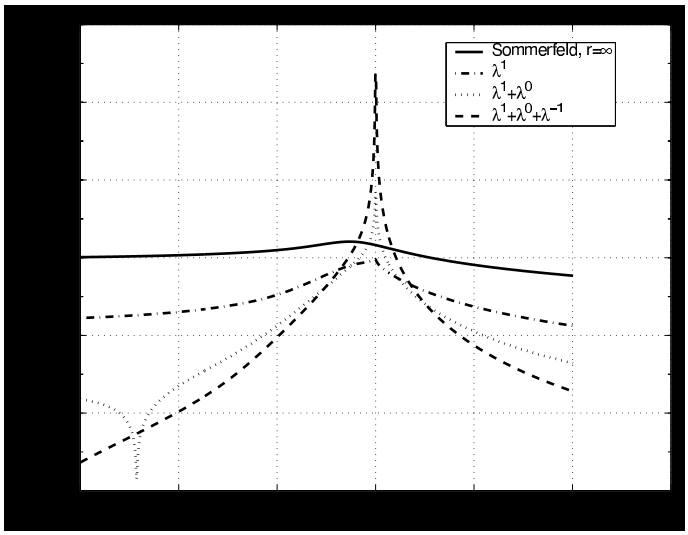


FIGURE 2.4.3. Zoom of the error plot Fig. 2.4.2 into the interval $5 \leq r \leq 30$

the approximation is the better the higher the order of the approximation for ν large. As in the far-field case above, the approximation fails for $nk = 1 = |\nu/r|$, as it becomes apparent from Fig. 2.4.5, which shows the error for $r = 10$, fixed. In conclusion, the experiments show that the factorization technique works well both in the far-field and in the high-frequency regime. In Section 5.3, p. 138, we will apply this technique to derive a high frequency approximation for the boundary condition in a general case.

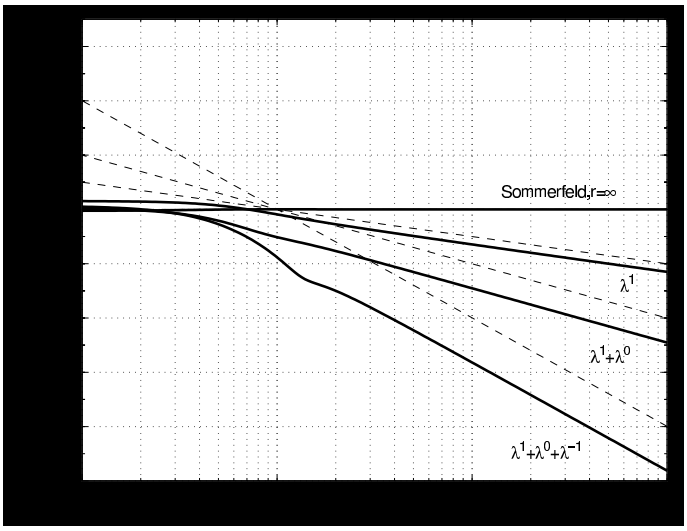


FIGURE 2.4.4. Error of the asymptotic boundary condition ($r = 0.1$). The dashed lines indicate the functions ν^0, ν^{-1}, ν^{-2}, ν^{-4}.

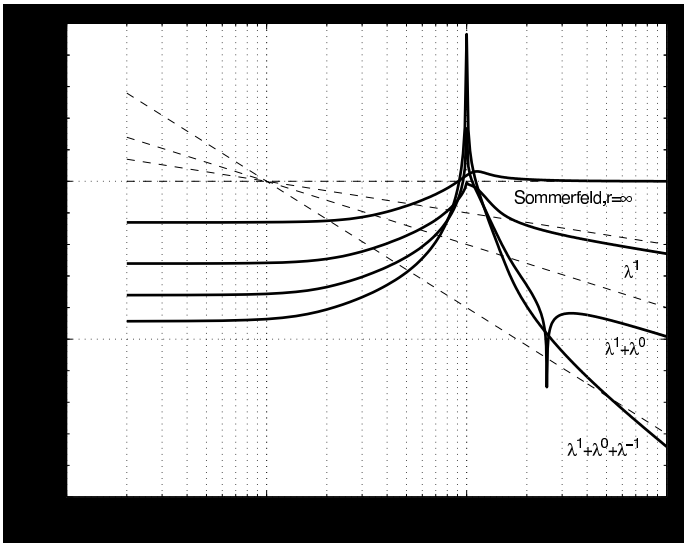


FIGURE 2.4.5. Error of the asymptotic boundary condition ($r = 10$). The dashed lines indicate the functions ν^0, ν^{-1}, ν^{-2}, ν^{-4}.

2.5. Further Methods

There are many other important methods tailored to solve scattering problems in exterior domains. Starting from an integral representation based on explicitly known Green's functions, both *multipole methods* and *panel clustering* seek to minimize the numerical effort to obtain a solution with prescribed accuracy.

Multipole Method. The multipole methods have been developed to have fast summation algorithms for situations, where the potential resulting from many charges has to be computed simultaneously at a number of positions. The first

multipole method has been proposed by Rokhlin **[76]**, where he applied the idea to discretized integral equations with a non-oscillatory kernel. Since the contributions of Greengard **[41]**, **[40]** and **[15]** the method has been established as standard tool. In 1993 Rokhlin extended the concept to the 3D Helmholtz equation **[77]**.

Panel Clustering. The panel clustering technique goes back to Hackbusch and Nowak **[47]** and has been improved by Hackbusch, Sauter and Lage in **[46, 79]**. As opposed to the multipole method, the panel clustering aims to make the boundary integral method efficient. The thesis of Giebermann **[34]** gives both an overview and many details of these fast summation methods.

Ficticious domain method. A further method for $k \geq 1$ is the so-called ficticious domain method, cf. Ernst **[30]** and Heikkola et. al. **[56]**. The idea is to surround the computational domain by sufficient large artificial domain, for which a good asymptotic approximation of the exact transparent boundary condition is known. The technique becomes fast, if the artificial layer can be efficiently discretized by tensorial meshes.

The difference of all these methods to our approach is that they are either based on an explicitely known Green's function or radiation condition, whereas our approaches do not need such a-priori knowledge. The last class of methods which we want to mention here works without using Green's functions.

Absorbing Layers. The currently perhaps most widely used and most successful approach is the perfectly matched layer technique (abbreviated to PML afterwards) proposed by Bérenger **[11]** in 1994 to solve the time-dependent scattering problem of Maxwell's equations in 2D, and extended in 1996 by him to 3D problems **[12]**. After Bérenger's initial proposal a large number of modifications and extensions to other problem classes than time-dependent Maxwell's equations has been made. Among them are the generalization to time-harmonic and curvilinear coordinates (instead of the usual Cartesian grids) by Collino and Monk **[17]**, the application to paraxial equations like the Fresnel equation by Collino **[18]**, the analysis of different possible auxiliary forms (so-called split-field, biaxial and uniaxial forms) **[91]** by Yang and Petropoulus and Petropoulus in **[75]**, and the convergence proof of Lassas and Somersalo **[67]**, to mention only a very few of the many interesting articles.

Before Bérenger's publication appeared, many other forms of absorbing layers had been developed, mainly motivated by the modeling of a physical absorption. However, caused by the success of the PML-technique, most of them are no longer in use.

The main idea of absorbing layers is to surround the artificial boundary by an artificial layer of finite thickness. The goal in the design of these layers is to avoid spurious reflections from the artificial boundary and the interior of the absorbing layer to the greatest possible extend. Bérenger showed how to modify Maxwell's equations to provide a perfectly matched absorber. That is, the waves are absorbed independently from their frequency and incidence direction. Moreover, the waves decay exponentially with the distance into the layer. Thus the PML layer can be truncated after a relatively short distance. The PML layer must be discretized like the interior problem and the discretized interior problem and the modified problem on the PML domain have to be solved simultaneously.

Surprisingly, it turned out that the PML method has much in common with the Laplace domain methods which forms the main topic of this work. In fact, from a certain point of view we can see the PML technique as one (among other possible)

realization of the pole condition, compare Section 4.4.6 for a short discussion of this issue.

CHAPTER 3

Characterization of Incoming and Outgoing Waves: Pole Condition

This chapter introduces our central condition, the pole condition. We start with the analysis of non-trivial examples in 1D and show that their far-field is characterized by the singularities of the Laplace transformed solutions. We discuss the physical meaning of the solution representation in terms of the Laplace transform, which shows that the Laplace transform is a convenient tool to represent functions on semi-infinite intervals. The essence of the examples leads to the definition of the pole condition. We extend the pole condition to both higher space dimensions and time-dependent problems. The crucial definitions are the following:

(1) Pole condition in 1D for continuous functions, Definition 3.1.1. This refers to potentials which become constant at infinity.
(2) Pole condition in arbitrary space dimensions for continuous functions, Definition 3.4.5. This is a generalization of the first point.
(3) Pole condition for time-dependent problems. A direct generalization to Schrödinger's equation.

In this section we present and analyze three characteristic examples, which can be treated almost analytically. These are

- The Helmholtz equation in 1D with constant coefficients.
- The Helmholtz equation in 1D with periodic coefficients.
- The radially symmetric Helmholtz equation in 2D with constant coefficients.

The Helmholtz equation in 1D is the most simple equation and therefore well suited to introduce the concept. The Bessel equation is a 1D equation, too, nevertheless it characterizes the situation in higher space dimension and is therefore of fundamental interest. We present an analytic solution, up to a convolution integral, of Bessel's equation in Laplace domain. This will motivate our treatment of the generalized Bessel (Hankel) functions, which provides the basis of the general theory for problems with variable coefficients. Finally, we consider the case of periodic coefficients. This is of interest for a number of practical problems and shows, how theory and algorithms can be extended to cover this problem class.

3.1. Helmholtz Equation in 1D with Constant Potential

Let us consider the Helmholtz equation in the semi-infinite interval $[a, \infty)$ and let us introduce the shifted coordinate-system with the independent variable x via $r = a + x, x \geq 0$,

$$\frac{d^2}{dx^2} u(x) + u(x) = 0, \quad x > 0.$$

Here we set $\text{w.r.o g.} k^2 = 1$. Laplace transform

$$\widehat{u}(s) = (Lu(x))(s) = \int_0^\infty \mathrm{d}x\, e^{-sx} u(x), \qquad \operatorname{Re} s > 0 \tag{3.1.1}$$

yields

$$\hat{u}(s) = \frac{su(0) + \partial_x u|_{x=0}}{s^2+1}\,.$$

In the following we abbreviate always: $u'(0) := \partial_x u|_{x=0}$. The complex valued function $\hat{u}(s)$ is holomorphic in $\mathbb{C} \setminus \{\pm i\}$ and possesses two isolated singularities, namely the points $\{\pm i\}$. Hence $\hat{u}(s)$, can be extended form $\operatorname{Re} s > 0$, where it has been defined to the whole complex plane except the singular points. A partial fraction decomposition supplies

$$\hat{u}(s) = \frac{1}{2}\frac{u(0)+iu'(0)}{s+i} + \frac{1}{2}\frac{u(0)-iu'(0)}{s-i}.$$

Obviously, the pole at $s = -i$ vanishes, if the 1D Sommerfeld condition holds: $u(0) + iu'(0) = 0$. We will call this condition pole condition. Thus, the Laplace-technique supplies two essential informations:

(1) It gives a natural splitting of a given function into asymptotically incoming and outgoing waves, because we can identify a complex function $\hat{u}_-(s) := \frac{1}{s+i}$ with an incoming wave $u_-(x) = \exp(-ix)$ and a function $\hat{u}_+(s) := \frac{1}{s-i}$ with an outgoing wave $u_+(x) = \exp(ix)$.
(2) It supplies automatically the relation between Neumann data $u'(0)$ and the Dirichlet data $u(0)$ needed to drop one of them.

DEFINITION 3.1.1. *Pole condition in 1D.* A function $u : \mathbb{R}_+ \to \mathbb{C}$ satisfies the pole condition if the function $\hat{u}$ defined by the Laplace transform (3.1.1) has a holomorphic extension to the lower half of the complex plane $\{s \in \mathbb{C} : \operatorname{Im} s < 0\}$.

REMARK 3.1.2. The pole condition is a condition concerning the behavior of u at infinity, i.e. it is independent of the choice of the boundary of the computational domain. If we replace the boundary at $x = a$ by a boundary at $a+\delta$, this property follows from

$$\begin{aligned}\int_0^\infty \mathrm{d}x\, e^{-sx}u(x) &= \int_0^\delta \mathrm{d}x\, e^{-sx}u(x) + \int_{-\delta}^\infty \mathrm{d}x\, e^{-sx}u(x)\\ &= \int_0^\delta \mathrm{d}x\, e^{-sx}u(x) + e^{-s\delta}\int_0^\infty \mathrm{d}x\, e^{-sx}u(x+\delta)\end{aligned}$$

and the fact that both $s \mapsto \int_0^\delta \mathrm{d}x\, e^{-sx}u(x)$ and $s \mapsto e^{-s\delta}$ are entire functions.

3.2. Helmholtz Equation in 1D with Periodic Potentials

We consider real, positive and piecewise continuous potentials with period $L > 0$, that is $p(x+L) = p(x)$, $x \geq 0$ along with the 1D Helmholtz equation

$$u''(x) + \left(k^2 + p(x)\right)u(x) = 0, \qquad x > 0 \tag{3.2.1}$$

with initial data $u(0)$ and $u'(0)$. If $u(x)$ is a solution to (3.2.1), it satisfies the *conservation property* $\operatorname{Im}(\overline{u}(x)u'(x)) = \text{const}$ for all $x \geq 0$. The quantity $P := \operatorname{Im}(\overline{u}(x)u'(x))$ is called power flux. The conservation property is derived by a multiplication of (3.2.1) by $\overline{u}$, followed by an integration by parts and a computation of the imaginary part of the resulting expression.

It is convenient to rewrite the Helmholtz equation (3.2.1) into a first order system

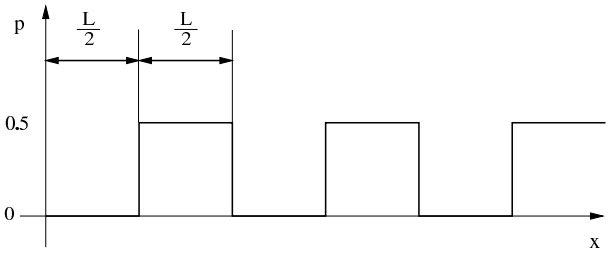


FIGURE 3.2.1. Example of a periodic potential $p(x)$

$$\left(\begin{array}{c} u \\ v \end{array}\right)' = \left(\begin{array}{cc} 0 & 1 \\ -k^2-p & 0 \end{array}\right)\left(\begin{array}{c} u \\ v \end{array}\right), \qquad x>0, \tag{3.2.2}$$

with initial data $u(0)$ and $v(0)=u'(0)$. The points $\{x_j=(j-1)L\,:\,j=1,2,\ldots\}$ define a partition of $\mathbb{R}_+$ into intervals $[x_j,x_{j+1}]$. We consider a solution of (3.2.2) on the interval $[x_j,x_{j+1}]$ with respect to the initial data $u_j:=u(x_j)$ and $v_j:=v(x_j)$. Let us denote the map from the initial data at x_j to the data at x_{j+1} as

$$\begin{aligned} \left(\begin{array}{c} u \\ v \end{array}\right)_{j+1} &= \mathrm{W}\left(\begin{array}{c} u \\ v \end{array}\right)_j, \qquad \mathrm{W}\in\mathbb{R}^{2\times 2},\ j=1,2,\ldots \\ \left(\begin{array}{c} u \\ v \end{array}\right)_1 &= \left(\begin{array}{c} u_1 \\ v_1 \end{array}\right)=\left(\begin{array}{c} u(0) \\ v(0) \end{array}\right). \end{aligned} \tag{3.2.3}$$

Here, we understood that $(u,v)^T$ has a continuous continuation to the boundary of the interval. The matrix W is regular since a backward integration exists on the same interval. Further, W does not depend on j, due to the periodicity of p. The repeated application of (3.2.3) yields the sequences

$$U=\{u_j\,:\,j=1,2,\ldots\} \quad \text{and} \quad V=\{v_j\,:\,j=1,2,\ldots\}.$$

Let us introduce the z-transforms of U and V by

$$\widehat{U}(z)=\sum_{j=1}^{\infty}\frac{u_j}{z^j}, \quad \text{and} \quad \hat{V}(z)=\sum_{j=1}^{\infty}\frac{v_j}{z^j} \qquad |z|\geq a>0$$

with a constant a such that the sums converge. Furthermore, we define a two element vector of sequences $(U,V)^T$ and a corresponding matrix-vector multiplication by

$$\mathrm{W}\left(\begin{array}{c} U \\ V \end{array}\right)=\left\{\mathrm{W}\left(\begin{array}{c} u \\ v \end{array}\right)_j\,:\,j=1,2,\ldots\right\}.$$

The product $\mathrm{W}\cdot\left(\widehat{U},\widehat{V}\right)^T$ is the conventional matrix-vector product. Obviously it holds

$$\begin{aligned} z\widehat{U}(z) &= u(0)+\sum_{j=1}^{\infty}\frac{u_{j+1}}{z^j} \\ z\widehat{V}(z) &= v(0)+\sum_{j=1}^{\infty}\frac{u_{j+1}}{z^j}, \end{aligned}$$

which allows us to write the z-transform of the left-hand side of (3.2.3) as

$$\begin{pmatrix} z\widehat{U}(z) - u(0) \\ z\widehat{V}(z) - v(0) \end{pmatrix}.$$

Consequently, we obtain the z-transform of the complete equation

$$(z - \mathrm{W}) \begin{pmatrix} \widehat{U}(z) \\ \widehat{V}(z) \end{pmatrix} = \begin{pmatrix} u(0) \\ v(0) \end{pmatrix}, \tag{3.2.4}$$

where here and in the following $z - \mathrm{W}$ denotes $z\mathbf{I} - \mathrm{W}$. Apparently, this is the direct analogue to the Laplace transform of the Helmholtz equation with constant coefficients. Obviously, the solution $\left(\widehat{U}(z), \widehat{V}(z)\right)^T$ has in general poles if z approaches the eigenvalues of W. We assign to each pole a special solution of the homogeneous equation, where one corresponds to an incoming and the other to an outgoing function. Our goal will be to find the conditions with respect to $(u(0), v(0))^T$ such that $\left(\widehat{U}(z), \widehat{V}(z)\right)^T$ has only one pole and the corresponding functions are outgoing functions. By means of the z-transform we will show that the sequence $(u, v)^T_{j=1,\ldots,\infty}$ has a representation

$$\begin{pmatrix} u \\ v \end{pmatrix}_j = C_1 \mathbf{q}^{(1)} t^j + C_2 \mathbf{q}^{(2)} \overline{t}^j \qquad \text{for all } j \geq 1 \tag{3.2.5}$$

with $C_1, C_2, t \in \mathbb{C}$ and $\mathbf{q}^{(1)}, \mathbf{q}^{(2)} \in \mathbb{C}^2$. In case of a vanishing periodic potential, that is $p = 0$, we have a representation:

$$\begin{pmatrix} u \\ v \end{pmatrix}_j = C_1 \begin{pmatrix} 1 \\ ik \end{pmatrix} e^{ikx_j} + C_2 \begin{pmatrix} 1 \\ -ik \end{pmatrix} e^{-ikx_j} \qquad \text{for all } j \geq 1. \tag{3.2.6}$$

Obviously t^j plays in the periodic case a role similar to $\exp(ikx_j)$ in the standard case. In fact, Lemma 3.2.3 at the end of this section, p. 45, shows that $|t| = 1$. Moreover, we will show that one term of (3.2.5) has positive power flux whereas the other must have a negative power flux. This corresponds exactly to the fact that the power flux P_+ which we can assign to the first term of (3.2.6) is $P_+ = k\,|C_1|^2 > 0$, whereas the power flux of the second term is $P_- = -k\,|C_2^2| < 0$. The representation (3.2.5) can be seen as a variant of Floquet's theorem. Additionally to the existence of at least one quasi-periodic solution $C\mathbf{q}t^j$ stated by Floquet's theorem, cf. e. g. **[16]**, Chapters 31 and 32, pp. 121, we obtained a complete representation formula due to exploitation of the special structure of the system (39).

The first idea to apply the pole condition technique to sequences is to require the holomorphic extension of $\widehat{U}(z)$ to the lower half of the complex plane. Of course, this is the right idea in the limiting case p small. For L sufficiently large, however, $u(x)$ may have a large phase increment over one period L. Therefore the simple correspondence: negative phase increment of $u(x) \longleftrightarrow$ singularity of $\widehat{U}(z)$ in the lower half of the complex plane does not hold true in general. As an example let us consider

$$u_j = e^{ij(2\pi-\alpha)} + e^{-ij(2\pi-\alpha)} = e^{-ij\alpha} + e^{ij\alpha}, \qquad j \geq 1$$

with an angle α, $0 < \alpha < \pi$. The sequence composed from u_j has a z-transform

$$\frac{e^{i(2\pi-\alpha)}}{z - e^{i(2\pi-\alpha)}} + \frac{e^{-i(2\pi-\alpha)}}{z - e^{-i(2\pi-\alpha)}}.$$

Hence, the first term which corresponds to a *positive* phase increment has its pole in the *lower* half plane, whereas the second term which corresponds to the *negative* phase increment has its pole in the *upper* half of the complex plane. Thus the situation known from the continuous 1D Helmholtz equation is reversed. Nevertheless, it turned out that there is another tool which allows us to decide which of the two terms of (3.2.5) corresponds to an *outgoing* wave. The important fact which is proved in Lemma 3.2.2 at the end of this section, p. 44, is that

$$\operatorname{Im}\left(\frac{\mathbf{q}^{(1)}(2)}{\mathbf{q}^{(1)}(1)}\right)\cdot\operatorname{Im}\left(\frac{\mathbf{q}^{(2)}(2)}{\mathbf{q}^{(2)}(1)}\right)<0 \tag{3.2.7}$$

holds true under general conditions. Consider, e. g. the case $\operatorname{Im}\left(\mathbf{q}^{(1)}(2)/\,\mathbf{q}^{(1)}(1)\right)>0$, consequently $\operatorname{Im}\left(\mathbf{q}^{(2)}(2)/\,\mathbf{q}^{(2)}(1)\right)<0$. If we set $C_2=0$ in (3.2.5), we have the simple series

$$\begin{pmatrix} u \\ v \end{pmatrix}_j = C_1 t^j \mathbf{q}^{(1)} \qquad \text{for all } j\geq 1. \tag{3.2.8}$$

It follows

$$\operatorname{Im}\left(\frac{v_j}{u_j}\right)=\operatorname{Im}\left(\frac{u_j'}{u_j}\right)=\operatorname{Im}\left(\frac{\mathbf{q}^{(1)}(2)}{\mathbf{q}^{(1)}(1)}\right)>0 \qquad \text{for all } j\geq 1.$$

Hence the *local* phase increment $\operatorname{Im}\left(u_j'/\,u_j\right)$ is positive at *all* sample points x_j and the sequence $(u,v)_{j=1,\ldots,\infty}^{T}$ is the sequence of sampled points of an *outgoing* function. Reversely, we characterize the sequence

$$\begin{pmatrix} u \\ v \end{pmatrix}_j = C_2 \overline{t}^j \mathbf{q}^{(2)} \qquad \text{for all } j\geq 1$$

with $\operatorname{Im}\left(\mathbf{q}^{(2)}(2)/\,\mathbf{q}^{(2)}(1)\right)<0$ as the sequence of sampled points of an *incoming* function. Note that together with a positive phase increment the power flux becomes positive. This follows trivially from

$$\operatorname{Im}\left(\frac{u_j'}{u_j}\frac{\overline{u_j}}{\overline{u_j}}\right)=\frac{P}{|u_j|^2}>0.$$

Assuming the property (3.2.7), we can give explicitely the Dirichlet-to-Neumann number $u'(0)/\,u(0)$ which guarantees that all local phase increments at the sample points are positive:

$$u'(0)/\,u(0)=\begin{cases} \dfrac{\mathbf{q}^{(1)}(2)}{\mathbf{q}^{(1)}(1)} & \text{if} \quad \operatorname{Im}\left(\dfrac{\mathbf{q}^{(1)}(2)}{\mathbf{q}^{(1)}(1)}\right)>0 \\ \dfrac{\mathbf{q}^{(2)}(2)}{\mathbf{q}^{(2)}(1)} & \text{if} \quad \operatorname{Im}\left(\dfrac{\mathbf{q}^{(1)}(2)}{\mathbf{q}^{(1)}(1)}\right)<0. \end{cases} \tag{3.2.9}$$

This is the main result with respect to our 1D example with periodic coefficients.

Having outlined the characterization of sequences, we want to show how to compute the quantities $\mathbf{q}^{(1)}$, $\mathbf{q}^{(2)}$ and t from our basic equation (3.2.4). The matrix W has a Schur decomposition with a unitary matrix $\mathrm{Q}^{(1)}\in\mathbb{C}^{2\times 2}$ such that

$$\mathrm{Q}^{(1)\mathrm{H}}\mathrm{W}\mathrm{Q}^{(1)}=\mathrm{T}^{(1)}=\mathrm{D}^{(1)}+\mathrm{N}^{(1)}$$

where $\mathrm{D}^{(1)}=\operatorname{diag}(\lambda_1,\lambda_2)$ and $\mathrm{N}^{(1)}\in\mathbb{C}^{2\times 2}$ is strictly upper triangular, cf. e. g. **[39**, Chap. 7, Theorem 7.1.3, p. 313**]**. To keep the representation simple, we consider only the case that the matrix W has complex eigenvalues with non-vanishing imaginary part. Since W is real, these eigenvalues are complex conjugates. Further,

because W is a 2×2 matrix, there exists a second unitary matrix $\mathrm{Q}^{(2)} \in \mathbb{C}^{2\times 2}$ such that

$$\mathrm{Q}^{(2)\mathrm{H}}\mathrm{W}\mathrm{Q}^{(2)} = \mathrm{T}^{(2)}$$

and

$$\mathrm{T}^{(1)} = \begin{pmatrix} t_{11} & t_{12} \\ & t_{22} \end{pmatrix}, \qquad \mathrm{T}^{(2)} = \begin{pmatrix} t_{22} & \widetilde{t}_{12} \\ & t_{11} \end{pmatrix},$$

that is, the main diagonal elements of $\mathrm{T}^{(1)}$ and $\mathrm{T}^{(2)}$ are reordered.This reordering can be obtained by the application of a complex Givens matrix U, cf. e.g. **[78]**, which realizes the transforms

$$\begin{array}{rcl} \mathrm{Q}^{(2)} & = & \mathrm{Q}^{(1)}\mathrm{U} \\ \mathrm{T}^{(2)} & = & \mathrm{U}^{\mathrm{H}}\mathrm{T}^{(1)}\mathrm{U}. \end{array} \tag{3.2.10}$$

Let us denote in the following the two column vectors of $\mathrm{Q}^{(1,2)}$ by $\mathbf{q}_1^{(1,2)}$and $\mathbf{q}_2^{(1,2)}$. A unitary transform of (3.2.4) yields

$$(z\mathrm{I} - \mathrm{T})\,\mathrm{Q}^{(1)\mathrm{H}} \begin{pmatrix} \widehat{U}(z) \\ \widehat{V}(z) \end{pmatrix} = \mathrm{Q}^{(1)\mathrm{H}} \begin{pmatrix} u(0) \\ v(0) \end{pmatrix}. \tag{3.2.11}$$

This equation has two singularities, namely for $z \in \{t_{11}, t_{22}\}$. Obviously, the solution of (3.2.11) has no pole in $z = t_{22}$ if

$$\begin{pmatrix} u(0) \\ v(0) \end{pmatrix} \in \operatorname{span}\left\{\mathbf{q}_1^{(1)}\right\}.$$

In this case, we obtain a first special solution to (3.2.11):

$$\begin{aligned} \mathrm{Q}^{(1)H} \begin{pmatrix} \widehat{U}(z) \\ \widehat{V}(z) \end{pmatrix}^{(1)} &= \left(z\mathrm{I} - \mathrm{T}^{(1)}\right)^{-1} \begin{pmatrix} C_1 \\ 0 \end{pmatrix} \\ &= \frac{1}{z}\left(\mathrm{I} + \frac{1}{z}\mathrm{T}^{(1)} + \frac{1}{z^2}\left(\mathrm{T}^{(1)}\right)^2 + \ldots\right)\begin{pmatrix} C_1 \\ 0 \end{pmatrix} \\ &= \frac{1}{z}\left(1 + \frac{1}{z}t_{11} + \frac{1}{z^2}t_{11}^2 + \ldots\right)\begin{pmatrix} C_1 \\ 0 \end{pmatrix} \end{aligned}$$

which yields

$$\begin{pmatrix} \widehat{U}(z) \\ \widehat{V}(z) \end{pmatrix}^{(1)} = \frac{C_1}{z}\left(1 + \frac{1}{z}t_{11} + \frac{1}{z^2}t_{11}^2 + \ldots\right).$$

Inverse z-transform supplies in fact the first term of (3.2.5). We may repeat the whole procedure replacing $\mathrm{Q}^{(1)}$ by $\mathrm{Q}^{(2)}$ and $\mathrm{T}^{(1)}$ by $\mathrm{T}^{(2)}$ to obtain

$$\begin{pmatrix} \widehat{U}(z) \\ \widehat{V}(z) \end{pmatrix}^{(2)} = \frac{C_2}{z}\left(1 + \frac{1}{z}t_{22} + \frac{1}{z^2}t_{22}^2 + \ldots\right).$$

Thus we have proved the representation formula (3.2.5).

REMARK 3.2.1. Following Ruhe **[78]** we can compute the unitary Givens matrix U using the ansatz

$$\mathrm{U} = \begin{pmatrix} \alpha & \overline{\beta} \\ -\beta & \overline{\alpha} \end{pmatrix}.$$

The coefficients are obtained as follows:

$$\begin{aligned}
(i) \quad |t_{11}-t_{22}| \le |t_{12}|: \quad \alpha &= \frac{1}{\sqrt{1+|\gamma|^2}} \\
\gamma &= \frac{t_{11}-t_{22}}{t_{12}} \\
\beta &= \gamma\alpha \\
(ii) \quad |t_{11}-t_{22}| > |t_{12}|: \quad \beta &= \frac{1}{\sqrt{1+|\gamma|^2}} \\
\gamma &= \frac{t_{12}}{t_{11}-t_{22}} \\
\alpha &= \gamma\beta.
\end{aligned}$$

To illustrate the properties arising from periodic media, we want to study, in brief, the reflection and transmission of the periodic potential given in Fig. 3.2.1 with $L = \pi/2$ and

$$\text{First example, low contrast}: \qquad p = \begin{cases} 0 & \text{for } 0 \le x < \frac{L}{2} \\ 0.05 & \text{for } \frac{L}{2} \le x < L \end{cases}, \qquad p(x+L) = x$$

as well as

$$\text{Second example, high contrast}: \qquad p = \begin{cases} 0 & \text{for } 0 \le x < \frac{L}{2} \\ 0.5 & \text{for } \frac{L}{2} \le x < L \end{cases}, \qquad p(x+L) = x$$

Corresponding to the layers, we set $k_1 = 1$, $k_2 = 1.05$ in the first and $k_1 = 1$, $k_2 = 1.5$ in the second example. To obtain the transmission matrix W, compare (3.2.3), we write the general solution of the Helmholtz equation in each of the two layers as

$$u_{1,2}(x) = c_1 e^{ik_{1,2}x} + c_2 e^{-ik_{1,2}x}. \tag{3.2.12}$$

Accordingly, we obtain the matrices

$$W_{1,2} = \begin{pmatrix} \cosh\left(ik_{1,2}\frac{L}{2}\right) & \frac{1}{ik_{1,2}}\sinh\left(ik_{1,2}\frac{L}{2}\right) \\ ik_{1,2}\sinh\left(ik_{1,2}\frac{L}{2}\right) & \cosh\left(ik_{1,2}\frac{L}{2}\right) \end{pmatrix}$$

supplying the composite matrix $\mathrm{W} = \mathrm{W}_1\mathrm{W}_2$. By (3.2.9) we compute the DtN-number $\mathrm{DtN} = u'(0)/u(0)$. Further, we define an amplitude reflection factor r with respect to the *first* layer

$$r \;:=\; \frac{c_2}{c_1} = \frac{ik_1 - \mathrm{DtN}}{ik_1 + \mathrm{DtN}}$$

where c_1 and c_2 are the constants from general representation (3.2.12). The reflection factor is the physical quantity of interest. Fig. 3.2.2 shows the computed reflection coefficients if the wavelength λ varies from $\pi/8$ to $\pi/8 + 2\pi$. The figures show the typical resonance effects of periodic layered media. The intervals where the reflection coefficient takes its maximum value, i. e. $r = 1$, have finite widths. In physics these intervals are called stopbands and are of particular interest for many applications. Among others, it can be seen from Fig. 3.2.2 that the larger the index contrast between the two layers, the larger the width of the stopband. These properties of 1D periodic layered media are well documented in the literature, cf. the book of März **[71]**, Chapters 2.2.4 and 7.2. The merit of our approach via the pole condition lies in the fact that no limiting process as e. g. number of layers $\to \infty$

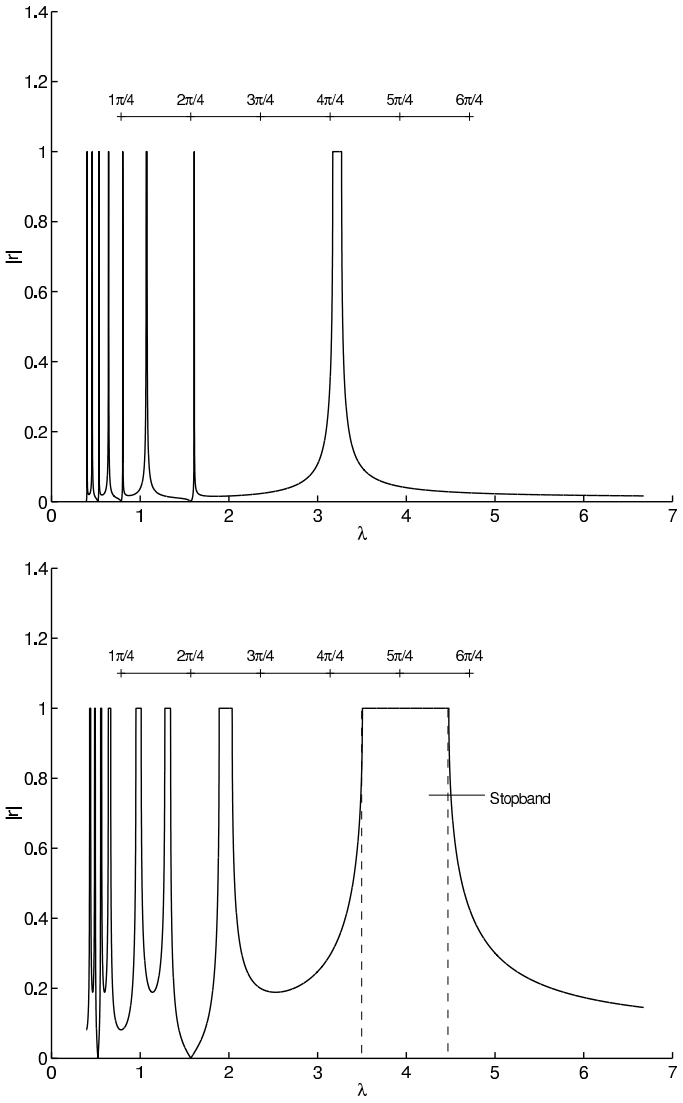


FIGURE 3.2.2. semi-infinite Bragg grating. Amplitude reflection coefficients vs. the wavelength for two different index contrasts. The upper figure refers to a layer stack with low index contrast ($k_1 = 2\pi/\lambda$, $k_2 = 1.05\,2\pi/\lambda$), the lower figure to a stack with high index contrast ($k_1 = 2\pi/\lambda$, $k_2 = 1.5\,2\pi/\lambda$)

has to be considered. Moreover, there are several ways to extend this to higher dimensions.

LEMMA 3.2.2. *Let the matrix* $\mathrm{W} \in \mathbb{R}^2$ *defined in (3.2.3) have a Schur transform* $\mathrm{Q}^{(1)H}\mathrm{W}\mathrm{Q}^{(1)} = \mathrm{T}^{(1)}$ *with a unitary matrix* $\mathrm{Q}^{(1)} := \left[\mathbf{q}_1^{(1)}, \mathbf{q}_2^{(1)}\right]$ *and an upper triangular matrix* $\mathrm{T}^{(1)}$ *such that* $\mathrm{Im}\,(t_{11}) \neq 0$. *Then there exists a second Schur transform* $\mathrm{Q}^{(2)H}\mathrm{W}\mathrm{Q}^{(2)} = \mathrm{T}^{(2)}$ *with* $\mathrm{Q}^{(1)} \neq \mathrm{Q}^{(2)} := \left[\mathbf{q}_1^{(2)}, \mathbf{q}_2^{(2)}\right]$ *and the following holds true:*

$$\mathrm{Im}\left(\frac{\mathbf{q}_1^{(1)}(2)}{\mathbf{q}_1^{(1)}(1)}\right) \cdot \mathrm{Im}\left(\frac{\mathbf{q}_1^{(2)}(2)}{\mathbf{q}_1^{(2)}(1)}\right) \leq 0.$$

In particular, if $w_{21} \neq 0$, then the expression becomes strictly negative.

PROOF. Since $\mathrm{Im}\,(t_{11}) \neq 0$ and W is real we have $t_{11} = \overline{t}_{22}$, hence $t_{11} \neq t_{22}$ and $t_{22} \neq w_{22}$. Thus the unitary matrix U defined in Remark 3.2.1 is different from the identity matrix. Hence it follows from the transform (3.2.10) that $\mathrm{Q}^{(1)} \neq \mathrm{Q}^{(2)}$, in particular $\mathbf{q}_1^{(1)} \neq \mathbf{q}_1^{(2)}$.

Taking the first column of each side of $\mathrm{Q}^{(1)H}\mathrm{W}\mathrm{Q}^{(1)} = \mathrm{T}^{(1)}$ we obtain

$$\mathrm{W}q_1^{(1)} = t_{11}q_1^{(1)}.$$

An evaluation in component-wise form yields

$$\frac{q_1^{(1)}(2)}{q_1^{(1)}(1)} = \frac{w_{21}}{t_{11} - w_{22}}.$$

Consequently we have

$$\mathrm{Im}\left(\frac{q_1^{(1)}(2)}{q_1^{(1)}(1)}\right) = \mathrm{Im}\left(\frac{w_{21}}{t_{11} - w_{22}}\right) \tag{3.2.13}$$

which is the first factor of the expression under investigation. We repeat the procedure with the second form of the Schur transform and obtain

$$\begin{aligned}
\mathrm{Im}\left(\frac{q_1^{(2)}(2)}{q_1^{(2)}(1)}\right) &= \mathrm{Im}\left(\frac{w_{21}}{t_{22} - w_{22}}\right) \\
&= \mathrm{Im}\left(\frac{w_{21}}{\overline{t}_{11} - w_{22}}\right) \\
&= \overline{\mathrm{Im}\left(\frac{w_{21}}{t_{11} - w_{22}}\right)} \\
&= -\mathrm{Im}\left(\frac{w_{21}}{t_{22} - w_{22}}\right). \qquad (3.2.14)
\end{aligned}$$

Computing the product of (3.2.13) and (3.2.14) we obtain finally the desired result

$$\mathrm{Im}\left(\frac{q_1^{(1)}(2)}{q_1^{(1)}(1)}\right) \cdot \mathrm{Im}\left(\frac{q_1^{(2)}(2)}{q_1^{(2)}(1)}\right) = -w_{21}^2 \left(\mathrm{Im}\left(\frac{1}{t_{22} - w_{22}}\right)\right)^2 \leq 0.$$

□

LEMMA 3.2.3. *Let the matrix* W *satisfy the same assumptions as in Lemma 3.2.2. Then the complex number t from the representation formula (3.2.5) satisfies $|t| = 1$.*

PROOF. The representation formula holds true for any initial data $(u(0), v(0))$. Let the initial data be chosen such that $C_2 = 0$ which means

$$\begin{pmatrix} u \\ v \end{pmatrix}_j = \begin{pmatrix} u \\ u' \end{pmatrix}_j = C_1 t^j \mathbf{q}^{(1)} \qquad j \geq 1.$$

From this we compute the power flux

$$\mathrm{Im}\,(\overline{u}u')_j = |C_1|^2\,|t|^{2j}\,\overline{\mathbf{q}_1}^{(1)}(1)\,q_1^{(1)}(2).$$

It follows form the conservation property $\mathrm{Im}\,(\overline{u}u')_j = \text{const}$ for all $j \geq 1$ that

$$\frac{\mathrm{Im}\,(\overline{u}u')_{j+1}}{\mathrm{Im}\,(\overline{u}u')_j} = |t|^2 = 1.$$

□

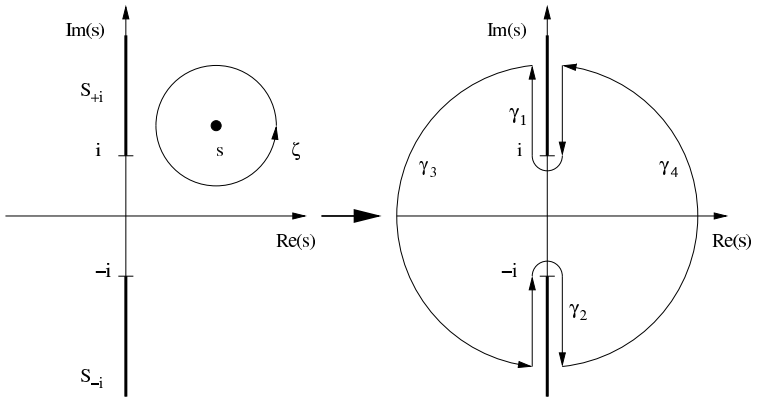


FIGURE 3.3.1. Definition of the cuts $S_{\pm i}$ and the contours $\gamma_1, \ldots, \gamma_4$. The contours around the cuts can be closed at infinity.

3.3. Radially Symmetric Helmholtz Equation in 2D

The following example is essentially for the understanding of the full 2D problem. Let us consider the complex function $\hat{u}(s)$

$$\hat{u}(s) := \frac{g(s)}{\sqrt{s^2+1}}. \tag{3.3.1}$$

Here, the square-root is understood as the square root on the principal plane, i. e. , for $a \in \mathbb{C}$ we define $\sqrt{a} = |a|^{1/2} e^{i\phi/2}, \phi = \arg(a), \phi \in (-\pi, \pi]$. With respect to the definition of $\hat{u}(s)$ this means, that we define the two branch cuts $S_{\pm i}$

$$S_{\pm i} = \{\pm i(1+\tau) : \tau \in \mathbb{R}_+\}. \tag{3.3.2}$$

In the more general cases, which we shall consider later, $\hat{u}$ is not given explicitely, but by differential or in the most general cases by integral equations. Then it will turn out to be more convenient to define the cuts parallel to the real axis, in order to construct the analytic continuation. For the present case, however, we will use the above given standard cuts.

The function $g(s)$ in the numerator of (3.3.1) is defined as follows: $g : \mathbb{C} \setminus (S_{+i} \cup S_{-i}) \to \mathbb{C}$, $g(s)$ is holomorphic in $\mathbb{C} \setminus (S_{+i} \cup S_{-i})$ and $\lim_{|s|\to\infty} |g(s)| = \text{const}$. The two singularities $\pm i$ and the infinite point are branch points of $\hat{u}(s)$ of order 1. The infinite point is a zero of $\hat{u}$, and the derivative $\hat{u}'(s)$ vanishes for $|s| \to \infty$, too. Hence, the infinite point is a regular point. In the following, we want to analyze the function $\hat{u}(s)$ in the same way as we analyzed the function $1/(s^2+1)$ occurring in the 1D example. The following lemma supplies the key for this investigation.

LEMMA 3.3.1. *The complex function $\hat{u}$ defined by (3.3.1) possesses a decomposition*

$$\hat{u}(s) = \frac{1}{2\pi i} \oint_{\gamma_1} \frac{g(\zeta)}{\sqrt{\zeta^2+1}\,(s-\zeta)}\, d\zeta + \frac{1}{2\pi i} \oint_{\gamma_2} \frac{g(\zeta)}{\sqrt{\zeta^2+1}\,(s-\zeta)}\, d\zeta,$$

where the paths γ_1 and γ_2 are contours in $\mathbb{C} \setminus (S_{+i} \cup S_{-i})$ enclosing the cuts $S_{\pm i}$ (see Fig. 3.3.1).

PROOF. For a small disc with radius ϵ we have, according to Cauchy's integral formula,

$$\hat{u}(s) = \frac{1}{2\pi i} \oint_{|\zeta - s| = \epsilon} \frac{\hat{u}(\zeta)}{\zeta - s} \, d\zeta \, .$$

Without to change the result we deform the path, as it is shown in Fig. 3.3.1, and obtain

$$\hat{u}(s) = \frac{1}{2\pi i} \oint_{\gamma} \frac{\hat{u}(\zeta)}{\zeta - s} \, d\zeta \, .$$

Here, $\gamma = \gamma_1 \cup \gamma_2 \cup \gamma_3 \cup \gamma_4$. Choosing for γ_3, γ_4 half-circles with radius r, we have only to show that the integrals along these paths vanish for $r \to \infty$. In polar coordinates r, ϕ, $\frac{\pi}{2} \leq \phi \leq \frac{3\pi}{2}$, we obtain on γ_3

$$\begin{aligned} \left| \int_{\gamma_3} \frac{\hat{u}(\zeta)}{\zeta - s} \, d\zeta \right| &\leq \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \left| \frac{|g(\zeta)|}{\sqrt{r^2 \mathrm{e}^{2i\phi} + 1} \, (r\mathrm{e}^{i\phi} - s)} i r \mathrm{e}^{i\phi} \right| \, d\phi \\ &\leq \int_{\frac{\pi}{2}}^{\frac{3\pi}{2}} \frac{r \, |g|}{\sqrt{|r^2 - 1|} |r - |s||} \, d\phi \, . \end{aligned}$$

Hence, the integral tends to zero for $r \to \infty$. The same applies to the integral on γ_4. □

We discuss two examples. First, we consider the more simpler treat case $g = 1$ on the entire complex plane. It will turn out that this corresponds to the Bessel function J_0. The second example deals with the Laplace transform of Bessel's equation, for which we will present a nearly analytic solution. We will show that a decomposition of the solution of the Laplace transformed Bessel equations according to Lemma 3.3.1 can be obtained.

We want to analyze the first term of the partial fraction decomposition of Lemma 3.3.1

$$\hat{u}_+(s) := \frac{1}{2\pi i} \oint_{\gamma_1} \frac{d\zeta}{\sqrt{\zeta^2 + 1} \, (s - \zeta)} \, .$$

We split the contour γ_1 into two parts $\gamma_1 = \gamma_1' + \gamma_1''$, where γ_1' lies to the left and γ_2'' to the right of the imaginary axis and $\zeta_1 \in \gamma_1'$ and $\zeta_2 \in \gamma_1''$

$$\hat{u}_+(s) = \frac{1}{2\pi i} \int_i^{i\infty} \frac{d\zeta_1}{\sqrt{\zeta_1^2 + 1} \, (s - \zeta_1)} + \frac{1}{2\pi i} \int_{i\infty}^{i} \frac{d\zeta_2}{\sqrt{\zeta_2^2 + 1} \, (s - \zeta_2)} \, .$$

Next, we let both contours get arbitrary close to the imaginary axis. Taking the definition of the square root into account and performing the limiting process $\zeta_{1,2} \to \zeta$, such that $\mathrm{Re}(\zeta_{1,2}) \to 0$, with

$$\begin{aligned} \zeta_1 &\to \zeta \qquad \text{and} \qquad \sqrt{\zeta_1^2 + 1} \to -\sqrt{\zeta^2 + 1} \\ \zeta_2 &\to \zeta \qquad \text{and} \qquad \sqrt{\zeta_2^2 + 1} \to \sqrt{\zeta^2 + 1} \end{aligned}$$

we obtain

$$\hat{u}_+(s) = -\frac{1}{\pi i} \int_i^{i\infty} \frac{d\zeta}{\sqrt{\zeta^2 + 1} \, (s - \zeta)} \, . \tag{3.3.3}$$

In order to find the standard form of the integral along the real axis, we compute

$$\hat{u}_+(s) = \frac{1}{\pi i} \int_{i - \infty}^{i - 0} \frac{d\zeta}{\sqrt{\zeta^2 + 1} (s - \zeta)} \tag{3.3.4}$$

$$= \frac{1}{\pi i} \int_0^{\infty} \frac{d\tau}{\sqrt{\tau^2 - 2i\tau} \, (s - (i - \tau))} \, . \tag{3.3.5}$$

In (3.3.4) we applied Cauchy's residue theorem in the left upper part of the complex plane and in (3.3.5) we introduced $\zeta \mapsto \tau = \zeta - i$, $\tau \in \mathbb{R}_+$ and reversed the direction of integration. Both equations (**??**) and (3.3.5) possess an interesting interpretation. Using $\zeta = i\tau$, applying $L^{-1}\left(\frac{1}{s+\tau}\right) = \exp(-\tau x)$ to (**??**), and reversing the order of integration we find

$$u_+(x) = \frac{1}{\pi}\int_1^\infty \frac{\mathrm{e}^{i\tau x}\, d\tau}{\sqrt{\tau^2-1}}, \tag{3.3.6}$$

i. e. , we can represent $u(x)$ as a superposition of *undamped* Fourier modes with positive frequencies (i. e. wavenumbers) $\tau \geq 1$. On the other hand, (3.3.5) allows a further representation of the same function $u(x)$, namely,

$$\begin{aligned} u_+(x) &= \frac{1}{i\pi}\int_0^\infty \frac{\mathrm{e}^{(i-\tau)x}\, d\tau}{\sqrt{\tau^2 - 2i\tau}}\,, \\ &= \frac{e^{i(x-\frac{\pi}{4})}}{\pi\sqrt{2x}}\int_0^\infty \frac{e^{-\tau}\, d\tau}{\sqrt{\tau}\,\sqrt{1-\frac{\tau}{2ix}}}\,. \end{aligned} \tag{3.3.7}$$

That is, we have equally valid $u(x)$ as a superposition of *damped* Fourier modes of a single frequency, due to the factor $\exp(ix)$, and with positive damping parts $\tau \geq 0$. Observe, that for $x \to \infty$ the integral converges to $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, which in turn supplies the well-known asymptotic formula for the Hankel function

$$u_+(x) = H_0^{(1)}(x) = \frac{e^{i(x-\pi/4)}}{\sqrt{2\pi x}}\left(1 + O\left(\frac{1}{x}\right)\right).$$

Differentiating (3.3.7) with respect to x and expanding the square-root term into a Taylor series supplies the asymptotic behavior

$$\frac{\partial_r u_+(r)}{u(r)} = i + O\left(\frac{1}{r}\right) \quad \text{for} \quad r \to \infty\,. \tag{3.3.8}$$

Obviously, (3.3.8) tells us that u_+ behaves asymptotically and locally like an outgoing plane wave.

REMARK. Each of the integral representations (3.3.6) and (3.3.7) coincides, up to normalization constants, to one of the two types of the classical integral representation formulas of Hankel's function. Referring to the book of Taylor, Eq. (3.3.6) corresponds to Eq. (6.32) in [**88**, Chap. 3.6, p. 230] and (3.3.7) corresponds to Eq. (6.33) [**88**, Chap. 3.6, p. 230].

Despite the simplicity of the above example, it contains all characteristic properties of the general Helmholtz-type scattering problem with variable coefficients, which we will analyze in Chap. 4. These basic properties are:

(1) The Laplace transformed solution contains two singularities. The singularities determine the far-field behavior of the function in the spatial domain. The singularity with positive imaginary part corresponds to an *outgoing* field, and the singularity with negative imaginary part corresponds to an *incoming* field. The "far-field" property of the Laplace transformed function, that is the behavior of $\widehat{u}$ for large $|s|$ determines the boundary condition at the artificial boundary.
(2) Different cuts connecting the singularities to infinity correspond to different forms of solution representations.

Bessel's equation. The radially symmetric Helmholtz equation in 2D separates into Bessel's equation for the distance variable r and an eigenvalue problem for the

angular variable, cf. Section 4.1, p. 59. The starting point of the classical derivation of transparent boundary conditions is always to find proper solutions of the Bessel equation

$$u''(r) + \frac{1}{r}u'(r) + \left(1 - \frac{\nu^2}{r^2}\right) u = 0\,, \quad r > \mathrm{a}\,,$$

with initial data $u(a)$ and $u'(a)$ exploiting the known properties of the Hankel functions at infinity. For the alternative derivation, we start again from the Bessel equation. We introduce a new coordinate system with the independent variable x obtained via $r \mapsto x := r - \mathrm{a}, x \geq 0$ and denote, for convenience, the shifted function $u(x + \mathrm{a})$ again by $u(x)$. This way, we obtain our basic differential equation

$$u''(x) + \frac{1}{x+\mathrm{a}}u'(x) + \left(1 - \frac{\nu^2}{(x+\mathrm{a})^2}\right) u = 0\,, \quad x > 0\,. \tag{3.3.9}$$

Now our concept is the following:

(1) Laplace transform of (3.3.9), which yields again a differential equation, this time in the dual domain.
(2) Solve the new differential equation by means of the variation of constant formula.
(3) Identify the singularities which correspond to incoming waves.
(4) Remove them by a proper choice of the initial conditions at $x = 0$.

In contrast to the classical approach, we can carry out the whole procedure without using properties of Hankel's functions.

Step 1. We multiply (3.3.9) with $x+\mathrm{a}$, introduce the new variable $v(x) = u(x)/(x+\mathrm{a})$, and take into account the rule $L(xu(x)) = -\hat{u}'(s)$, see A.1.2, where the prime denotes the complex derivative with respect to the argument s, to obtain

$$\frac{d}{ds}\left[\begin{pmatrix} s^2+1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix}\right] = \begin{pmatrix} \mathrm{a}(s^2+1)+s & -\nu^2 \\ -1 & \mathrm{a} \end{pmatrix}\begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix} - \mathrm{a}\begin{pmatrix} su(0)+u'(0) \\ 0 \end{pmatrix}.$$

Equivalently, we have

$$\frac{d}{ds}\begin{pmatrix} \widehat{u} \\ \widehat{v} \end{pmatrix} = \begin{pmatrix} a - \frac{s}{s^2+1} & -\frac{\nu^2}{s^2+1} \\ -1 & a \end{pmatrix}\begin{pmatrix} \widehat{u} \\ \widehat{v} \end{pmatrix} - a\begin{pmatrix} \frac{su(0)+u'(0)}{s^2+1} \\ 0 \end{pmatrix}. \tag{3.3.10}$$

In view of the solution procedure, it is convenient, to extract the translation part and to denote the remaining matrix with A

$$\begin{aligned} &\frac{d}{ds}\begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix} = \mathrm{a}\begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix} + \mathrm{A}\begin{pmatrix} \hat{u} \\ \hat{v} \end{pmatrix} - \mathrm{a}\begin{pmatrix} \frac{su(0)+u'(0)}{s^2+1} \\ 0 \end{pmatrix} \\ &\text{where } \mathrm{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and } \mathrm{A} = \begin{pmatrix} -\frac{s}{s^2+1} & -\frac{\nu^2}{s^2+1} \\ -1 & 0 \end{pmatrix} \end{aligned} \tag{3.3.11}$$

Step 2. Our goal is a solution formula of the type

$$\begin{pmatrix} \widehat{u}(s) \\ \widehat{v}(s) \end{pmatrix} = \Phi \int_s^\infty ds_1 \Phi^{-1} r\left(u(0), u'(0)\right). \tag{3.3.12}$$

Φ is the matrix of fundamental solutions and must obey the matrix differential equation

$$\Phi'(s) = \left(x_0 \mathbf{I} + \mathbf{A}(s)\right)\Phi(s)\,, \tag{3.3.13}$$

and $\mathbf{r}(u(0), u'(0))$ is the source vector of (3.3.11) which depends on the initial data. The determination of Φ evolves in three sub-steps: (i) remove the translation part, (ii) factorize the solution to drop the part which corresponds to $\nu = 0$, (iii)

transform the variable $s \to \tau$, to arrive at a problem with constant coefficients.
Step (i). We introduce

$$\Phi = \Phi^{(1)} \mathrm{e}^{\mathrm{a}s}\,,$$

which results in

$$\Phi^{(1)\prime}(s) = \mathbf{A}(s)\Phi^{(1)}(s)\,.$$

This way, a column $(u, v)^T$ of Φ, where we have dropped here the hat used so far to mark a Laplace transformed quantity, is transformed into a corresponding column of $\Phi^{(1)}$

$$\begin{pmatrix} u \\ v \end{pmatrix} \mapsto \begin{pmatrix} u^{(1)} \\ v^{(1)} \end{pmatrix}.$$

Step (ii). We factorize the first components of every column of $\Phi^{(1)}$

$$\begin{aligned} u^{(1)} &= u^{(2)} w, \quad w = \frac{1}{\sqrt{s^2+1}} \\ v^{(1)} &= v^{(2)}\,, \end{aligned}$$

which results in

$$\Phi^{(2)\prime}(s) = -\frac{1}{\sqrt{s^2+1}} \begin{pmatrix} 0 & \nu^2 \\ 1 & 0 \end{pmatrix} \Phi^{(2)}(s)$$

Step (iii). Finally, we transform $\Phi^{(2)}(s(\tau)) = \Phi^{(3)}(\tau)$ with

$$\frac{d\tau}{ds} = \frac{1}{\sqrt{s^2+1}}, \quad \tau = -\log\left(\sqrt{s^2+1} - s\right)$$

to obtain

$$\Phi^{(3)\prime}(s) = -\begin{pmatrix} 0 & \nu^2 \\ 1 & 0 \end{pmatrix} \Phi^{(3)}(s),$$

which in turn has a solution

$$\Phi^{(3)}(s) = \begin{pmatrix} e^{\nu\tau} & e^{-\nu\tau} \\ -\frac{1}{\nu} e^{\nu\tau} & \frac{1}{\nu} e^{-\nu\tau} \end{pmatrix}.$$

Tracing the whole procedure backwards, we find

$$\Phi^{(3)} \mapsto \Phi^{(2)} \mapsto \Phi^{(1)} \mapsto \Phi = \begin{pmatrix} \frac{(\sqrt{s^2+1}-s)^{-\nu}}{\sqrt{s^2+1}} & \frac{(\sqrt{s^2+1}-s)^{\nu}}{\sqrt{s^2+1}} \\ -\frac{1}{\nu}(\sqrt{s^2+1}-s)^{-\nu} & \frac{1}{\nu}(\sqrt{s^2+1}-s)^{\nu} \end{pmatrix} \mathrm{e}^{\mathrm{a}s}\,. \tag{3.3.14}$$

Applying (3.3.12), we obtain finally the desired representation of $\widehat{u}(s)$.

$$\begin{aligned} \hat{u}(s) = {} & \frac{\mathrm{a}}{2} \frac{(\sqrt{s^2+1}-s)^{-\nu}}{\sqrt{s^2+1}} \mathrm{e}^{\mathrm{a}s} \int_s^\infty \mathrm{e}^{-\mathrm{a}s_1} \frac{(\sqrt{{s_1}^2+1}-s_1)^{\nu}}{\sqrt{{s_1}^2+1}} (s_1 u(0) + u'(0))\, ds_1 + \\ & \frac{\mathrm{a}}{2} \frac{(\sqrt{s^2+1}-s)^{\nu}}{\sqrt{s^2+1}} \mathrm{e}^{\mathrm{a}s} \int_s^\infty \mathrm{e}^{-\mathrm{a}s_1} \frac{(\sqrt{{s_1}^2+1}-s_1)^{-\nu}}{\sqrt{{s_1}^2+1}} (s_1 u(0) + u'(0))\, ds_1 \end{aligned} \tag{3.3.15}$$

Step 3. According to (3.3.15), we can represent $\hat{u}(s)$ in factorized form

$$\hat{u}(s) = \frac{1}{\sqrt{s^2+1}} g(s)\,,$$

with

$$(3.3.16)\quad g(s) = \frac{\mathrm{a}}{2}(\sqrt{s^2+1}-s)^{-\nu}\mathrm{e}^{\mathrm{a}s}\int_s^\infty \mathrm{e}^{-\mathrm{a}s_1}\frac{(\sqrt{{s_1}^2+1}-s_1)^{\nu}}{\sqrt{{s_1}^2+1}}(s_1u(0)+u'(0))\,ds_1 + \frac{\mathrm{a}}{2}(\sqrt{s^2+1}-s)^{\nu}\mathrm{e}^{\mathrm{a}s}\int_s^\infty \mathrm{e}^{-\mathrm{a}s_1}\frac{(\sqrt{{s_1}^2+1}-s_1)^{-\nu}}{\sqrt{{s_1}^2+1}}(s_1u(0)+u'(0))\,ds_1\,.$$

Consequently, $\hat{u}(s)$ may have singularities at $s=\pm i$. In order to drop the singularity at $s=-i$, the following necessary condition must be fulfilled.

> Pole condition for the Bessel equation
> $$g(i)=0.$$

Step 4. This, in turn, supplies the necessary condition

$$\begin{aligned} u'(0) &= \mathrm{DtN}_\nu u(0) \\ \mathrm{DtN}_\nu &= -\frac{\int_{-i}^\infty ds\,\mathrm{e}^{-\mathrm{a}s}s\dfrac{i^{-\nu}(\sqrt{s^2+1}-s)^{\nu}+i^{\nu}(\sqrt{s^2+1}-s)^{-\nu}}{\sqrt{s^2+1}}}{\int_{-i}^\infty ds\,\mathrm{e}^{-\mathrm{a}s}\dfrac{i^{-\nu}(\sqrt{s^2+1}-s)^{\nu}+i^{\nu}(\sqrt{s^2+1}-s)^{-\nu}}{\sqrt{s^2+1}}}. \end{aligned}$$

In **[86]** we showed by estimating the integrals that the DtN-number has the following properties

(1) $\mathrm{DtN}_\nu \xrightarrow{\nu\to\infty} -\nu/\mathrm{a}$.
(2) $|\mathrm{DtN}_\nu/\nu| \le K < \infty$ for all ν.
(3) $\mathrm{Im}\,(\mathrm{DtN}_\nu) > 0$ for all $0 \le \nu < \infty$.
(4) $|\mathrm{DtN}_\nu| \ge K > 0$ for all $\nu \ge 0$.

Similar results were shown by Harari and Hughes **[55]** and Ihlenburg **[60**, Chapter 3.2.2**]** dealing with properties of Hankel's functions. In Chapter 4 we shall again discuss these spectral properties from the viewpoint of our general theory.

3.4. Generalizations

In the following we show, how to extend the above constructions to higher dimensions. Starting with a separable problem, which consists of a sequence of 1D problems, we generalize the concept to general situations.

Fields Outside a Sphere. In order to formulate the pole condition for dimensions $d>1$, we consider first fields outside a given sphere with radius a. There are two ways to generalize the concept. The first one uses a *separation* into functions depending on angular- and distance coordinates, the second uses a more general *parameterization* of the exterior domain. The first way is the more simpler one, the second way allows for a more general class of problems.

In the one-dimensional case we considered exterior functions $u(x)$, $x \ge a$. Now we study functions $u : \mathbb{R}^d \setminus \Omega_a \to \mathbb{C}$ with Ω_a the interior of a sphere with radius a. Let $Q : (a+R_+)\times S^{d-1} \to \mathbb{R}^d \setminus \Omega_a$ be an homeomorphism mapping generalized spherical coordinates to Euclidean coordinates outside the sphere. Let us write $(uQ)(\mathbf{x}) =: \widetilde{u}(\rho, \mathbf{x}_{S^{d-1}})$ and drop the tilde. Motivated by the diagonalization of

the Laplace-Beltrami operator on S^{d-1} based on the complete orthonormal system $\{(\phi_j, \lambda_j) : j = 0, \ldots, \infty\}$ of eigenvectors and eigenvalues, we consider first the case that u has a representation

$$u(\rho, \mathbf{x}_{S^{d-1}}) = \sum_{j=0}^{\infty} \phi_j(\mathbf{x}_{S^{d-1}})\, v_j(\rho), \quad \langle \phi_i, \phi_j \rangle_{L^2(S^{d-1})} = \delta_{ij}, \qquad \rho \geq a. \tag{3.4.1}$$

For $d = 2$ the functions ϕ_j are trigonometric monomials, for $d = 3$ spherical harmonics. This will be the starting point of our analysis in Section 4.2. Here we want to use it only for our qualitative discussion. Now, a natural generalization of the one-dimensional pole conditions to higher dimensions is

DEFINITION 3.4.1. (Pole condition for generalized Fourier modes on $\partial\Omega$.) Functions with a separable representation of type (3.4.1) satisfy the pole condition if each of the functions v_j , $j = 0, \ldots, \infty$, satisfies the one-dimensional pole condition with respect to the distance variable $(\rho - a)$.

However, we want to have a characterization of u independent oft the underlying partial differential equation. Therefore, in the general case, it does not make sense to use operator properties like the diagonalization above to define the pole condition. Therefore, we extend this to the case where no explicit series representation of u is given.

DEFINITION 3.4.2. (Pole condition for functions exterior to a sphere.) A function $u(\cdot, \mathbf{x}_{S^{d-1}})$ satisfies the pole condition if it satisfies the one-dimensional pole condition for all $\mathbf{x}_{S^{d-1}} \in S^{d-1}$.

This definition is motivated by the following observation. Let $d = 2$ and let us choose the ϕ_j from (3.4.2) with a very small support $\operatorname{meas}\operatorname{supp} v_j \to 0$. Then, the support of the corresponding function in $\mathbb{R}^2 \setminus \Omega_a$ is an angle segment with an interior angle $\to 0$. In the limiting case, the segment degenerates to a semi-infinite line. Our practical realization of the algorithms, however, follow a generalization of the first concept (3.4.1). Let us consider the case that u has a representation

$$u(\rho, \mathbf{x}_{S^{d-1}}) = \sum_{j=0}^{\infty} \psi_j(\mathbf{x}_{S^{d-1}})\, v_j(\rho), \qquad \rho \geq a \tag{3.4.2}$$

where the functions $\psi_j : S^{d-1} \to \mathbb{C}$ form a set of linear independent functions. Clearly, (3.4.2) is a generalization of our first approach (3.4.5). In the numerical algorithms, the functions ψ_j will play a role as trial functions resulting automatically from a finite-element discretization, whereas the functions v_j has to be computed according to the pole condition. We call this representation a local separable representation, since we assume a factorization with respect to distance and angular coordinates, but we don't require a global orthogonality of any kind of the functions ψ_j. Since the pole condition for representations of type (3.4.2) plays a dominant role for all of our the numerical concepts, we want to give it an extra definition

DEFINITION 3.4.3. (Pole condition for locally separable representations.) Functions with a separable representation of type (3.4.2) satisfy the pole condition if each of the functions v_j , $j = 0, \ldots, \infty$, satisfies the one-dimensional pole condition with respect to the distance variable $(\rho - a)$.

Our numerical algorithms for the general Helmholtz scattering problem with potentials depending both on the angular and the distance coordinates will be based on the representation (3.4.2) and pole condition 3.4.3.

General Case. We generalize the above given special forms of the pole condition and give it in a general form. To this end we assume the following

CONDITION 3.4.4. Let $\Omega \subset \mathbb{R}^d$ be the interior domain and $\partial\Omega$ its boundary. We assume that there exists a homeomorphism Q such that

(1) $Q : \mathbb{R}_+ \times \partial\Omega \to \mathbb{R}^d \setminus \Omega$.
(2) $\lim_{\xi\to\infty} |Q(\xi, \mathbf{x}_{\partial\Omega})| / \xi = C > 0$ for all $\mathbf{x}_{\partial\Omega} \in \partial\Omega$.

Item (1) supplies a generalized distance variable ξ, which plays the same role as the radius $\rho - a$ above. Since the interpretation of the distance variable with an radius is not always useful, we decided to use the symbol ξ (motivated by $|\mathbf{x}|$). Item (2) ensures that asymptotically the metric induces by Q is an Euclidean metric. By means of Q we can generalize the pole condition in a direct manner. Moreover, the general pole condition implies our understanding of outgoing functions.

DEFINITION 3.4.5. *General pole condition.* A function $u(\mathbf{x})$ satisfies the pole condition if the Laplace transform of $(uQ)(\cdot, \mathbf{x}_{\partial\Omega})$ has a holomorphic extension to the lower half-plane $\{\mathrm{s} \in \mathrm{C} : \mathrm{Im\,s} < 0\}$ for all $\mathbf{x}_{\partial\Omega} \in \partial\Omega$.

DEFINITION 3.4.6. A function $u(\mathbf{x})$ satisfying the pole condition Definition 3.4.5 is called outgoing.

Our definition of the general pole condition is based on a given parameterization Q. However, whether a function satisfies the pole condition or not does not depend on the special choice of Q. Using additional assumptions, we will show that if a function u satisfies the pole condition along a given path, it satisfies the pole condition in the neighborhood of this path, too. To discuss a particular example, we introduce to following notation:

- $\gamma_0 := \{(\xi, \eta_0), \xi \geq 0\}$ is a given path.
- $\widehat{u}_0(s)$ is the Laplace transform of $u_0(\xi, \eta_0) := (uQ)(\xi, \eta_0)$ along this path.
- $\gamma_1 := \{(\xi, \eta_0 + \delta\eta(\xi)), \xi \geq 0\}$ is a perturbation of γ_0, where $\delta\eta(\xi)$ denotes the perturbation in the angular variable in dependence of the distance variable ξ.
- $\widehat{u}_1(s)$ is the Laplace transform of $u_1(\xi, \eta_0 + \delta\eta) := (uQ)(\xi, \eta_0 + \delta\eta)$ along this path.

Our basic additional assumption is that the function $u(\xi, \eta)$ in the neighborhood of the path γ_0 has a representation $u(\xi, \eta) = f(\eta - \eta_0)u_0(\xi, \eta_0)$ and f has a convergent Taylor representation around 0.

LEMMA 3.4.7. *Let a parameterization $Q(\xi, \eta)$ be given. Let $\widehat{u}_0(s)$ satisfy the pole condition along the path γ_0. Let there be $\delta\eta$ and $u(\xi, \eta)$ such that*

(1) *$\delta\eta := p\left(\frac{1}{\xi + a}\right)$, where $p(t) = \sum_{m=1}^{\infty} p_m t^m$ with a convergence radius $\rho > 1/\widetilde{a}$, $0 < \widetilde{a} < a$ and*
(2) *$u(\xi, \eta) = \left(\sum_{j=0}^{\infty} c_j (\delta\eta)^j\right) u(\xi, \eta_0)$ absolutely and uniformly for $0 < |\eta - \eta_0| \leq \rho_\eta$ and $\xi \geq 0$.*

Then, $\widehat{u}_1(s)$ computed along the perturbed path γ_1 has a holomorphic extension to the lower half of the complex plane.

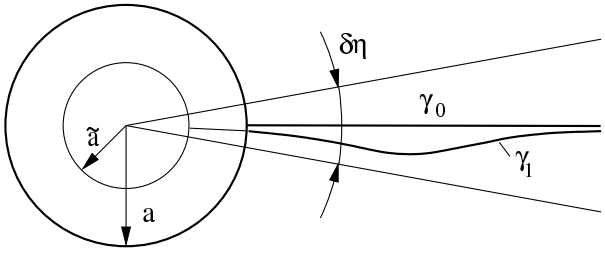


FIGURE 3.4.1. The path γ_0, on which the pole condition is satisfied and the disturbed path γ_1.

PROOF. We refer to Fig. 3.4.1. It holds

$$(\delta\eta)^j = \sum_{m=j}^{\infty} p_{m,j} \left(\frac{1}{\xi + a}\right)^m \qquad \text{for } \xi > \tilde{a} - a,\ j \geq 1$$

with real coefficients $p_{m,j}$ since the convergence radius of the product of two power series with convergence radius ρ is again ρ. Thus we get

$$\begin{aligned} u(\xi,\eta) &= u(\xi,\eta_0)\left(1 + \sum_{j=1}^{\infty} c_j \left(\sum_{m=j}^{\infty} p_{m,j}\left(\frac{1}{\xi+a}\right)^m\right)\right) \\ &= u(\xi,\eta_0) + u(\xi,\eta_0)\sum_{j=1}^{\infty}\left(\frac{1}{\xi+a}\right)^j \left(\sum_{m=1}^{j} c_j p_{j,m}\right) \\ &= u(\xi,\eta_0) + u(\xi,\eta_0)\sum_{j=1}^{\infty}\left(\frac{1}{\xi+a}\right)^j q_j. \qquad \text{for all } \xi \geq 0,\ 0 < |\eta - \eta_0| \leq \rho_\eta, \end{aligned}$$

where the coefficients q_j are defined by $q_j := \sum_{m=1}^{j} c_j p_{j,m}$. Since both series are absolute convergent, the reordering is legitimate and supplies a convergent power series. Thus we have a decomposition of function on the perturbed path into the function on the unperturbed path, which satisfies the pole condition and a remainder. The series is absolutely convergent for $\xi \geq \tilde{a} - a$, hence $\lim_{j\to\infty} |q_j| / \tilde{a}^j \to 0$ and therefore $|q_j| < \tilde{a}^j$ for $j \to \infty$. We apply the Laplace transform and find

$$\widehat{u}(s) = \widehat{u}_0(s) + \int_0^{\infty} \mathrm{d}\xi\, e^{-s\xi} u(\xi,\eta_0)\left(\sum_{j=1}^{\infty}\left(\frac{1}{\xi+a}\right)^j q_j.\right), \qquad \operatorname{Re} s > 0.$$

By the convolution lemma A.1.8 of the appendix this yields

$$\widehat{u}(s) = \widehat{u}_0(s) + \int_s^{\infty} \mathrm{d}s_1\, f(s_1 - s)\, \widehat{u}_0(s_1) \tag{3.4.3}$$

$$f(s) = e^{-as} \sum_{j=1}^{\infty} \frac{q_j}{(j-1)!} s^{j-1}. \tag{3.4.4}$$

We have to show that the integral expression in (3.4.3) is holomorphic in the lower half of the complex plane. To this end we study first the convergence of $g(s) := \sum_{j=1}^{\infty} q_j/(j-1)!\, s^{j-1}$. By the root criterion

$$\begin{aligned}\lim_{j\to\infty}\left(\frac{|q_j|}{(j-1)!}\,|s|^{j-1}\right)^{\frac{1}{j}} &\leq \lim_{j\to\infty}\left(\frac{\widetilde{a}^j}{(j-1)!}\,|s|^{j-1}\right)^{\frac{1}{j}}\\ &\leq a\lim_{j\to\infty}\frac{|s|^{-\frac{1}{j}}}{((j-1)!)^{\frac{1}{j}}}\\ &= 0.\end{aligned}$$

Thus the series converges on the entire complex plane. Next we have to show that the complex derivative of $\exp(-as)\ g(s)$, hence the complex derivative of $g(s)$ exists. We can compute $g'(s)$ term-wise, since

(i) $g(s)$ converges uniformely on the entire complex plane,
(ii) the derivatives of the individual terms of the series are continuous, and
(iii) the sum of the differentiated terms converges uniformely on the entire complex plane, which can be shown by the same technique as before.

Hence, the function $f(s)$ is holomorphic on the entire complex plane. Further, $g(s)$ is exponentially bounded, which follows from

$$\begin{aligned}|g(s)| &= \left|\sum_{j=1}^{\infty}\frac{q_j}{(j-1)!}s^{j-1}\right|\\ &\leq C\left|\sum_{j=1}^{\infty}\frac{1}{(j-1)!}\,|\widetilde{a}s|^{j-1}\right|\\ &= Ce^{|\widetilde{a}s|}\end{aligned}$$

with $C=\sup_{j\geq 1}\left|q_j/\widetilde{a}^{j-1}\right|$. The same holds for $|g'(s)|$ with $C=\sup_{j\geq 1}\left|q_{j+1}/\widetilde{a}^{j-1}\right|$. Thus it holds

$$\begin{aligned}|f(s)| &\leq Ce^{-a\operatorname{Re}s+|\widetilde{a}s|}\\ |f'(s)| &\leq Ce^{-a\operatorname{Re}s+|\widetilde{a}s|}.\end{aligned}$$

Using this we can show that (3.4.3) together with (3.4.4) is holomorphic in the lower half of the complex plane. A change of variables yields

$$\widehat{u}(s)=\widehat{u}_0(s)+\int_0^{\infty}\mathrm{d}s_1\,f(s_1)\,\widehat{u}_0(s_1+s)\,. \tag{3.4.5}$$

We prove that the complex derivative with respect to s exists in the lower half of the complex plane. The following holds true:

(i) for each fixed s_1 the complex derivative of $f(s_1)\,\widehat{u}_0(s_1+s)$ with respect to s exists,
(ii) for each fixed s the integral exists, due to the exponential bounds,
(iii) since $|\partial_s(f(s_1)\widehat{u}_0(s_1+s))|=|f(s_1)\,\partial_s\widehat{u}_0(s_1+s)|\leq C\,|f(s_1)|$ there exists a function $F(s_1):=Ce^{-a\operatorname{Re}s_1+|\widetilde{a}s_1|}$ integrable on the interval $\operatorname{Re}(s_1)\in[0,\infty)$, $\operatorname{Im}s_1<0$ such that $F(\operatorname{Re}s_1)\geq\left|(f(s_1)\widehat{u}_0(s_1+s))'\right|$ for all s in the lower half of the complex plane.

Hence we can apply Lebesgue's Dominated Convergence Theorem (cf. e. g. Zeidler **[95**, Appendix, p. 440**]**) which shows that the integral term can be differentiated

with respect to the parameter s and the result remains bounded. Thus the complex derivative of $\widehat{u}(s)$ exists in the entire lower half of the complex plane. □

Time-Dependent Schrödinger Equation. We want to extend the definition of outgoing waves to time-dependent problems, where we concentrate only to so-called abstract Schrödinger type problems

$$\begin{aligned} \partial_t u(\mathbf{x},t) &= -iAu(\mathbf{x},t), \qquad t>0 \\ u(\mathbf{x},0) &= u_0(\mathbf{x}) \end{aligned}$$

where the operator $A : D(A) \subset X \to X$ is an self-adjoint operator on the complex Hilbert space X. Prototypes of this equation are the Schrödinger equation and the Fresnel equation discussed in Section 1.1. In these examples we have $A : D(A) \subset L^2\left(\mathbb{R}^d\right) \to L^2\left(\mathbb{R}^d\right)$, and together with the symmetry of A we find immediately the conservation property $\partial_t (u,u)_{L^2(\mathbb{R}^d)} = 0$. In what follows, we will not study existence and uniqueness of the abstract evolution problem for operators A depending both on space and time, but we want to extend the pole condition to functions $(\mathbf{x},t) \mapsto u(\mathbf{x},t)$ and to apply this condition to derive discrete transparent boundary conditions satisfying discrete analogies to the continuous conservation property. To this end we give the obvious extension of the pole condition Definition3.4.5which refers only to spatial problemsto Schrödinger-type evolution problems.

> DEFINITION 3.4.8. A continuous, time-dependent function $u(x,t)$ satisfies the pole condition if it satisfies the time-ind ependent pole condition Definition3.4.5 for each fixed time $t \geq 0$.

REMARK. The Schrödinger equation with a self-adjoint operator A is time-reversible, hence it could be natural the extend this definition to the time interval $-\infty < t < \infty$. On the other hand, practical computations always start with initial data given at a fixed initial time. Thus it seems to be sufficient to restrict the definition to the half-interval.

CHAPTER 4

Existence and Uniqueness Statements for Separable Problems

We restate our basic Helmholtz scattering problem (1.0.1) in a slightly modified form. Since our focus is on the solution of the exterior problem and the coupling to the interior problem, we do not consider interior sources and boundaries. This is done to simplify the presentation and does not mean any restriction with respect to the theory. Our starting point is the system

$$
\begin{aligned}
\Delta u(\mathbf{x}) + k^2(\mathbf{x})\, u(\mathbf{x}) &= 0 && \text{in } \Omega_{\text{int}} \\
\Delta u_{\text{out}}(\mathbf{x}) + k^2(\mathbf{x})\, u_{\text{out}}(\mathbf{x}) &= 0 && \text{outside } \overline{\Omega}_{\text{int}} \\
u(\mathbf{x}) &= u_{\text{src}}(\mathbf{x}) + u_{\text{out}}(\mathbf{x}) && \text{on } \partial\Omega \\
\partial u(\mathbf{x}) &= \partial u_{\text{src}}(\mathbf{x}) + \partial u_{\text{out}}(\mathbf{x}) && \text{on } \partial\Omega \\
\partial_n u_{\text{out}}(\mathbf{x}) &= B u_{\text{out}}(\mathbf{x}) && \text{on } \partial\Omega.
\end{aligned} \tag{4.0.6}
$$

Introducing the sesquilinear forms

$$a \;:\; H^1(\Omega_{\text{int}}) \times H^1(\Omega_{\text{int}}) \to \mathbb{C} \tag{4.0.7}$$

$$b \;:\; H^{1/2}(\Omega_{\text{int}}) \times H^{1/2}(\Omega_{\text{int}}) \to \mathbb{C} \tag{4.0.8}$$

$$
\begin{aligned}
\text{by} \qquad a(v,u) &= (\nabla v, \nabla v) - \left(v, k^2(\mathbf{x}) u\right) \\
b(v,u) &= \left\langle v|_{\partial\Omega_{\text{int}}}, B\, u|_{\partial\Omega_{\text{int}}} \right\rangle,
\end{aligned}
$$

which are derived by a formal application of Green's theorem, we obtain the *interior* problem in weak form: Find $u \in H^1(\Omega_{\text{int}})$ such that for all $v \in H^1(\Omega_{\text{int}})$

$$a(v,u) - b(v,u) = \left\langle v|_{\partial\Omega_{\text{int}}}, \partial_n u_{\text{src}} \right\rangle - b(v, u_{\text{src}}). \tag{4.0.9}$$

Here the source data $u_{\text{src}}|_{\partial\Omega_{\text{int}}}$, $\partial_n u_{\text{src}}|_{\partial\Omega_{\text{int}}}$ must be given such that the corresponding functionals are continuous. In the sequel of this chapter we consider mainly the *exterior* problem and its coupling to the interior problem.

4.1. Separable Coordinates

Since we deal in the following only with the exterior domain Ω_{ext}, we suppress the subscript and write simply Ω instead of Ω_{ext}. Accordingly, we suppress the subscript *out* and write simply u instead of u_{out}. Further we write (4.0.6) explicitely in Cartesian coordinates

$$
\begin{aligned}
\Delta_{xy} u(x,y) + k^2 u(x,y) &= 0 && \text{in } \Omega_{xy} \\
\partial_n u(x,y) &= -B(x,y) u(x,y) && \text{on } \partial\Omega_{xy},
\end{aligned} \tag{4.1.1}
$$

where $\Delta_{xy} = \nabla_{xy}^T \nabla_{xy}$ with ∇_{xy} the column vector $(\partial_x, \partial_y)^T$ and Ω_{xy} and $\partial\Omega_{xy}$ are parameterizations of the exterior domain Ω and the the boundary $\partial\Omega$ with respect to Cartesian coordinates, that is

$$\begin{aligned}\Omega_{xy} &= \{(x,y)\in\mathbb{R}^2 : (x,y)\in\Omega\}\\ \partial\Omega_{xy} &= \{(x,y)\in\mathbb{R}^2 : (x,y)\in\partial\Omega\}\ .\end{aligned}$$

We want to transform the description based on Cartesian x,y-coordinates to other coordinates, denoted by ξ and η. To this end we introduce the C^1-diffeomorphism Q from the new domain

$$\Omega_{\xi\eta} = (\xi_0,\infty)\times[\eta_{\min},\eta_{\max})$$

onto the original domain Ω_{xy} by

$$\begin{aligned}Q : \Omega_{\xi\eta} &\to \Omega_{xy}\\ Q(\xi,\eta) &= \begin{pmatrix} x\\ y\end{pmatrix}.\end{aligned}$$

We consider the independent variable ξ with $0<\xi_0\le\xi$ as generalized distance variable and η with $\eta_{\min}\le\eta\le\eta_{\max}$ as generalized angular variable. Let Q satisfy Condition 3.4.4. We denote the Jacobi matrix Q' by

$$J := Q' = \begin{pmatrix} \partial_\xi x & \partial_\eta x\\ \partial_\xi y & \partial_\eta y\end{pmatrix}$$

and its determinant by $|J| = \partial(x,y)/\partial(\xi,\eta)$. Given a differentiable function f, it holds, by the definition of the gradient,

$$\nabla_{xy} f = J^{-T}\nabla_{\xi\eta} f,$$

where f on the left-hand side has to be read as $f(x,y)$ and f on the right-hand side has to be considered as function of ξ and η via $f(x(\xi,\eta), y(\xi,\eta))$. Considering the integral $\int_{\Omega_{xy}} g\Delta_{xy} f\,dx\,dy$ with an arbitrary twice differentiable function f and a test function $g\in C_0^\infty(\Omega_{xy})$ we find the transform of the Laplacian in terms of the the transformed gradient

$$\begin{aligned}\int_{\Omega_{xy}} g\Delta_{xy} f\,dx\,dy &= -\int_{\Omega_{xy}} (\nabla_{xy} g)^T \nabla_{xy} f\,dx\,dy\\ &= -\int_{\Omega_{\xi\eta}} \left(J^{-T}\nabla_{\xi\eta} g\right)^T J^{-T}\nabla_{\xi\eta} f\ |J|\,d\xi\,d\eta\\ &= \int_{\Omega_{xy}} g\,|J|^{-1}\nabla_{\xi\eta}^T\left(J^{-1}J^{-T}|J|\nabla_{\xi\eta}\right) f\,dx\,dy.\end{aligned}$$

The transformed Laplacian follows

$$\Delta_{xy} f = |J|^{-1}\nabla_{\xi\eta}^T\left(J^{-1}J^{-T}|J|\nabla_{\xi\eta}\right) f. \tag{4.1.2}$$

Next we restrict our consideration only to those transforms which can be derived from a conformal mapping.

REMARK. This does not mean that our consideration is based essentially on a conformal mapping and therefore mainly restricted to two-dimensional problems. We use this only as a convenient way to derive separated equations. For an example in higher space dimensions see Section 4.1.

Let us consider the coordinate transform Q as complex mapping $(x+iy) = Q(\xi+i\eta)$ with an analytic function Q. Consequently we have

$$d(x+iy) = Q' d(\xi+i\eta)$$

which supplies

$$J = \begin{pmatrix} \operatorname{Re} Q' & -\operatorname{Im} Q' \\ \operatorname{Im} Q' & \operatorname{Re} Q' \end{pmatrix},$$

from which we find the well-known property of conformal mapping

$$|J| = Q'\overline{Q'} \quad \text{and} \quad J^{-1}J^{-T}\,|J| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Consequently the Helmholtz equation in transformed coordinates based on conformal mapping reads

$$(\partial_\xi^2 + \partial_\eta^2)\, u + |J|\, k^2 u = 0. \tag{4.1.3}$$

DEFINITION 4.1.1. We call the Helmholtz equation separable with respect to ξ, η-coordinates, if the Helmholtz term $|J|\, k^2$ possesses a representation

$$|J(\xi,\eta)|\, k^2(\xi,\eta) = f_\xi + f_\eta$$

with functions $f_\xi$and f_η depending only on ξ and η, respectively.

If the transformed Helmholtz equation (4.1.3) is separable, we try in the usual way to find solutions of the form $u(\xi,\eta) = u_1(\xi)u_2(\eta)$ by a solution of the separated equations

$$\partial_\xi^2 u_1 + f_\xi u_1 - K u_1 = 0 \tag{4.1.4}$$

$$\partial_\eta^2 u_2 + f_\eta u_2 = -K u_2\,. \tag{4.1.5}$$

Since we will use in both cases under consideration, namely cylindric and elliptic-hyperbolic coordinates, the scaling $\xi \mapsto \widetilde{\xi} = \exp\xi$ we rewrite (4.1.4) using this transform. Depressing the tilde, we obtain

$$\xi^2\partial_\xi^2 u_1 + \xi\partial_\xi u_1 + f_\xi u_1 - K u_1 = 0\,.$$

The exponential re-scaling of the distance variable ξ introduced by the technique of the conformal mapping ensures – as we will see – the required asymptotic equivalence between a distance measured in Cartesian coordinates and a distance measured by the generalized distance variable ξ.

Cylindric Coordinates. We use the mapping $x + iy = Q(\xi + i\eta) := \exp(\xi + i\eta)$ with the angular variable in the interval $-\pi < \eta \le \pi$. We obtain $|J| = \exp(2\xi)$, and after the scaling $\xi \mapsto \widetilde{\xi} = \exp\xi$ and dropping the tilde, we have the separated equations

$$\xi^2\partial_\xi^2 u_1 + \xi\partial_\xi u_1 + (\xi^2 k^2(\xi) - K)\, u_1 = 0 \tag{4.1.6}$$

$$\partial_\eta^2 u_2 = -K u_2. \tag{4.1.7}$$

Obviously, the Helmholtz equation in cylindric coordinates separates for ξ-dependent wavenumbers. For our further consideration, especially for the derivation of the boundary conditions, we allow position-dependent wavenumbers of the type

$$k^2(\xi) = k_0^2 + \frac{k_{-1}}{\xi} + \ldots + \frac{k_{-n}}{\xi^n}\,. \tag{4.1.8}$$

Elliptic-Hyperbolic Coordinates. We use the mapping

$$x + iy = Q(\xi + i\eta) := \frac{q}{2}\left(\mathrm{e}^{\xi+i\eta} + \mathrm{e}^{-(\xi+i\eta)}\right)$$

with the angular variable in the interval $-\pi < \eta \le \pi$. We compute

$$|J| = \frac{q^2}{4}\left(\mathrm{e}^{2\xi} + \mathrm{e}^{-2\xi}\right) - \frac{q^2}{2}\cos 2\eta\,.$$

Allowing only for an position-independent wavenumber k, and and carrying out the exponential scaling as above, we obtain the separated equations

$$\xi^4\partial_\xi^2 u_1 + \xi^3\partial_\xi u_1 + \left(\frac{k^2q^2}{4}\left(\xi^4+1\right) - K\xi^2\right)u_1 = 0 \tag{4.1.9}$$

$$\partial_\eta^2 u_2 - \frac{k^2q^2}{2}\cos 2\eta\, u_2 = -Ku_2\,. \tag{4.1.10}$$

Generalization of the Potential. Both of the separated equations in cylindric and elliptic-hyperbolic coordinates, (4.1.6) and (4.1.9), possesses the common structure

$$\partial_\xi^2 u_1 + \xi^{-1}\partial_\xi u_1 + \left(\sum_{j=0}^{\infty} a_j\xi^{-j}\right)u_1 = 0 \tag{4.1.11}$$

$$\partial_\eta^2 u_2 + f(\eta)u_2 = -Ku_2\,, \tag{4.1.12}$$

where the function $f(\eta)$ is a smooth, bounded and periodic function in the angle-like variable η.

Thus, these examples show exemplarily that the transform between curvilinear coordinates results merely in a change of the potential, where terms of the form a_j/ξ^j, $j \geq 0$ are added. A similar statement has been obtained in Section 3.4, Lemma 3.4.7, where we considered perturbed paths. Therefore we will base our theory on potentials of the above form.

Generalization to Higher Space Dimensions. In Section 4.2 we will need the Laplacian in $\mathbb{R}^d$ exterior to a sphere. Therefore we add here a generalization of the conformal mapping used above to arbitrary space dimensions. We consider the map

$$\begin{aligned} Q &: \mathbb{R}_+ \times S^{d-1} \to \mathbb{R}^d \\ Q\left(\xi, \mathbf{x}^0\right) &= \mathbf{y}, \end{aligned}$$

with $\mathbf{y} = (y_1, \ldots, y_d)$ the Cartesian coordinates in $\mathbb{R}^d$. In generalization of the conformal mapping we assume the following:

(1) Two orthogonal intersecting curves remain orthogonal intersecting under the map Q.
(2) It holds the scaling $e^{2\xi} = \sum_{j=1}^d y_j^2$.
(3) Q has a representation $Q = e^\xi \widetilde{Q}$, where $\widetilde{Q} : S^{d-1} \to \mathbb{R}^{d\times d}$ depends only on the angular coordinates.

As in Section 4.1 we use the Jacobian $J = Q'$. Let $d\mathbf{s}_{1,2} = \left(d\xi, d\mathbf{x}^0\right)^T_{1,2}$ be two orthogonal vectors. With $d\mathbf{y}_{1,2} = J \cdot d\mathbf{s}_{1,2}$ and the orthogonality assumption: $\langle d\mathbf{y}_1, d\mathbf{y}_2\rangle = 0$ implies $\langle d\mathbf{s}_1, d\mathbf{s}_2\rangle = 0$, it follows that the matrix

$$J^T J = e^{2\xi}\begin{pmatrix} f & \\ & \Phi \end{pmatrix}, \qquad f : S^{d-1} \to \mathbb{R}, \quad \Phi :\ S^{d-1} \to \mathbb{R}^{(d-1)\times(d-1)}$$

is a diagonal matrix. Using further

$$|J| = e^{d\xi} g, \qquad \text{with } g : S^{d-1} \to \mathbb{R},$$

we compute the transformed Laplacian:

$$\begin{aligned}\Delta_{\xi,S^{d-1}}u &= |J|^{-1}\nabla^T_{\xi,S^{d-1}}\left(J^{-1}J^{-T}|J|\nabla_{\xi,S^{d-1}}\right)u \\ &= \frac{1}{ge^{d\xi}}\begin{pmatrix}\partial_\xi & \nabla_{S^{d-1}}\end{pmatrix} ge^{(d-2)\xi}\begin{pmatrix} f^{-1}\left(\mathbf{x}^0\right)\partial_\xi \\ \Phi^{-1}\left(\mathbf{x}^0\right)\nabla_{S^{d-1}}\end{pmatrix}u.\end{aligned}$$

As in the conformal mapping approach we let $r := e^{\xi}$, which results in

$$\begin{aligned}\Delta_{r,S^{d-1}}u &= \frac{f^{-1}\left(\mathbf{x}^0\right)}{r^{d-1}}\partial_r r^{d-1}\partial_r u + \frac{1}{r^2}\Delta_{S^{d-1}}u \\ \Delta_{S^{d-1}} &:= \frac{1}{g\left(\mathbf{x}^0\right)}\nabla^T_{S^{d-1}}g\left(\mathbf{x}^0\right)\Phi^{-1}\left(\mathbf{x}^0\right)\nabla_{S^{d-1}}, \qquad \mathbf{x}^0\in S^{d-1},\end{aligned}$$

where $\Delta_{S^{d-1}}$ denotes the Laplace-Beltrami operator on S^{d-1}. Using the second assumption above, the special choice $u = r^2$ shows $f = 1$, and we obtain finally the representation required in the following section

$$\Delta_{r,S^{d-1}}u = \frac{1}{r^{d-1}}\partial_r r^{d-1}\partial_r u + \frac{1}{r^2}\Delta_{S^{d-1}}u. \tag{4.1.13}$$

4.2. Preparation of Main Theory

Having outlined the structure of several example problems with potentials depending on the distance, we now want to analyze the scattering theory based on the Laplace transform and the pole condition on a model problem in arbitrary space dimensions. We shall supply the following key results:

(1) The pole condition is equivalent to Sommerfeld's radiation condition, at least for a finite number of modes. This will be proofed in Theorem 4.4.4.
(2) It generalizes Sommerfeld's radiation condition to problems with position dependent potentials (Theorem 4.4.4).
(3) It ensures existence and uniqueness of the interior scattering problem (Section 4.4.4 and 4.4.5).
(4) It supplies a new representation formula for the exterior solution of the scattering problem, which is not based on a Green's function (Theorem 4.4.1).
(5) The new representation formula supplies directly the far-field.
(6) It generalizes parts of the important theorems of Wilcox and Karp concerning the series representation of solutions in the exterior domain (Theorem 4.4.4).
(7) The pole condition supplies a constructive tool to design numerical algorithms. Three different forms of the pole condition are given in Section 4.4.2.

In a joint work with Hohage and Zschiedrich **[58]**, we present proofs of the two missing points:

(1) The pole condition is equivalent to Sommerfeld's radiation condition even in the infinite dimensional case.
(2) The boundary condition derived from the pole condition ensures existence and uniqueness of the exterior scattering problem.

To contain the following sections self-consistent, we start again from the Helmholtz equation for a model problem–simple enough to derive the desired properties–and introduce the corresponding variational formulation, regardless of the presentation given in the previous sections.

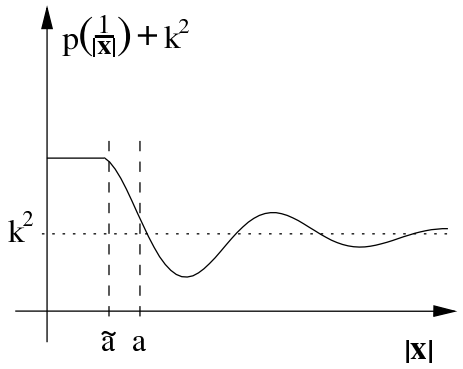


FIGURE 4.2.1. Potential of the model problem

Model Problem. We consider the Helmholtz problem (4.0.6) inside and outside the artificial domain $\Omega_a = \{\mathbf{x} \in \mathbb{R}^d : |\mathbf{x}| < a\}$[1] , $d \geq 1$, with boundary $\Gamma_a = \{\mathbf{x} \in \mathbb{R}^d : |\mathbf{x}| = a\}$ and a radial symmetric potential. Further we set $\Omega_{\text{int}} := \Omega_{\text{a}}$ and $\Omega_{\text{ext}} = \mathbb{R}^d \setminus \overline{\Omega_{\text{a}}}$. However, we underline that we can easily accommodate the setting to the more general situations described in the example section, e. g. to interior domains contained in Ω_{a} and to problems with interior boundaries and sources and arbitrary piecewise continuous potentials. We investigate the model problem

$$\Delta u + p\left(\tfrac{1}{|\mathbf{x}|}\right) u + k^2 u = 0 \qquad \text{in } \mathbb{R}^d \ . \tag{4.2.1}$$

For the potential $p : \left[0, \frac{1}{\widetilde{a}}\right) \to \mathbb{R}$ we assume a power series representation

$$p(t) = \sum_{j=2}^{\infty} p_j t^j \quad \text{with convergence radius } \frac{1}{\widetilde{a}} > \frac{1}{a} > 0,$$

where a is the above introduced radius of the artificial domain Ω_{a}, see Fig. 4.2.1. The auxiliary radius $\widetilde{a} < a$ is needed later to obtain convenient bounds with respect to Laplace transformed functions. Additionally we assume that u exterior to Ω_a can be composed from incoming and outgoing functions, that is $u(\mathbf{x}) = u_{\text{in}}(\mathbf{x}) + u_{\text{out}}(\mathbf{x})$ for $\mathbf{x} \in \Omega_{\text{ext}}$. Further we assume that both the incoming and the outgoing functions are solutions of the exterior Helmholtz equation. We consider the incoming field $u_{\text{in}}(\mathbf{x})$, $\mathbf{x} \in \Omega_{\text{ext}}$, as given and wish to compute the field $u(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$, such that the field $u - u_{\text{in}}$ is an outgoing field in the exterior domain.

Let us consider – as part of the full problem – the Helmholtz equation (4.2.1) for u_{out} on the exterior domain, where e. g. Neumann data on $\partial\Omega_{\text{a}}$ are imposed. This problem is *not* uniquely solvable. This can already be seen at the simplest situation $d = 1$ and the uniqueness of linear ODE's of second order. To obtain uniqueness of the exterior boundary value problem we need a second condition. As mentioned several times before, the established standard is to require the Sommerfeld radiation condition

$$\lim_{r \to \infty} r^{\frac{d-1}{2}} \left(\partial_r u_{\text{out}} - iku_{\text{out}}\right) = 0, \qquad r = |\mathbf{x}| \tag{4.2.2}$$

[1]Different to the previous sections we write in the following a instead of r_0 or ξ_0 since we will it use several times as index of different quantities.

uniformly for all directions $\frac{\mathbf{x}}{|\mathbf{x}|}$. It can be shown (cf. **[21]**) that for a non-vanishing solution u_{out} to (4.2.1) and (4.2.2) the energy flux is positive

$$\operatorname{Im} \int_{\partial\Omega_{\text{int}}} \overline{\partial_n u_{\text{out}}}\, u_{\text{out}} > 0.$$

This implies the uniqueness of the exterior solution, cf. **[60]**. In the following section we will repeat, in brief, how Sommerfeld's radiation condition, together with the positivity of the energy flux and properly defined Dirichlet-to-Neumann maps are used to prove existence and uniqueness of the interior scattering problem. This will direct our attention to the essential properties which we must derive from the pole condition.

Dirichlet-to-Neumann Map. We will study the interior problem by means of a variational formulation. To this end we multiply (4.2.1) by a test function $-\overline{v}$ and integrate over the interior of the domain Ω_a. A formal application of Green's theorem yields

$$\int_{\Omega_a} \mathrm{d}x \left(\nabla\overline{v}\nabla u - \left(p\left(\frac{1}{|\mathbf{x}|}\right) + k^2 \right) \overline{v} u \right) - \int_{\Gamma_a} \mathrm{d}s\, \overline{v}\partial_n u = 0 \tag{4.2.3}$$

where ∂_n is the unit normal derivative with a normal vector $\mathbf{n}$ pointing to the exterior of Ω_a. Next we introduce the so-called Dirichlet-to-Neumann map $B : H^{1/2}(\Gamma_a) \to H^{-1/2}(\Gamma_a)$ which maps Dirichlet data $u_{\text{out}}|_{\Gamma_a}$ of the outgoing solution satisfying (4.2.1) and (4.2.2) to its Neumann data $\partial_n u_{\text{out}}|_{\Gamma_a}$. Later we will consider solutions in the exterior of Ω_a based on the pole condition. For the moment, we concentrate on the solution of the problem interior to Ω_a. According to (4.2.3), we introduce the sesquilinear form

$$\begin{array}{lcl} a & : & H^1(\Omega_a) \times H^1(\Omega_a) \to \mathbb{C} \\ a(v,u) & := & \int_{\Omega_a} \mathrm{d}x \left(\nabla\overline{v}\nabla u - \left(p\left(\frac{1}{|\mathbf{x}|}\right) + k^2 \right) \overline{v} u \right) - \int_{\Gamma_a} \mathrm{d}s\, \overline{v} B u \end{array} \tag{4.2.4}$$

and the anti-linear functional

$$\begin{array}{lcl} F & : & H^1(\Omega_a) \to \mathbb{C} \\ F(v) & := & \int_{\partial\Omega_a} \mathrm{d}s\, \overline{v} \left(\partial_n u_{\text{src}} - B u_{\text{src}} \right), \end{array} \tag{4.2.5}$$

with given source data $\partial_n u_{\text{src}} \in H^{-1/2}(\Gamma_a)$ and $u_{\text{src}} \in H^{1/2}(\Gamma_a)$ so that the functional (4.2.5) is continuous. Now, the variational problem reads: Find $u \in H^1(\Omega_a)$ such that

$$a(v,u) = F(v) \qquad \text{for all } v \in H^1(\Omega_a). \tag{4.2.6}$$

The following theorem states the general existence and uniqueness result for the variational problem (4.2.6). Our further elaboration will not concentrate on the existence and uniqueness aspects, but it is important to know which properties the boundary operator has to satisfy.

THEOREM 4.2.1. *Let B be an operator with the following properties:*

(i) $B : H^{1/2}(\Gamma_a) \to H^{-1/2}(\Gamma_a)$ *is linear and bounded.*
(ii) *There exists a compact operator $\widetilde{B} : H^{1/2}(\Gamma_a) \to H^{-1/2}(\Gamma_a)$ such that* $\operatorname{Re} \int_{\Gamma_a} \mathrm{d}s \left(-B + \widetilde{B} \right) \overline{\phi}\phi \geq 0$ *for all $\phi \in H^{1/2}(\Gamma_a)$.*

(iii) $\operatorname{Im} \int_{\Gamma_a} \mathrm{d}s\, B\overline{\phi}\phi > 0$ *for all* $\phi \in H^{1/2}(\Gamma_a)$.

Then, the variational problem given by (4.2.4), (4.2.5) and (4.2.6) has a unique solution for u *for all right-hand sides* F*, and depends continuously on* F*.*

PROOF. Following **[21**, Theorem 5.7**]**, we repeat the standard arguments in brief and leave out many details.

Condition (i). Since $v \in H^1(\Omega_a)$, the restriction to the boundary Γ_a results in $v|_{\Gamma_a} \in H^{1/2}(\Gamma_a)$. Let $w := Bu$. Hence $w \in H^{-1/2}(\Gamma_a)$ and the integral $\int_{\Gamma_a} \mathrm{d}s\, \overline{v}w$ exists. Thus the condition guarantees that the sesquilinear form (4.2.4) is well defined.

Condition (ii). We use the operator $\widetilde{B}$ to decompose the sesquilinear form (4.2.4) into two parts $a(v,u) = a_0(v,u) + a_1(v,u)$:

$$\begin{aligned} a_{0,1} &: H^1(\Omega_a) \times H^1(\Omega_a) \to \mathbb{C} \\ a_0(v,u) &:= \int_{\Omega_a} \mathrm{d}x\, \nabla\overline{v}\nabla u + \int_{\Gamma_a} \mathrm{d}s\, \overline{v}\left(\widetilde{B} - B\right) u + \int_{\Omega_a} \mathrm{d}x\, \overline{v}u \\ a_1(v,u) &:= -\int_{\Omega_a} \mathrm{d}x\, \left(p\left(\frac{1}{|\mathbf{x}|}\right) + k^2\right) \overline{v}u - \int_{\Gamma_a} \mathrm{d}s\, \overline{v}\widetilde{B}u - \int_{\Omega_a} \mathrm{d}x\, \overline{v}u \end{aligned}$$

Clearly, $\operatorname{Re} a_0(u,u) \geq c_1 \|u\|^2_{H^1(\Omega_a)}$ holds true for some constant $c_1 > 0$. Further we see

$$\begin{aligned} \int_{\Omega_a} \mathrm{d}x\, \left(p\left(\frac{1}{|\mathbf{x}|}\right) + k^2\right) \overline{u}u &\leq \sup_{\mathbf{x}\in\Omega_a} \left|p\left(\frac{1}{|\mathbf{x}|}\right) + k^2\right| \int_{\Omega_a} \mathrm{d}x\, \overline{u}u \\ \text{and} \qquad \int_{\Gamma_a} \mathrm{d}s\, \overline{u}\widetilde{B}u &= \left\langle \operatorname{Tr} u, \widetilde{B}\operatorname{Tr} u \right\rangle_{L^2(\Gamma_a)}. \end{aligned}$$

Together this yields the Gårding inequality

$$\operatorname{Re} a(u,u) \geq c_1 \|u\|^2_{H^1(\Omega_a)} - c_2 \|u\|^2_{L^2(\Omega_a)} - \operatorname{Re} \left\langle \operatorname{Tr} u, \widetilde{B}\operatorname{Tr} u \right\rangle_{L^2(\Gamma_a)}$$

for all $u \in H^1(\Omega_a)$ with positive constants c_1, c_2 and $\operatorname{Tr} : H^1(\Omega_a) \to H^{1/2}(\Gamma_a)$ the trace operator. By Riesz' representation theorem, we can associate with the sesquilinear form $a(\cdot,\cdot)$ an operator $A : H^1(\Omega_a) \to H^1(\Omega_a)$. This operator can be split into the sum $A_0 + A_1$, where A_0 corresponds to $a_0(\cdot,\cdot)$ and A_1 to $a_1(\cdot,\cdot)$. The operator A_0 is linear, continuous and boundedly invertible on $H^1(\Omega_a)$. The latter fact follows form the Lax-Milgram theorem. The operator A_1 consists of two terms. The term corresponding to $\left\langle \operatorname{Tr} u, \widetilde{B}\operatorname{Tr} u \right\rangle_{L^2(\Gamma_a)}$ is compact, since $\operatorname{Tr}'\widetilde{B}\operatorname{Tr} : H^1(\Omega_a) \to H^1(\Omega_a)$ is compact. Further, the term corresponding to $\int_{\Omega_a} \mathrm{d}x\, \left(p\left(\frac{1}{|\mathbf{x}|}\right) + k^2\right) \overline{v}u$ is compact, since the embedding operator $H^1(\Omega_a) \hookrightarrow L^2(\Omega_a)$ is compact. Thus the operator $A_0 + A_1$ is of the form boundedly invertible plus compact operator, hence it is Fredholm of index zero.

Condition (iii). Let us consider the solution of the homogeneous operator equation $(A_0 + A_1)\, u = 0$, or equivalently, of the variational equation $a(v,u) = 0$ for all $v \in H^1(\Omega_a)$. Taking the imaginary part of (4.2.6) and using condition (iii) shows that u has vanishing Cauchy data on Γ_a. Hence, by virtue of the Cauchy-Kowalewskaja Theorem and elliptic regularity results, u must vanish everywhere. Since the homogeneous equation has only the trivial solution, the original equation has a unique solution. Moreover, the solution depends continuously on the right-hand side, that is the inverse operator $(A_0 + A_1)^{-1} : H^1(\Omega_a) \to H^1(\Omega_a)$ is continuous. □

Usually, the conditions (i)-(iii) are proved based on properties of the Hankel functions, see e.g. **[21, 60]**. This, however, works only for cases where a solution

representation based on Hankel functions is available. Our alternative approach presented below supplies these properties completely without using special functions. Moreover, it covers also cases with $p \neq 0$.

Laplace Transformed Helmholtz Equation. In order to simplify our analysis, we scale and shift the solution to obtain a new function

$$U\left(r, \mathbf{x}^0\right) := (r+a)^{\frac{d-1}{2}} u\left((r+a)\mathbf{x}^0\right) \tag{4.2.7}$$

with $r \geq 0$ and $\mathbf{x}^0 \in S^{d-1}$. This scaling realizes $\|U(r,\cdot)\|_{L^2(S^{d-1})} = \|u\|_{L^2((r+a)S^{d-1})}$. It is well-known for the case of a vanishing potential p, (see again e. g. **[21]**), that the Sommerfeld condition (4.2.2) implies the asymptotic formula

$$U\left(r, \mathbf{x}^0\right) = U_\infty\left(\mathbf{x}^0\right)\left(e^{ikr} + \mathcal{O}\left(\frac{1}{r}\right).\right)$$

We will show that the same asymptotic formula holds true for our case of a non-vanishing potential $p\left(\frac{1}{r}\right) \to 0$ as $r \to \infty$. Moreover, it is one of the most interesting aspects of our approach that it yields the far-field U_∞ directly both in theory and in numerical algorithms.

Let $\{(\phi_j, \lambda_j) : j \in \mathbb{N}\}$ be a complete orthonormal system of eigenfunctions and eigenvalues of the Laplace-Beltrami operator on S^{d-1}. In two dimensions the eigenfunctions are trigonometric functions, in three dimensions spherical harmonics. We denote by $U_j(r) := \int_{S^{d-1}} \mathrm{d}s\, \overline{\phi}_j U(r,\cdot)$ the Fourier coefficients with respect to the given orthonormal system. We want to study first the case that the solution U possesses the factorized representation consisting only of *one* Fourier mode

$$U\left(r, \mathbf{x}^0\right) = U_j(r)\phi_j\left(\mathbf{x}^0\right), \qquad \mathbf{x}^0 \in S^{d-1}.$$

We decompose the Laplacian in $\mathbb{R}^d$ into a radial and a corresponding angular part using the representation formula (4.1.13)

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{d-1}{r+a}\frac{\partial}{\partial r} + \frac{1}{(r+a)^2}\Delta_{S^{d-1}}.$$

Based on this decomposition, we reformulate the Helmholtz equation (4.2.1) in terms of the Fourier coefficients:

$$U_j''(r) + \left[k^2 + \frac{\frac{1}{4}(d-1)(3-d) + \lambda_j}{(r+a)^2} + p\left(\frac{1}{r+a}\right)\right] U_j(r) = 0. \tag{4.2.8}$$

Compared with Bessel's equation, there is no first order derivative, due to our scaling with the factor $(r+a)^{\frac{d-1}{2}}$.

Now we are ready to apply the Laplace transform (4.2.8). A direct application of the transformation formulas proved in the appendix supplies the key equations

$$\left(k^2 + s^2\right)\widehat{U}_j + \int_s^\infty \mathrm{d}s_1\, P_j(s_1 - s)\widehat{U}_j(s_1) \quad = \quad sU_j(0) + U_j'(0) \tag{4.2.9}$$

with

$$P_j(s) \quad := \quad p_*(s) + e^{-as}s\left(\frac{1}{4}(d-1)(3-d) + \lambda_j\right) \tag{4.2.10}$$

$$p_*(s) \quad := \quad e^{-as}\sum_{m=2}^{\infty} \frac{p_m}{(m-1)!}s^{m-1}. \tag{4.2.11}$$

The integral equation (4.2.9) is a linear Volterra integral equation of the second kind. It provides the basis of all our following considerations. Note, that the restriction of (4.2.11) to terms with index larger than two is of technical nature and not a restriction of the applicability of the method. In the following we will frequently use the quantity $P_j(s)$ without referring to a given mode. To this end we introduce

$$P(s) := p_*(s) + ce^{-as}s, \qquad P(0) = 0 \tag{4.2.12}$$

with a real constant c.

4.3. Main Theoretical Results

The integral equation (4.2.9) is the central object of our analysis. We shall derive from it all desired results. To simplify our investigation we set $k = 1$, which can always be obtained by a rescaling $r \mapsto kr$ and define the new function

$$w(s) := \widehat{U}(s)\left(s^2 + 1\right) \tag{4.3.1}$$

where $\widehat{U}$ stands for one of the modes $\widehat{U}_j$. With these notations (4.2.9) can be written as

$$w(s) + (Jw)(s) = sU(0) + U'(0) \tag{4.3.2}$$

with the integral operator

$$(Jw)(s) = \int_s^\infty \mathrm{d}s_1\, P(s - s_1)\frac{w(s_1)}{s_1^2 + 1} \tag{4.3.3}$$

and P is of the type (4.2.12). It is our first goal, to establish the solvability of the integral equation (4.3.2). To this end we introduce two cuts in the complex plane and define a metric appropriate for our purposes. Let the cuts be $S_{\pm i} := \{\pm i - t : t > 0\}$, see Fig. 4.3.1 and the domain $V := \mathbb{C} \setminus (S_+ \cup S_-)$. We want to analyze functions on V which are not continuous across the cuts. To deal with functions whose difference quotient, or Hölder difference quotient remains bounded for points with a small distance from each other, we define the following metric

$$d(s_1, s_2) = \sqrt{|s_1 - s_2|^2 + |\phi(s_1) - \phi(s_2)|^2}$$

with an auxiliary function $\phi : V \to \mathbb{R}$

$$\phi(s) := \begin{cases} -\mathrm{Re}\, s & \text{for } |\mathrm{Im}\, s| < 1,\ \mathrm{Re}\, s < 0 \\ 0 & \text{else.} \end{cases}$$

Fig. 4.3.1 explains the construction of the metric. If two points lie outside the strip between the cuts, their distance is the usual one. If one lies inside and one outside the area, there distance is the usual one plus the "distance" needed to surround the cut.

Let $\overline{V}$ denote the completion of V with respect to d. Then $\overline{V}$ is the union of all points of V with the set of points

$$s_\pm := \lim_{\epsilon \downarrow 0} s \pm i\epsilon, \text{ where } s \in S_{-i} \cup S_{+i}.$$

The curves s_+ and s_- are the "upper" and the "lower" side of the cuts. Now we can define a "jump" function $[w]$ which corresponds to a given continuous (with respect to d) function $w : \overline{V} \to \mathbb{C}$ by

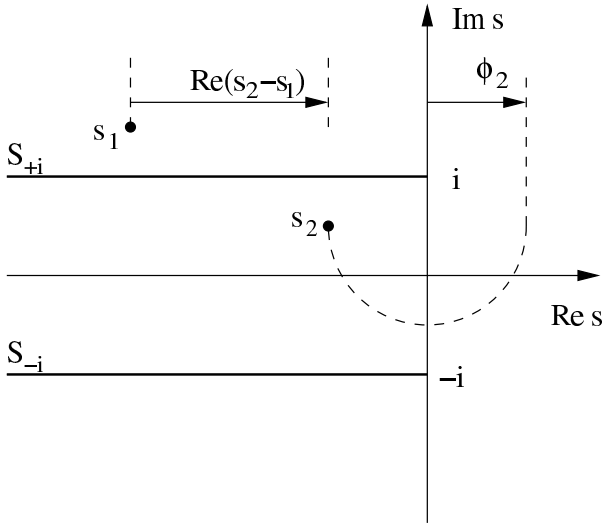


FIGURE 4.3.1. Definition of the cuts and construction of the metric d

$$
\begin{aligned}
[w] &: S_{-i} \cup S_{+i} \to \mathbb{C} \\
[w](s) &:= w(s_+) - w(s_-).
\end{aligned} \tag{4.3.4}
$$

If w is continuous on $\mathbb{C}$ (with respect to the usual distance in $\mathbb{C}$) so is $[w]$ on $S_{-i} \cup S_{+i}$ and $\lim_{s \to \pm i} [w] = 0$. We introduce the Banach space X with norm $\|\cdot\|_X$ by

$$
\begin{aligned}
X &= \{ w \in C(\overline{V}) : w \text{ holomorphic}, \\
&\quad w = o(|s|^2) \text{ uniformly for } |s| \to \infty \}
\end{aligned} \tag{4.3.5}
$$

$$
\|w\|_X = \sup_{s \in \overline{V}} \frac{|w(s)|}{1 + |s|^2}. \tag{4.3.6}
$$

Further, in order to analyze the solvability of the integral equation (4.3.3) we need to isolate the singularities at $s = \pm i$. To this end we denote by $|s|_1 := |\mathrm{Re}\, s| + |\mathrm{Im}\, s|$ and introduce the diamond shaped regions

$$
D_\pm = \left\{ s \in \overline{V} : |s \pm i|_1 < \frac{1}{2} \right\}.
$$

Now we can formulate the essential regularity results of the integral operator (4.3.3):

LEMMA 4.3.1. *Let* $0 < \alpha < 1$. *Then there exists a constant* c *such that for all* $w \in X$

$$
|(Jw)(s)| \leq c \left(\sup_{\mathrm{Re} s_1 \geq \mathrm{Re}\, s} \frac{|w(s_1)|}{1 + |s_1|^2} \right), \; s \in \overline{V} \setminus (D_+ \cup D_-) \tag{4.3.7}
$$

$$
|(Jw)'(s)| \leq c \left(\sup_{\mathrm{Re} s_1 \geq \mathrm{Re}\, s} \frac{|w(s_1)|}{1 + |s_1|^2} \right), \; s \in \overline{V} \setminus (D_+ \cup D_-) \tag{4.3.8}
$$

$$
|(Jw)(s) - (Jw)(\sigma)| \leq c d(s, \sigma)^\alpha \|w\|_X, \; s, \sigma \in (D_+ \cup D_-) \tag{4.3.9}
$$

$$
|(Jw)(s)| \leq c \|w\|_X, \; s \in \overline{V}. \tag{4.3.10}
$$

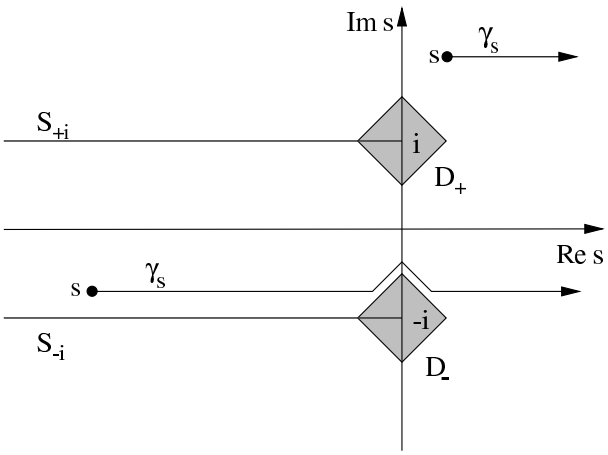


FIGURE 4.3.2. Paths γ_s outside the domains $D_\pm$

PROOF. We prove first the bound (4.3.7). The operator is an integral over a holomorphic function, hence we may deform the contour without changing the result. First we consider the rather simple case that $\mathrm{Re}\, s > 1/2$, that is s is on the right-hand side of $D_\pm$. With $s_1 = t + s$, $t \in \mathbb{R}_+$, we get from (4.3.3)

$$\begin{aligned} |(Jw)(s)| &= \left| \int_0^\infty \mathrm{d}t\, P(t)\, \frac{w(t+s)}{(t+s)^2+1} \right| \\ &\leq \left| \int_0^\infty \mathrm{d}t\, P(t) \right| \left(\sup_{\mathrm{Re}\, s_1 \geq \mathrm{Re}\, s} \frac{|w(s_1)|}{1+|s_1|^2} \right). \end{aligned}$$

Now we have from (4.2.12) and (4.2.11) and the convolution Lemma A.1.8

$$|P(t)| \leq c_1 e^{-a\,t+(a-\tilde{a})t} + c_2 t e^{-a\,t}.$$

Clearly, there exists a positive constant c such that

$$|P(t)| \leq c e^{-\tilde{a}\,t}, \qquad t \geq 0.$$

Thus $\left|\int_0^\infty \mathrm{d}t\, P(t)\right| \leq c/\tilde{a}$ and we have proved (4.3.7) for this particular choice of the starting point s. Now we extend this to the general situation $s \in \overline{V} \setminus (D_+ \cup D_-)$. For $\mathrm{Im}\, s > 0$ we use the contour parameterized by

$$\begin{aligned} \gamma_s(t) &:= s + t + i\psi_s(t) \\ \psi_s(t) &:= \begin{cases} \frac{1}{2} - |s+t-i|_1 & \text{if } |s+t-i|_1 \leq \frac{1}{2} \quad \text{and} \quad \mathrm{Im}\,(s+t-i) > 0 \\ -\frac{1}{2} + |s+t-i|_1 & \text{if } |s+t-i|_1 \leq \frac{1}{2} \quad \text{and} \quad \mathrm{Im}\,(s+t-i) \leq 0 \\ 0 & \text{else,} \end{cases} \end{aligned}$$

see Fig. (4.3.2). For $\mathrm{Im}\, s \leq 0$ the contour follows accordingly replacing $+i$ by $-i$.

This path has the following properties:

(1) It neither intersect the cuts $S_{\pm i}$ nor the domains $D_\pm$,
(2) meas $(\mathrm{supp}\, \psi_s) \leq 1$,
(3) $|\psi_s'| \leq 1$,
(4) $|\psi_s| \leq 1/2$,
(5) $\lim_{t\to\infty} \Psi_s(t) = 0$, and $\Psi_s(0) = 0$.

Using this path, which is an extension of the foregoing one, we write now

$$\begin{aligned}|(Jw)(s)| &= \left|\int_0^\infty \mathrm{dt}\,\frac{\mathrm{d}\gamma_s}{\mathrm{d}t}P(t+i\psi_s(t))\,\frac{w(t+i\psi_s(t)+s)}{(t+i\psi_s(t)+s)^2+1}\right| \\ &= \left|\int_0^\infty \mathrm{dt}\,\frac{\mathrm{d}\gamma_s}{\mathrm{d}t}P(t+i\psi_s(t))\,\frac{|t+i\psi_s(t)+s|^2+1}{(t+i\psi_s(t)+s)^2+1}\,\frac{w(t+i\psi_s(t)+s)}{|t+i\psi_s(t)+s|^2+1}\right|.\end{aligned}$$

We want to analyze the influence of $i\psi_s(t)$ in the argument of $P(t+i\psi_s(t))$. Our goal is to show that the deformation of the contour caused by Ψ_s is small enough to maintain an exponential bound with respect to $|P(s)|$. By the convolution lemma A.1.8 of the appendix we know $|P(\tau+i\sigma)| \le ce^{-a\tau+(a-\widetilde{a})|\tau+i\sigma|}$, $\sigma,\tau\in\mathbb{R}$. Let τ be the solution of

$$-a\tau+(a-\widetilde{a})\left|\tau+\frac{1}{2}i\right| = -\frac{\widetilde{a}}{2}\tau, \tag{4.3.11}$$

that is

$$\tau = \frac{|a-\widetilde{a}|}{2\sqrt{\left(a-\frac{\widetilde{a}}{2}\right)^2-(a-\widetilde{a})^2}}.$$

Since $|\Psi_s|\le 1/2$, equation (4.3.11) extends to the inequality

$$-at+(a-\widetilde{a})\,|t+i\Psi_s(t)| \le -\frac{\widetilde{a}}{2}t \qquad \text{for all t}\ge\tau.$$

Thus we obtain

$$|P(t+i\Psi_s(t))| \le ce^{-at+(a-\widetilde{a})|t+i\Psi_s(t)|} \le ce^{-\widetilde{a}/2\,t} \qquad \text{for t}\ge\tau.$$

Since P is continuous, this holds also true for $0\le t\le\tau$, possibly with a larger constant c:

$$|P(t+i\Psi_s(t))| \le ce^{-\widetilde{a}/2\,t} \qquad \text{for t}\ge 0.$$

Further, the quotients

$$\begin{aligned}\sup_{s\in V\setminus(D_+\cup D_-)}\sup_{t\ge 0}\left|\frac{|t+i\psi_s(t)+s|^2+1}{(t+i\psi_s(t)+s)^2+1}\right| &< \infty \\ \sup_{s\in V\setminus(D_+\cup D_-)}\sup_{t\ge 0}\left|\frac{(t+s)^2+1}{|t+i\psi_s(t)+s|^2+1}\right| &< \infty\end{aligned}$$

are bounded. Hence

$$\begin{aligned}|(Jw)(s)| &\le c\left(\int_0^\infty \mathrm{dt}\,e^{-\widetilde{a}/2\,t}\right)\left(\sup_{\mathrm{Re\,s_1}\ge\mathrm{Re\,s}}\frac{|w(s_1)|}{1+|s_1|^2}\right) \\ &= \frac{2c}{\widetilde{a}}\left(\sup_{\mathrm{Re\,s_1}\ge\mathrm{Re\,s}}\frac{|w(s_1)|}{1+|s_1|^2}\right).\end{aligned}$$

This is (4.3.7).

Next we want to prove (4.3.8). First we observe that

$$(Jw)'(s) = \int_s^\infty \mathrm{ds_1}\,P'(s_1-s)\,\frac{w(s_1)}{s_1^2+1}$$

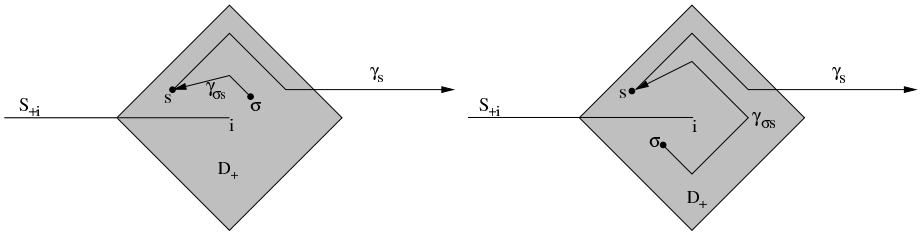


FIGURE 4.3.3. The two different choices of the paths inside the domains $D_{\pm}$

since by assumption $P(0) = 0$. Further, P' possesses the same exponential bound as P, by virtue of the convolution lemma A.1.8. Hence we can repeat the above procedure to obtain (4.3.8).

Now to the harder part. We want to show the Hölder continuity (4.3.9). Without loss of generality we may assume that $s, \sigma \in D_+$ and that $|s-i|_1 \geq |\sigma - i|_1$. We say that s and σ are on *opposite* sides (with respect to S_i) if both $\operatorname{Re} s < 0$, $\operatorname{Re} \sigma < 0$ *and* $(\operatorname{Im}(s-i))(\operatorname{Im}(\sigma - i)) < 0$ or if s and σ are limits of such points. First we assume that s and σ are not on opposite sides of S_i. We consider a path $\gamma_{\sigma,s}$ connecting the points σ and s. Let $\gamma_{\sigma,s}$ be the shortest path such that

$$|s_1 - i|_1 \geq |\sigma - i|_1 , \qquad \text{for all } s_1 \in \gamma_{\sigma s}.$$

Fig. 4.3.3 illustrates this situation.

The length of the path can be estimated by $l(\gamma_{\sigma s}) \leq 3\delta$, where δ is the distance between the points s and σ, $\delta = |s - \sigma|$. We extend the path $\gamma_{\sigma s}$ from s to $+\infty$ as shown in Fig . 4.3.3 by a path γ_s such that

$$|s_1 - i|_1 \geq |s - i|_1 , \qquad \text{for all } s_1 \in \gamma_s.$$

Next, we compute the difference

$$\begin{aligned}
(Jw)(s) - (Jw)(\sigma) &= \int_s^\infty \mathrm{d}s_1 \, P(s_1 - s) \frac{w(s_1)}{s_1^2 + 1} - \int_\sigma^\infty \mathrm{d}s_1 \, P(s_1 - \sigma) \frac{w(s_1)}{s_1^2 + 1} \\
&= -\int_\sigma^s \mathrm{d}s_1 \, P(s_1 - \sigma) \frac{w(s_1)}{s_1^2 + 1} \\
&\quad + \int_s^\infty \mathrm{d}s_1 \, (P(s_1 - s) - P(s_1 - \sigma)) \frac{w(s_1)}{s_1^2 + 1} \\
&= \int_{\gamma_{\sigma s}} \mathrm{d}s_1 \, P(s_1 - \sigma) \frac{w(s_1)}{s_1^2 + 1} \\
&\quad + \int_{\gamma_s} \mathrm{d}s_1 \, (P(s_1 - s) - P(s_1 - \sigma)) \frac{w(s_1)}{s_1^2 + 1}.
\end{aligned}$$

Using $|s| \leq |s|_1 \leq \sqrt{2}\,|s|$ we obtain for $s_1 \in \gamma_{\sigma,s}$

$$\begin{aligned}
\left| \frac{P(s_1 - \sigma)}{s_1^2 + 1} \right| &= \left| \frac{1}{s_1 + i} \right| \left| \frac{P(s_1 - \sigma)}{s_1 - i} \right| \\
&\leq \left| \frac{\sqrt{2}}{s_1 + i} \right| \frac{|P(s_1 - \sigma)|}{|s_1 - i|_1}.
\end{aligned}$$

We show that the second factor in the last inequality is bounded. First, $|s_1 - \sigma|$ goes faster to zero than $|s_1 - i|$ as $s_1 \to \sigma$, since by construction $|s_1 - i|_1 \geq |s_1 - \sigma|_1$.

Second, $|P(t)| = \mathcal{O}(|t|)$ as $|t| \to 0$, due to the properties (4.2.10) and (4.2.11). Hence we find

$$\left| \frac{P(s_1 - \sigma)}{s_1^2 + 1} \right| \leq c, \qquad \text{for all } s_1 \in \gamma_{\sigma,s}.$$

Together with the bound on $l(\gamma_{\sigma s})$ this yields for the first integral

$$\left| \int_{\gamma_{\sigma s}} \mathrm{d}s_1 \, P(s_1 - \sigma) \, \frac{w(s_1)}{s_1^2 + 1} \right| \leq c\delta \, \|w\|_X .$$

We estimate the second integral. We apply the mean value theorem

$$\begin{aligned} &\left| \int_{\gamma_s} \mathrm{d}s_1 \, \left(P(s_1 - s) - P(s_1 - \sigma) \right) \frac{w(s_1)}{s_1^2 + 1} \right| \\ &\quad \leq \quad \delta \int_{\gamma_s} |\mathrm{d}s_1| \sup_{z \in |[s,\sigma]|} |P'(s_1 - z)| \left| \frac{w(s_1)}{s_1^2 + 1} \right| . \end{aligned}$$

The derivative of $P(\cdot)$ is uncritical. We consider the term $1/\left(s_1^2 + 1\right)$. To bound the integrand inside of D_+ we take into account that due to the choice of γ_s we have

$$\begin{aligned} & & |s_1 - i|_1 &\geq |s_1 - s|_1 - |s - i|_1 && \text{inverse triangle inequality} \\ & & &\geq |s_1 - s|_1 - |s_1 - i|_1 \\ \text{and} & & 2\,|s_1 - i|_1 &\geq |s - i|_1 + |\sigma - i|_1 \\ & & &\geq |s - \sigma|_1 \\ & & &\geq \delta. \end{aligned}$$

The sum of these inequalities supplies

$$4\,|s_1 - i|_1 \geq |s_1 - s|_1 + \delta.$$

This implies that if $s_1 \to s$ there is a residual distance between s_1 and i larger than $\delta/4$. Together with $|s_1 - s|_1 \geq \mathrm{Re}\,(s_1 - s)$ and $|s_1 + i|_1 \geq 1$ (we are in D_+) we get

$$\begin{aligned} \left| s_1^2 + 1 \right| &\geq \frac{1}{2} \, |s_1 + i|_1 \, |s_1 - i|_1 \\ &\geq \frac{1}{8} \left(\mathrm{Re}\,(s_1 - s) + \delta \right). \end{aligned}$$

This is the result which we need to integrate the critical term $1/\left(s_1^2 + 1\right)$ inside D_+. Outside of D_+ the bound

$$\left| \frac{w(s_1)}{s_1^2 + 1} \right| \leq c \, \|w\|_X$$

holds true. With the bound on $P'(\cdot)$ from Lemma A.1.8 and $t^* := \sup\{t \geq 0 : s + t \in D_+\}$ we obtain

$$\begin{aligned}
&\left|\int_{\gamma_s} \mathrm{d}s_1 \left(P\left(s_1-s\right)-P\left(s_1-\sigma\right)\right)\frac{w(s_1)}{s_1^2+1}\right| \\
&\quad\leq\quad c\delta\left\|w\right\|_X\left(\int_0^{t^*}\mathrm{d}t\,\frac{1}{t+\delta}+\int_{t^*}^{\infty}\mathrm{d}t\,e^{-\widetilde{a}t/2}\right) \\
&\quad\leq\quad c\delta\left\|w\right\|_X\left(\ln\frac{t^*+\delta}{\delta}+\frac{\widetilde{a}}{2}\right) \\
&\quad\leq\quad c\delta^{\alpha}\left\|w\right\|_X.
\end{aligned}$$

Since $\delta\leq d(s,\sigma)$, we have proved (4.3.9) if the points s and σ not on opposite sides of S_i. The other case can be traced back to the above one using the triangle inequality and our last estimates

$$\begin{aligned}
\left|(Jw)(\sigma)-(Jw)(s)\right| &\leq \left|(Jw)(\sigma)-(Jw)(i)\right|+\left|(Jw)(i)-(Jw)(s)\right| \\
&\leq c\left(\left|\sigma-i\right|^{\alpha}+\left|i-s\right|^{\alpha}\right).
\end{aligned}$$

Without loss of generality we may assume that $\mathrm{Im}\,(\sigma-i)<0$ and $\mathrm{Im}\,(s-i)>0$, i.e. $d(\sigma,s)=\left|\mathrm{Re}\,\sigma\right|+\left|s-\sigma\right|$. Further, inserting the following inequalities in the last expression

$$\begin{aligned}
\left|\mathrm{Im}\,(\sigma-i)\right|+\left|\mathrm{Im}\,(i-s)\right| &= \left|\mathrm{Im}\,(\sigma-s)\right| \\
&\leq d(\sigma,s), \\
\left|\mathrm{Re}\,(s-i)\right| &= \left|\mathrm{Re}\,s\right| \\
&\leq \left|\mathrm{Re}\,\sigma\right|+\left|\mathrm{Re}\,(s-\sigma)\right| \\
&\leq d(s,\sigma), \\
\left|\mathrm{Re}\,(\sigma-i)\right| &= \mathrm{Re}\,\sigma \\
&\leq \left|d(s,\sigma)\right|,
\end{aligned}$$

we obtain the desired result (4.3.9). The last equation (4.3.10) follows directly from (4.3.9) and (4.3.7). □

The following is the essential prerequisite to prove the unique solvability of our integral equation (4.3.2).

LEMMA 4.3.2. *The integral operator J defined in (4.3.3) is a compact operator from X to X.*

PROOF. It follows from the definition of J in (4.3.3) that Jw is holomorphic for $w\in V$. Further, by the above lemma, equation (4.3.7), Jw remains bounded for $w\in\overline{V}$. Together with the global bound (4.3.10) from Lemma 4.3.1 we see that J is an operator $J:X\to X$. To prove compactness we use an iterated application of the Arzelá-Ascoli theorem. Let $(w_n)_{n\in\mathbb{N}}$ be a bounded sequence in X, $\left\|w_n\right\|_X\leq 1$ for all $n\in\mathbb{N}$. We have to show that the sequence $v_n=Jw_n$ has a convergent subsequence in X. First we consider the restriction of v_n to some compact subset $K\subset\overline{V}$. By Theorem 4.3.1, equations (4.3.9) and (4.3.10), the set of restricted functions $\{v_n|_K\}$ is equi-continuous on K with respect to the metric d. Hence, by the Arzelá-Ascoli theorem, there exists a subsequence of v_n which converges with respect to the maximum norm on K, $\left\|\phi\right\|_{\infty,K}=\sup_{s\in K}\left|\phi(s)\right|$. Next we introduce a sequence of sets K_j which become larger and larger. We define $K_j=\left\{s\in\overline{V}\,:\,|s|\leq j, j\in\mathbb{N}\right\}$. Our induction argument is the following. By our conclusion above, there exists a subsequence $\left(v_{n_1(l)}\right)_l$ which converges with respect to $\left\|\cdot\right\|_{\infty,K_1}$. Next we consider

K_2. Applying the same argument, we obtain a subsequence $\left(v_{n_2(l)}\right)_l$ which converges with respect to $\|\cdot\|_{\infty,K_2}$. Repeating this process of selecting subsequences, we arrive at an array $v_{n_j(l)}$ with the property that each row is a subsequence of the forgoing row. The diagonal subsequence $v_{n(l)} := v_{n_l(l)}$ converges to some function v with respect to the supremum norm on each K_j. In particular, $\lim_{l\to\infty} v_{n(l)}(s) = v(s)$ for all $s \in \overline{V}$. It remains to show that $\left\|v_{n(l)} - v\right\|_X \to 0$. Let $\epsilon > 0$. By virtue of the global bound Lemma 4.3.1 and equation (4.3.10) there exists a constant $C > 0$ such that $|v_{n(l)}(s)| \leq C$ for all $s \in \overline{V}$ and $l \in \mathbb{N}$. Therefore we find

$$\frac{|v(s) - v_{n(l)}(s)|}{1+|s|^2} \leq \epsilon, \qquad \text{for all} l \in \mathbb{N} \qquad \text{and } |s| \geq \sqrt{\frac{2C}{\epsilon}}.$$

Now let $J \geq \sqrt{2C/\epsilon}$,$J \in \mathbb{N}$. Since $v_{n(l)}$ converges to v with respect to $\|\cdot\|_{\infty,K_J}$, there exists $L \in \mathbb{N}$ such that

$$\sup_{s\in K_J} \frac{|v(s) - v_{n(l)}(s)|}{1+|s|^2} \leq \left\|v(s) - v_{n(l)}(s)\right\|_{\infty,K_J} \leq \epsilon$$

for $l \geq L$. Putting these inequalities together, we find

$$\left\|v(s) - v_{n(l)}(s)\right\|_X \leq \epsilon \qquad \text{for} l \geq L.$$

□

Now we are ready to state the first essential result.

THEOREM 4.3.3. *The integral equation (4.3.2) has a unique solution in X.*

PROOF. Since J is a compact operator from X onto X, the assertion follows from Riesz theory if we can show the uniqueness of the solution.Therefore we consider the homogeneous equation

$$w_0 + Jw_0 = 0$$

and want to show that necessarily $w_0 = 0$.

First we deduce that w_0 must be bounded. We find

$$\begin{aligned} |w_0(s)| &= |Jw_0(s)|, \qquad s \in \overline{V} \\ &\leq \frac{c}{a}\|w_0\|_X \\ &= \frac{c}{a}\sup_{s\in\overline{V}} \frac{|w_0(s)|}{1+|s|^2} \end{aligned}$$

by Lemma 4.3.1, equation (4.3.10), the norm definition (**??**) and the fact that $w_0 = o\left(|s|^2\right)$. It follows from Lemma 4.3.1, equation (4.3.7) that there exists a $s^* \in \overline{V} \setminus (D_+ \cup D_-)$ which specializes

$$|Jw_0(s)| \leq c\left(\sup_{\mathrm{Re}s_1 \geq \mathrm{Re}s} \frac{|w_0(s_1)|}{1+|s_1|^2}\right), s \in \overline{V} \setminus (D_+ \cup D_-)$$

to

$$|Jw_0(s^*)| \leq \frac{1}{4}\sup_{\mathrm{Re}s_1 \geq \mathrm{Re}s^*} |w_0(s_1)|, \qquad s^* \in \overline{V} \setminus (D_+ \cup D_-).$$

To compute the supremum we consider only the case $\mathrm{Re}\, s^* > 0$. We make $\mathrm{Re}\, s^*$ even larger, until the bound

$$\sup_{\operatorname{Re} s_1 \geq \operatorname{Re} s^*} |w(s_1)| \leq 2\,|w(s^*)|$$

holds. This must be possible, otherwise w cannot be a bounded function. Thus we obtain

$$\begin{aligned} |w_0(s^*)| &= |Jw_0(s^*)| \\ &\leq \frac{1}{2}\,|w(s^*)|\,. \end{aligned}$$

It follows that $w(s^*) = 0$. This implies $w(s) = 0$ for all s with $\operatorname{Re} s \geq \operatorname{Re} s^*$. Since w is holomorphic and continuous, it follows that $w(s) = 0$ for all $s \in \overline{V}$. □

Cut Functions. Now we want to study the above mentioned cut functions, using the definition of the jump functions (4.3.4) at p. 67,

$$\psi_\pm(t) := \frac{\left[\widehat{U}\right](\pm i - t)}{w(\pm i)}, \qquad t > 0. \tag{4.3.12}$$

The normalization with respect to the values $w(\pm i)$ allows an intuitive formulation of the theory. For the moment we assume $w(\pm i) \neq 0$, later we will see how to deal with the case $w(+i) = 0$ and the case $w(-i) = 0$. It is a remarkable result that we can compute the cut functions without knowing the values $U(0)$ and $U'(0)$, which define the right-hand side of the underlying integral equation (4.3.2).The introduction of the cut functions is of basic interest both for the theory and for the numerical algorithms for the following reasons:

(1) The cut functions show that we can always eliminate one singularity by a proper choice of the boundary conditions.
(2) It can be shown that, once one of the singularities is removed in one complex half-plane, the $\widehat{U}$ becomes a holomorphic function in the corresponding half plane.
(3) The cut functions allow the construction of an *explicit representation formula* for the spatial function in the exterior domain.
(4) They can be computed efficiently.
(5) Their computation supplies *automatically* the far field, without any other numerical effort.

LEMMA 4.3.4. *If $w(\pm i) \neq 0$, the cut functions $\psi_\pm$ defined by (4.3.12) for $t > 0$ satisfy the integral equations*

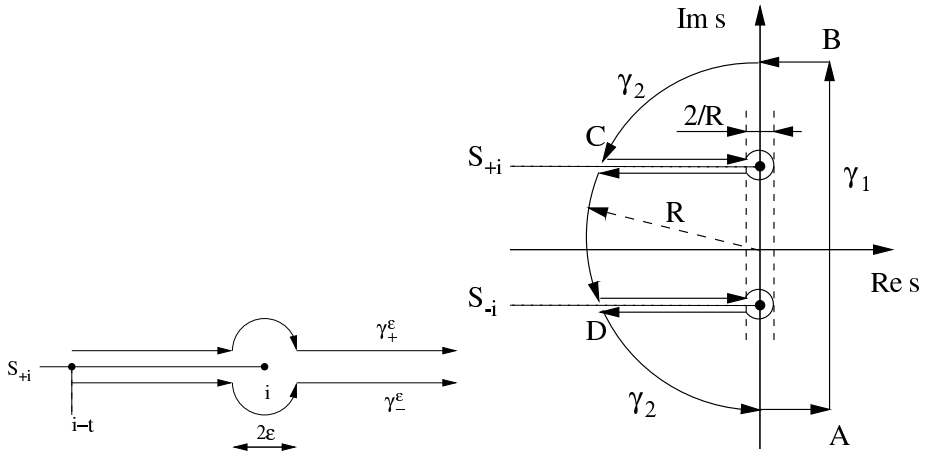
$$\psi_+(t) + \int_0^t \mathrm{d}t_1\, \frac{P(t-t_1)}{t(t-2i)}\psi_+(t_1) = \pi \frac{P(t)}{t(t-2i)}, \tag{4.3.13}$$

$$\psi_-(t) + \int_0^t \mathrm{d}t_1\, \frac{P(t-t_1)}{t(t+2i)}\psi_-(t_1) = \pi \frac{P(t)}{t(t+2i)}. \tag{4.3.14}$$

Here $P(\cdot)$ denotes the infinite series defined in (4.2.10), with dropped index j.

PROOF. We prove only the first of these equations since the proof of the second is similar. First we write our basic integral equation (4.3.2) with w replaced according to (4.3.1)

$$(s_\pm^2 + 1)\,\widehat{U}(s_\pm) + \int_{\gamma_\pm^\epsilon} \mathrm{d}s_1\, P(s_1 - s_\pm)\,\widehat{U}(s_1) = s_\pm U(0) + U'(0)$$



(a) Paths for the derivation of the cut functions

(b) Contours for the drivation of the representation formulas

FIGURE 4.3.4.

for $s = i - t \in S_i$ and $\epsilon > 0$. The paths $\gamma_\pm^\epsilon$ are shown in Fig. 4.3.4. Subtracting the equation with the minus sign from the one with the plus sign yields

$$t(t-2i)\left[\widehat{U}\right](i-t) + \int_{-t}^{-\epsilon} \mathrm{d}t\, P(t+t_1)\left[\widehat{U}\right](i+t_1) + \int_0^{2\pi} \mathrm{d}\phi\, (i\epsilon)e^{-i\phi} P\left(t-\epsilon e^{-i\phi}\right) \frac{w\left(i-\epsilon e^{-i\phi}\right)}{\left(2i-\epsilon e^{-i\phi}\right)\left(-\epsilon e^{-i\phi}\right)} = 0.$$

Carrying out the limit $\epsilon \to 0$ and dividing by $w(i)t(t-2i)$ yields (4.3.13). □

Referring to the book of Kreß **[66**, Theorem 10.15**]** we know that the Volterra integral equations for the cut functions (4.3.13), (4.3.14) have a unique solution, since both the kernel and the right-hand sides are bounded. This is, of course what we expect, since our basic integral equation (4.3.2) is uniquely solvable. Next we will derive essential properties of the cut functions $\psi_\pm$ near 0 and ∞. We will need these properties for the characterization of the near and the far-field of the corresponding spatial solution. Since $\psi_-(t) = \overline{\psi_+(t)}$, we give the first lemma only for ψ_+.

LEMMA 4.3.5. *The cut function ψ_+ defined in (4.3.12) belongs to $C^\infty(\mathbb{R}_+)$. The derivatives of ψ_+ at 0 can be computed recursively as follows:*

$$\psi_+(0) = \pi \lim_{t\to 0} \frac{P(t)}{t(t-2i)} \tag{4.3.15}$$

$$\psi_+^{(k+1)}(0) = \pi \lim_{t\to 0} \left(\frac{P(t)}{t(t-2i)}\right)^{(k+1)} + \frac{(k+1)!}{2i} \sum_{j=1}^{k+1} \frac{1}{(2i)^{k+1-j}(j+1)!} \sum_{n=1}^{j} P^{(n)}(0)\psi_+^{(j-n)}(0). \tag{4.3.16}$$

PROOF. Introducing the action of the integral operator by $(K\psi_+)(t)$, we write the Volterra integral equation (4.3.13)

$$\begin{aligned} \psi_+(t) + (K\psi_+)(t) &= \pi \frac{P(t)}{t(t-2i)}, \qquad t > 0 \\ (K\psi_+)(t) &:= \int_0^t \mathrm{d}t_1 \, \frac{P(t-t_1)}{t(t-2i)} \psi_+(t_1). \end{aligned} \tag{4.3.17}$$

By repeated partial integration we obtain

$$\begin{aligned} \int_0^t \mathrm{d}t_1 \, P(t-t_1) \, \frac{t^j}{j!} &= \sum_{l=1}^{\infty} P^{(l)}(0) \int_0^t \mathrm{d}t_1 \, \frac{(t-t_1)^l}{l!} \frac{t^j}{j!} \\ &= \sum_{l=1}^{\infty} P^{(l)}(0) \, \frac{t^{l+j+1}}{(l+j+1)!}. \end{aligned}$$

The change of the order of integration and summation is justified, since the Taylor expansion of $P(\cdot)$ converges uniformly around 0. The last term is an analytic function in t. Next, suppose that a given function $v(t)$ is a polynomial

$$v(t) := \sum_{j=0}^{\infty} \frac{v^{(j)}(0)}{j!} t^j.$$

Then, we compute

$$\begin{aligned} (Kv)(t) &= \frac{1}{t(t-2i)} \int_0^t \mathrm{d}t_1 \, P(t-t_1) \sum_{j=0}^{\infty} \frac{v^{(j)}(0)}{j!} t^j \\ &= \frac{1}{t-2i} \sum_{l=1}^{\infty} \sum_{j=0}^{\infty} P^{(l)}(0) v^{(j)}(0) \, \frac{t^{l+j}}{(l+j+1)!} \\ &= -\frac{1}{2i} \sum_{r=0}^{\infty} \left(\frac{t}{2i} \right)^r \sum_{m=1}^{\infty} \frac{t^m}{(m+1)!} \sum_{n=1}^{m} P^{(n)}(0) v^{(m-n)}(0) \\ &= -\frac{1}{2i} \sum_{l=0}^{\infty} t^l \sum_{j=1}^{l} \left(\frac{1}{2i} \right)^{l-j} \frac{1}{(j+1)!} \sum_{n=1}^{j} P^{(n)}(0) v^{(j-n)}(0). \end{aligned} \tag{4.3.18}$$

In particular, $(Kv)(t)$ is analytic at $t = 0$.

We prove by induction on n that $\psi_+ \in C^n([0,\infty))$ for $n \in \mathbb{N}_0$. We start with $n = 0$. Note that the right-hand side of (**??**) can be continuously extended to a function in $C^\infty([0,\infty))$, due to the fact $P(0) = 0$. First we consider ψ_+ on the interval $[0, t^*]$ and show that there exists $t^* > 0$ such that $\|K\| < 1$, where the integral operator K is restricted accordingly to $K : C([0,t^*]) \to C([0,t^*])$. The definition of $P(\cdot)$ implies

$$P(t) = \mathcal{O}(t) \qquad (t \to 0).$$

Based on the definition of K in (**??**) it follows

$$\begin{aligned}
|(K\psi_+)(t^*)| &= \left| \int_0^{t^*} \mathrm{d}t_1 \frac{P(t^*-t_1)}{t^*(t^*-2i)} \psi_+(t_1) \right| \\
&\leq C \left| \int_0^{t^*} \mathrm{d}t_1 \frac{t^*-t_1}{t^*(t^*-2i)} \psi_+(t_1) \right| \\
&\leq C \left(\max_{0\leq t\leq t^*} |\psi_+(t)| \right) \int_0^{t^*} \mathrm{d}t_1 \left| \frac{t^*}{t^*2} \right| \\
&= C \left(\max_{0\leq t\leq t^*} |\psi_+(t)| \right) \frac{t^*}{2}.
\end{aligned}$$

Hence for $t^* < 2/C$ it follows $\|K\| < 1$ and $\lim_{t\to 0}(K\psi_+)(t^*) = 0$. Thus the integral operator vanishes in comparison to ψ_+ as $t \to 0$, which shows the first statement of the lemma, equation (**??**). Moreover, by Banach's fixed-point theorem there exists a unique continuous solution on the interval $[0, t^*]$. Since the kernel of the integral operator K is bounded for all $t \geq t^*$ we obtain simultaneuously $\psi_+ \in C([0,\infty))$.

Induction step. Now let $\psi_+ \in C^k([0,\infty))$, for some $k \geq 0$. Then there exists a residual function R_k such that

$$\psi_+(t) = \sum_{j=0}^{k} \psi_+^{(j)}(0) \frac{t^j}{j!} + R_k(t) \qquad \text{and } R_k(t) = o\left(t^k\right) \ (t \to 0).$$

We compute the action of the integral operator on the residual function

$$\begin{aligned}
|(KR_k)(t)| &\leq \frac{1}{t\,|t-2i|} \int_0^t \mathrm{d}t_1 \, |P(t-t_1)| \sup_{0\leq t_1\leq t} |R_k(t_1)| \\
&= o\left(t^{k+1}\right) \ (t \to 0)
\end{aligned}$$

since $P(t) = \mathcal{O}(t)$ $(t \to 0)$. Hence, the application of the integral operator increases the order by one, $KR_k \in C^{k+1}([0,\infty))$, also we see $(KR_k)^{(k+1)}(0) = 0$. Now it follows from the integral equation (**??**), the analyticity of its right-hand side, and our series representation (4.3.18), by a component-wise comparison, that $\psi_+ \in C^{k+1}([0,\infty))$ and that $\psi_+^{(k+1)}(0)$ satisfies (**??**). □

LEMMA 4.3.6. *Let positive numbers ϵ_1, $\widetilde{a}$ and c be given such that $a - \widetilde{a} - \epsilon_1 > 0$, and let*

$$|P(t)| \leq c e^{-(a-\widetilde{a}-\epsilon_1)t}, \qquad t > 0.$$

Let ϵ_2 such that $0 < \epsilon_2 < a - \widetilde{a} - \epsilon_1$. Then, there exists $0 < \zeta < t^* := \sqrt{c/(a-\widetilde{a}-\epsilon_1-\epsilon_2)}$ such that for all $t^* \leq t$

$$|\psi_\pm(t)| \leq (a - \widetilde{a} - \epsilon_1)\left(\pi + t^* |\psi_+(\zeta)|\right) e^{-\epsilon_2(t-t^*)}.$$

PROOF. The idea of the proof is based on Gronwall's lemma, cf. [**25**, Chapt. 3.1, p. 63]. As in Lemma 4.3.5, we start from the integral equation for the cut function ψ, (**??**).

$$\psi_+(t) + \int_0^t \mathrm{d}t_1 \, \frac{P(t-t_1)}{t(t-2i)} \psi_+(t_1) = \pi \frac{P(t)}{t(t-2i)}, \qquad t > 0.$$

Next we take into account the structure of $P(\cdot)$, given by its definition (4.2.12) together with (4.2.11), and our estimate Lemma A.1.8, that is

$$\begin{aligned} |P(t)| &\leq c_1 e^{(\widetilde{a}-a)t} + c_2 t e^{-at}, \qquad t>0, \qquad \text{with } \widetilde{a} < a \\ &\leq c t e^{(\widetilde{a}-a)t} \\ &\leq c e^{(\widetilde{a}-a+\epsilon_1)t}, \qquad \text{with } \epsilon_1 > 0 \text{ such that } \widetilde{a}-a+\epsilon_1 < 0. \end{aligned}$$

Hence, there are always positive constants ϵ_1 and c satisfying the assumptions of the lemma. We introduce the notation

$$|P(t)| \leq c e^{-\delta_a t}, \qquad \delta_a := a - \widetilde{a} - \epsilon_1 > 0.$$

As first step, we split the integration interval into two intervals $[0, t^*]$ and $[t^*, t]$. Accordingly, the integral equation is given by

$$\begin{aligned} \psi_+(t) &= \pi \frac{P(t)}{t(t-2i)} - \int_0^{t^*} \mathrm{d}t_1 \, \frac{P(t-t_1)}{t(t-2i)} \psi_+(t_1) \\ &\quad - \int_{t^*}^{t} \mathrm{d}t_1 \, \frac{P(t-t_1)}{t(t-2i)} \psi_+(t_1). \end{aligned}$$

Next we multiply by $\exp(\delta_a t)$ and take the absolute values

$$\begin{aligned} |e^{\delta_a t}\psi_+(t)| &\leq \pi \left| e^{\delta_a t} \frac{P(t)}{t(t-2i)} \right| + \int_0^{t^*} \mathrm{d}t_1 \, e^{\delta_a t} \left| \frac{P(t-t_1)}{t(t-2i)} \right| |\psi_+(t_1)| \\ &\quad \int_{t^*}^{t} \mathrm{d}t_1 \, e^{\delta_a t} \left| \frac{P(t-t_1)}{t(t-2i)} \right| |\psi_+(t_1)|, \qquad t \geq t^*. \end{aligned} \tag{4.3.19}$$

We consider the integral terms separately. First we have, by the mean value theorem,

$$\begin{aligned} \int_0^{t^*} \mathrm{d}t_1 \, e^{\delta_a t} \left| \frac{P(t-t_1)}{t(t-2i)} \right| |\psi_+(t_1)| &= \frac{1}{|t(t-2i)|} \int_0^{t^*} \mathrm{d}t_1 \, e^{\delta_a t} P(t-t_1) \, |\psi_+(t_1)| \\ &= \frac{1}{|t(t-2i)|} t^* e^{\delta_a t} P(t-\zeta) \, |\psi_+(\zeta)|, \qquad 0 < \zeta < t^* \\ &\leq \frac{1}{|(t^*-2i)|} e^{\delta_a t} c e^{-\delta_a (t-\zeta)} \, |\psi_+(\zeta)| \\ &= \frac{c e^{\delta_a \zeta}}{|(t^*-2i)|} \, |\psi_+(\zeta)|. \end{aligned}$$

Next, let us choose real positive t^*, ϵ_2 such that

$$|t^*(t^*-2i)| > \frac{c}{\delta_a - \epsilon_2}, \qquad \delta_a - \epsilon_2 > 0,$$

which is satisfied for

$$t^* = \sqrt{\frac{c}{\delta_a - \epsilon_2}}.$$

This allows to bound the second integral as

$$\begin{aligned}\int_{t^*}^{t} \mathrm{d}t_1\, e^{\delta_a t} \left|\frac{P(t-t_1)}{t(t-2i)}\right| |\psi_+(t_1)| &\leq \int_{t^*}^{t} \mathrm{d}t_1\, e^{\delta_a t} c e^{-\delta_a (t-t_1)} \left|\frac{1}{t(t-2i)}\right| |\psi_+(t_1)| \\ &= \int_{t^*}^{t} \mathrm{d}t_1\, e^{\delta_a t_1} c \left|\frac{1}{t(t-2i)}\right| |\psi_+(t_1)| \\ &\leq \int_{t^*}^{t} \mathrm{d}t_1\, e^{\delta_a t_1} (\delta_a - \epsilon_2) |\psi_+(t_1)| \,.\end{aligned}$$

We bound the first term of the right-hand side of (4.3.19) by the same technique

$$\begin{aligned}\pi \left| e^{\delta_a t} \frac{P(t)}{t(t-2i)} \right| &\leq \pi \left| e^{\delta_a t} \frac{c e^{-\delta_a t}}{t(t-2i)} \right| \\ &= \pi \left| \frac{c}{t(t-2i)} \right| \\ &\leq \pi (\delta_a - \epsilon_2) \,.\end{aligned}$$

Hence, all together we find the inequality

$$\begin{aligned}\phi(t) &\leq \rho + \int_{t^*}^{t} \mathrm{d}t_1\, (\delta_a - \epsilon_2) |\phi(t_1)| \,, \qquad t \geq t^* \\ \phi(t) &:= \left| e^{\delta_a t} \psi_+(t) \right| \\ \rho &:= \pi (\delta_a - \epsilon_2) + \frac{c e^{\delta_a \zeta}}{|(t^* - 2i)|} |\psi_+(\zeta)| \,.\end{aligned}$$

It holds, by Gronwall's lemma

$$\begin{aligned}\phi(t) &\leq \rho e^{\int_{t^*}^{t} \mathrm{d}t_1\, (\delta_a - \epsilon_2)} \\ &= \rho e^{(\delta_a - \epsilon_2)(t - t^*)} .\end{aligned}$$

This yields

$$|\psi_+(t)| \leq \rho e^{-(\delta_a - \epsilon_2) t^* - \epsilon_2 t}$$

which results via

$$\begin{aligned}|\psi_+(t)| &\leq (\delta_a - \epsilon_2) \left(\pi e^{-\delta_a t^*} + t^* |\psi_+(\zeta)| \right) e^{-\epsilon_2 (t - t^*)} \\ &\quad 0 < \zeta < t^* \\ &\quad 0 < \epsilon_2 < \delta_a .\end{aligned}$$

finally in

$$|\psi_+(t)| \leq \delta_a \left(\pi + t^* |\psi_+(\zeta)|\right) e^{-\epsilon_2 (t - t^*)}, \qquad t^* \leq t,\, 0 < \epsilon_2 < \delta_a,\, 0 < \zeta < t^*$$

Since ψ_+ and ψ_- are complex conjugates, the same statement holds for ψ_-. □

The statements of Lemma 4.3.5 and Lemma 4.3.6 are – besides their value for the theoretical characterization – of vital interest for the numerical implementation. Since $\psi_\pm \in C^\infty(\mathbb{R}_+)$ we may apply high order discretization methods with good effect, and since $\psi_\pm$ decays exponentially for large distances on the cut, we may restrict the size of the support of approximating functions to a moderate size. In fact, this rapid decay is one of the characteristic differences between the numerical algorithms based on the cut function formulation and the direct computation of $\widehat{U}$ with the pole condition realized along the real axis.

4.4. Important Consequences

In the following we will use the theory of Section 4.2 to prove the essential results needed both for the theoretical justification of the pole condition approach as well as for the numerical implementation.

4.4.1. Representation formula. First we want to prove one of the most practical results of our analysis – a representation formula of the spatial function U in terms of the cut functions $\psi_\pm$. The derived formulas give not only the exterior solution, we will also use them as linking conditions between the interior problem and the exterior problem in our algorithms. Moreover, they form the basis of our understanding of the perfectly matched layer method.

THEOREM 4.4.1. *The function $U(r)$ and its derivatives $U^{(k)}(r)$, $(k \geq 1)$, have holomorphic extensions to the right half of the complex plane and satisfy the representation formulas*

$$
\begin{aligned}
(4.4.1) \qquad U(r) &= w(i)e^{ir}\left(\frac{1}{2i} - \frac{1}{2\pi i}\int_0^\infty \mathrm{d}t\, e^{-tr}\psi_+(t)\right) \\
&\quad + w(-i)e^{-ir}\left(-\frac{1}{2i} + \frac{1}{2\pi i}\int_0^\infty \mathrm{d}t\, e^{-tr}\psi_-(t)\right) \\
(4.4.2) \qquad U^{(k)}(r) &= w(i)e^{ir}\left(\frac{i^k}{2i} - \frac{1}{2\pi i}\int_0^\infty \mathrm{d}t\, (i-t)^k e^{-tr}\psi_+(t)\right) \\
&\quad + w(-i)e^{-ir}\left(-\frac{(-i)^k}{2i} - \frac{1}{2\pi i}\int_0^\infty \mathrm{d}t\, (-i-t)^k e^{-tr}\psi_-(t)\right)
\end{aligned}
$$

for $\mathrm{Re}(r) \geq 0$.

PROOF. The idea is to compute the inverse of the Laplace transformed spatial function $\widehat{U}$ by means of the inversion formula utilizing Cauchy's contour integral, contour deformation and our results on the decay behavior of $\widehat{U}$ and $\psi_\pm$ for large arguments. Let the contour parallel to the imaginary axis be denoted by $\gamma_1(t) := 1 + it$, $-R \leq t \leq R$, see Fig. 4.3.4. The inversion formula is

$$U(r) = \lim_{R\to\infty} \frac{1}{2\pi i}\int_{\gamma_1} \mathrm{d}s\, e^{rs}\widehat{U}(s).$$

In standard manner we use Cauchy's integral theorem

$$\frac{1}{2\pi i}\left(\int_{\gamma_1} \mathrm{d}s\, e^{rs}\widehat{U}(s) + \int_{\gamma_2} \mathrm{d}s\, e^{rs}\widehat{U}(s)\right) = 0$$

to compute the inversion integral based on the contributions from the path γ_2. The integrals from B to C, from C to D and from D to A vanish as $R \to \infty$, since $\left|\widehat{U}(s)\right| = \mathcal{O}\left(|s|^{-1}\right)$ as $|s| \to \infty$ and since the integrand decays exponentially as $\mathrm{Re}\, s \to -\infty$. The integrals along $S_\pm$ converge to $w(i)\int_0^\infty \mathrm{d}t\, \exp\left(r\left[\pm i - t\right]\right)\psi_\pm(t)$ as $R \to \infty$. A computation similar to that in the proof of Lemma 4.3.4 shows that the integral around i converges to $\pi w(i)e^{ir}$, and the integral around $-i$ to $\pi w(-i)e^{-ir}$ as $R \to \infty$. Summing the parts, we obtain (4.4.1). Differentiating this equation and changing the order of differentiating and integration, which is possible by Lebesgue's Dominated Convergence Theorem, we obtain the second equation, (4.4.2). □

Thus we have generalized the partial fraction decomposition technique used in the introductory examples to the general case. Eq. (4.4.1) supplies in a very natural way the desired decomposition of a function into its incoming and its outgoing part.

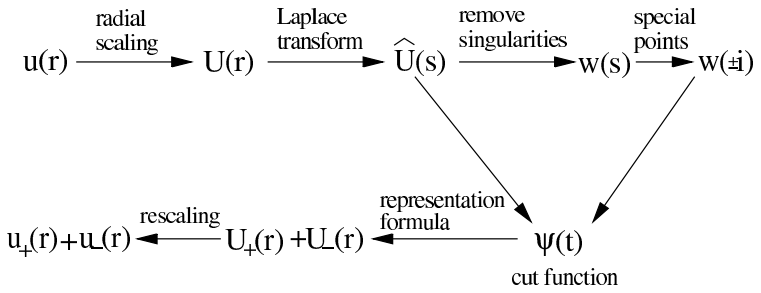


FIGURE 4.4.1. Decomposition of u into an incoming part u_- and an outgoing part u_+

Further we note that the incoming part is proportional to $w(-i)$, the outgoing part to $w(+i)$. Since with $w(i)$ the pole of $\widehat{U}(s)$ vanishes, too, we refer to $w_\pm(i)$ as pole strengths. Now the proof of the following corollary is very easy. Nevertheless it contains the central property that the pole condition can always be satisfied by a proper choice of the initial conditions.

COROLLARY 4.4.2. *The matrix $L(r)$ defined by*

$$\begin{pmatrix} U(r) \\ U'(r) \end{pmatrix} = L(r) \begin{pmatrix} w(i) \\ w(-i) \end{pmatrix}$$

using the representation formulas (4.4.1) and (4.4.2) is regular for all $r \geq 0$. Hence there exist complex numbers $U(r)$, $U'(r)$ such that $w(-i) = 0$. In this case, the spatial function U satisfies the pole condition, i.e. $\widehat{U}$ is holomorphic in the lower half of the complex plane.

PROOF. Let

$$L(r) \begin{pmatrix} w(i) \\ w(-i) \end{pmatrix} = 0.$$

Since U solves a linear second order differential equation, $U(r) = U'(r) = 0$ implies $U = 0$, thus $w(i) = w(-i) = 0$. If the pole strength $w(-i)$ vanishes, the cut function ψ_- vanishes, too, since $\psi_- = w(-i)\,[U]\,(-i - t)$. Therefore, $[w]\,(-i - t) = 0$ for all $t > 0$, which tells us that w is continuous in the lower half of the complex plane. Using Morera's Theorem and a contour deformation around the cut S_{-i} it can be shown that w is even holomorphic in the lower half plane. As $w(-i) = 0$, $\widehat{U}$ is holomorphic in the lower half-plane. □

4.4.2. Equivalent Formulations of the Pole Condition. We want to summarize the statements obtained so far concerning the pole condition. Before doing so, we summarize the interdependence of the different quantities introduced so far. Fig. 4.4.1 gives the corresponding graphical representation. After the radial scaling we obtain our basic spatial function $U(r)$, which is than Laplace transformed to get $\widehat{U}(s)$. The function $\widehat{U}(s)$ is multiplied with $s^2 + 1$ which yields a function $w(s)$ without any singularity. Based on $\widehat{U}(s)$ a cut function $\psi(t)$ is defined living only on the cuts $S_\pm$, and together with the values of w at $s = \pm i$, the spatial function $U(r)$ can be recovered, decomposed into an outgoing and an incoming part.

The representation formulas of Theorem 4.4.1 tell us in which way the function $\widehat{U}$ is expressed by the cut functions $\psi_\pm$ together with the pole strengths $w(\pm i)$. If $w(+i) = w(-i) = 0$, $\widehat{U}$ vanishes identically, consequently $U(r) \equiv 0$. If $w(-i) = 0$, then $\widehat{U} = \widehat{U}_+$, hence the corresponding spatial function $U(r)$ is a pure outgoing

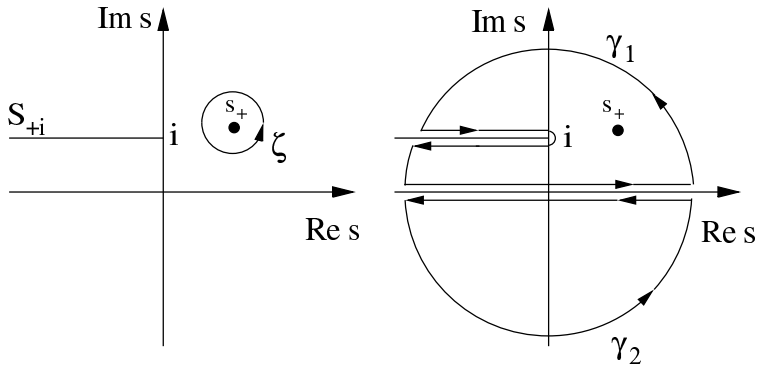


FIGURE 4.4.2. Deformation of the contour from a small circle (left) to the final contour

function and $\widehat{U}$ satisfies the pole condition, i.e. it is holomorphic in the lower half of the complex plane. Thus we may write the current form of the pole condition simply as

Pole condition in terms of the pole strength

$$w(-i) = 0.$$

This is the most simple form of the pole condition. However, this form is not constructive because it does not show how to realize the pole condition. On the other hand, the vanishing pole strength $w(-i) = 0$ means that $U(r)$ satisfies the pole condition if it has a representation

Pole condition in terms of the representation formulas
–cut function approach–

$$U(r) = w(i)e^{ir}\left(\frac{1}{2i} - \frac{1}{2\pi i}\int_0^\infty \mathrm{d}t\, e^{-tr}\psi_+(t)\right)$$

with some $w(i) \in \mathbb{C}$ and $\psi_+ a$ solution of

$$\psi_+(t) + \int_0^t \mathrm{d}t_1 \frac{P(t-t_1)}{t(t-2i)}\psi_+(t_1) = \pi\frac{P(t)}{t(t-2i)}.$$

Practically this requires the computation of the cut function and the pole strength $w(i)$. This, in turn, requires to compute the representation formulas (4.4.1) and (4.4.2) at $r = 0$ for $k = 0, 1$. We shall discuss this in detail in the application chapter, where we will use these formulas to construct a numerical algorithm, which we call in accordance with this theoretical framework *cut-function approach*.

There is a third form of the pole condition, formulated for the first time in **[81]**.

LEMMA 4.4.3. *Let $\widehat{U}$ satisfy the pole condition. Then,*

$$0 = \int_\infty^{-\infty} \mathrm{d}\tau \frac{\widehat{U}(\tau)}{\tau - s_+} \qquad \textit{for all} s_+ \in V \textit{ with } \mathrm{Im}\,(s_+) > 0.$$

PROOF. We apply Cauchy's integral theorem. Let s_+ given as defined above. First, we choose a small circle around s_+, see Fig. 4.4.2. Since s_+ is the only singularity inside the small circle, it holds

$$\widehat{U}\left(s_+\right) = \frac{1}{2\pi i}\oint_{\gamma} \mathrm{d}\zeta\, \frac{U(\zeta)}{\zeta - s_+}.$$

We deform the contour as shown in Fig. 4.4.2 and split the path γ into the two paths γ_1 and γ_2. This defines two functions $\widehat{U}_1(s_+)$ and $\widehat{U}_2(s_+)$ both depending on the radius R, and we have $\widehat{U}(s_+) = \widehat{U}_1(s_+) + \widehat{U}_2(s_+)$. We let $R \to \infty$ and denote the limiting functions again by $\widehat{U}_1(s_+)$ and $\widehat{U}_2(s_+)$. Now, since $U(\zeta)$ is holomorphic, we obtain

$$0 = \frac{1}{2\pi i}\oint_{\gamma_2} \mathrm{d}\zeta\, \frac{\widehat{U}(\zeta)}{\zeta - s_+} \qquad \text{for all} \, s_+ \in V \, \text{with} \, \mathrm{Im}\,(s_+) > 0.$$

Additionally, it holds by construction, $\widehat{U}(s) = \mathcal{O}\left(|s|^{-1}\right)$ as $|s| \to \infty$. Thus we have $\left|\widehat{U}(s)/\left(s - s_+\right)\right| = \mathcal{O}\left(|s|^{-2}\right)$ and the contour integral on the large semi-circle vanishes. This is the desired result. □

REMARK. The decomposition $\widehat{U}(s_+) = \widehat{U}_1(s_+) + \widehat{U}_2(s_+)$ via the Cauchy integral corresponds to the partial fraction decomposition discussed in connection with our introductory 1D example, which has motivated us to study the properties of the singularities in more general cases.

Now we are ready to give the third form of the pole condition.

Pole condition in terms of the axis integral
–real axis approach–

$$0 = \int_{\infty}^{-\infty} \mathrm{d}\tau\, \frac{\widehat{U}(\tau)}{\tau - s_+}, \qquad \tau \in \mathbb{R}, \ \text{for fixed} \ s_+ \in V \, \text{with} \, \mathrm{Im}\,(s_+) > 0.$$

with $\widehat{U}$ a solution of

$$\widehat{U}(\tau) + \int_{\tau}^{\infty} \mathrm{d}\tau_1\, \frac{P(\tau_1 - \tau)\widehat{U}(\tau_1)}{(k^2 + \tau^2)} = \frac{\tau U(0) + U'(0)}{(k^2 + \tau^2)}, \qquad \tau \in \mathbb{R}.$$

Observe the remarkable symmetry of these formulas in comparison to those related to the cut function approach. Observe further, that the formulas of the real axis approach does not possess any singularity on the real axis, which neither makes it necessary to carry out the transform (4.2.7) nor to exclude the term $1/r$ in the definition of the potentials $p_*(1/r)$, (4.2.11).

4.4.3. Asymptotic Expansion of the Far-Field.

THEOREM 4.4.4. *For all $m \geq 0$, U and U' of an outgoing function satisfy the asymptotic formulas*

$$\frac{U(r)}{w(i)} = e^{ir}\left(\frac{1}{2i} - \frac{1}{2\pi i}\sum_{j=1}^{m} \frac{\psi_+^{(j-1)}(0)}{r^j} + \mathcal{O}\left(\frac{1}{r^{m+1}}\right)\right) \tag{4.4.3}$$

$$\frac{U'(r)}{w(i)} = e^{ir}\left(\frac{1}{2} - \frac{1}{2\pi i}\sum_{j=1}^{m} \frac{i\psi_+^{(j-1)}(0) - (j-1)\psi_+^{(j-2)}}{r^j} + \mathcal{O}\left(\frac{1}{r^{m+1}}\right)\right) \tag{4.4.4}$$

for $r \to \infty$ such that $|\arg r| \le \phi < \pi/2$. Here $0 \cdot \psi_+^{(-1)}(0) := 0$.

PROOF. We start from the outgoing part of the representation formula (4.4.1)

$$\frac{U(r)}{w(i)} = e^{ir}\left(\frac{1}{2i} - \frac{1}{2\pi i}\int_0^\infty \mathrm{d}t\, e^{-tr}\psi_+(t)\right).$$

Note that the integral term is the Laplace transform of ψ_+, where the dual variable is r. Using the asymptotic formula

$$(Lf)(s) = \sum_{j=1}^{m} \frac{f^{(j-1)}(0)}{s^j} + \mathcal{O}\left(s^{-m-1}\right), \qquad s \to \infty,\ |\arg s| \le \phi < \pi/2,$$

which holds for any bounded function $f \in C^m(\mathbb{R}_+)$, (cf. [**26**, Vol. 2, p. 47] and Lemma 4.3.5 we obtain

$$\begin{aligned}\frac{U(r)}{w(i)} &= e^{ir}\left(\frac{1}{2i} - \frac{1}{2\pi i}\left(L\psi_+\right)(r)\right)\\ &= e^{ir}\left(\frac{1}{2i} - \frac{1}{2\pi i}\sum_{j=1}^{m}\frac{\psi_+^{(j-1)}(0)}{r^j} + \mathcal{O}\left(r^{-m-1}\right)\right).\end{aligned}$$

This is (4.4.3). The asymptotic formula for U' follows analogously from the representation formula for U', equation (4.4.2) written as

$$\frac{U'(r)}{w(i)} = e^{ir}\left(\frac{1}{2} - \frac{1}{2\pi i}\left(L\left[(i-t)\,\psi_+\right]\right)(r)\right)$$

and the identity

$$\frac{\mathrm{d}^j}{\mathrm{d}t^j}\left((i-t)\psi_+(t)\right)\Big|_{t\to 0} = i\psi_+^j(0) - j\psi_+^{(j-1)}(0).$$

□

COROLLARY 4.4.5. *The Hankel functions $H_\nu^{(1)}$ of the first kind of order ν satisfy*

$$H_\nu^{(1)}(r) = \sqrt{\frac{2}{\pi r}}e^{i\left(r - \frac{\nu\pi}{2} - \frac{\pi}{4}\right)}\left(\sum_{k=0}^{m}\left\{\prod_{j=1}^{k}\frac{\nu^2 - \left(j-\frac{1}{2}\right)^2}{-2ijr}\right\} + \mathcal{O}\left(r^{-m-1}\right)\right)$$

for $r \to \infty$ such that $|\arg r| \le \phi < \pi/2$.

PROOF. We take

$$\begin{aligned}p_*(t) &= 0\\ P(t) &= e^{-at}t\left(\frac{1}{4} - \nu^2\right)\end{aligned}$$

which results from (4.2.12) and (4.2.10) for a constant potential. With the normalization

$$w(i) = \sqrt{\frac{8}{\pi}}\exp\left(i\frac{\pi}{4} - \frac{\nu\pi}{2}\right)$$

we get $H_\nu^{(1)}(r+a) = (r+a)^{1/2}U(r)$. Carrying out the limit $a \to 0$ and using (**??**), (**??**) from Lemma 4.3.5 and the identity

$$\frac{1}{t(t-2i)} = \frac{1}{-2it}\sum_{j=0}^{\infty} t^j (2i)^{-j}$$

supplies

$$\begin{aligned}
\psi_+(0) &= 2\pi \frac{\nu^2 - \frac{1}{4}}{2i}\frac{w(i)}{2} \\
\psi_+^{(k+1)}(0) &= \left(\frac{k+1}{2i} + \frac{1}{2i(k+2)}\left(\frac{1}{4} - \nu^2\right)\right)\psi_+^{(k)}(0) \\
&= \frac{1}{-2i(k+2)}\left(\nu^2 - \left(k+\frac{3}{2}\right)^2\right)\psi_+^{(k)}(0).
\end{aligned}$$

□

4.4.4. Spectral Properties of the Dirichlet-to-Neumann Map . We define a complete orthonormal system of eigenfunctions of the Laplace-Beltrami operator on Γ_a by $\phi_j^a\left(a\mathbf{x}^0\right) := a^{-\frac{d-1}{2}}\phi_j\left(\mathbf{x}^0\right)$. Then, for $f \in H^{1/2}\left(\Gamma_a\right)$, we expect that the solution to our underlying Helmholtz equation (4.2.1) satisfying the pole condition and the boundary condition $\mathrm{Tr}_{\Gamma_a} u = f$ is given by

$$u\left(\left(r+a\right)\mathbf{x}^0\right) := \sum_{j=1}^{\infty}\left\langle \phi_j^a, f\right\rangle_{\Gamma_a}(r+a)^{-\frac{d-1}{2}} U_j(r)\phi_j\left(\mathbf{x}^0\right)$$

for $r \geq 0$ and $\mathbf{x}^0 \in S^{d-1}$, where the functions $U_j(r)$ are solutions of (4.2.8). This will be shown below. Differentiating the last equation with respect to r we formally find the following formula for the Dirichlet-to-Neumann map

$$\frac{\partial}{\partial r}u\left(a\mathbf{x}^0\right) = \sum_{j=1}^{\infty}\mathrm{DtN}\left(\lambda_j, a\right)\left\langle \phi_j^a, f\right\rangle_{\Gamma_a}\phi_j(\mathbf{x})$$

with the eigenvalues

$$\begin{aligned}
\mathrm{DtN}\left(\lambda_j, a\right) &= \frac{\left((r+a)^{-\frac{d-1}{2}}U_j(r)\right)'\Big|_{r=0}}{a^{-\frac{d-1}{2}}U_j(0)} \\
&= \frac{U_j'(0)}{U_j(0)} - \frac{d-1}{2a}. \qquad (4.4.5)
\end{aligned}$$

Since the Sobolev norm of index $s \in \mathbb{R}$ is given by

$$\begin{aligned}
\|f\|_{H^s(\Gamma_a)}^2 &= \left\|\left(I - \Delta_{\Gamma_a}\right)^{s/2} f\right\|_{L^2(\Gamma_a)}^2 \\
&= \sum_{j=1}^{\infty}\left(1-\lambda_j\right)^s\left|\left\langle \phi_j^a, f\right\rangle\right|^2
\end{aligned}$$

(cf. e. g. the book of Taylor **[88]**) the properties (i)-(iii) in Theorem 4.2.1 are equivalent to

$$|\mathrm{DtN}(\lambda_j, a)| = \mathcal{O}\left(\sqrt{|\lambda_j|}\right), \qquad j \to \infty \tag{4.4.6}$$

$$\mathrm{Re}\,(\mathrm{DtN}(\lambda_j, a) + l_j) \leq 0 \text{ for some sequence } |l_j| = o\left(\sqrt{|\lambda_j|}\right) \tag{4.4.7}$$

$$\mathrm{Im}\,\mathrm{DtN}(\lambda_j, a) > 0 \text{ for all } j. \tag{4.4.8}$$

The most difficult part is to establish the spectral property (4.4.6). Unfortunately, we cannot give a complete proof for the spectral property. In our earlier work we showed how to prove this property in the context of the pole condition for the case of a constant potential and two space dimensions **[86]**. Since the main tool was an explicit representation of the functions $\widehat{U}$, obtained via a sequence of transforms, this technique does not generalize to arbitrary potentials. In a joined work with Hohage and Zschiedrich **[58]** a complete proof is presented. However, the techniques applied there do not fit well in our presentation here, so we will restrict ourselves to an incomplete, but convincing motivation of the desired property.

We establish the property (4.4.8) first. The idea is to apply the conservation property of the energy flux, cf. **[86]**.

LEMMA 4.4.6. *If the spatial function U_j satisfies the pole condition, then*

$$\mathrm{Im}\left(U_j'(r)\overline{U_j(r)}\right) = \left|\frac{w(i)}{2}\right|^2$$

for all $r \geq 0$. In particular, $U_j(r) \neq 0$ for all $r \geq 0$ and

$$\mathrm{Im}\left(\frac{U_j'(r)}{U_j(r)}\right) > 0.$$

PROOF. We start from our basic equation for U in the spatial domain (4.2.8)

$$\begin{aligned} U_j''(r) + K(r)U_j(r) &= 0. \\ K(r) &:= \left[k^2 + \frac{\frac{1}{4}(d-1)(3-d) + \lambda_j}{(r+a)^2} + p\left(\frac{1}{r+a}\right)\right] \end{aligned}$$

with K a real function. Multiplying by $\overline{U_j}$, integrating the expression on the interval $[0, r]$ yields

$$\begin{aligned} 0 &= \int_0^r \mathrm{d}r_1\, \overline{U_j}\left(U_j''(r) + K(r)U_j(r)\right) \\ &= \overline{U_j(r)}U_j'(r) - \overline{U_j(0)}U_j'(0) + \int_0^r \mathrm{d}r_1\, \left(\overline{U_j'(r)}U_j'(r) + K(r)\overline{U_j(r)}U_j(r)\right). \end{aligned}$$

Taking the imaginary part of this equation, we obtain $\mathrm{Im}\left(\overline{U_j(r)}U_j'(r)\right) = \mathrm{Im}\left(\overline{U_j(0)}U_j'(0)\right) =$ const. The constant can be evaluated using Theorem 4.4.4 via

$$\lim_{r\to\infty} \mathrm{Im}\left(\overline{U_j(r)}U_j'(r)\right) = \left(-\frac{1}{2i}\overline{w(i)}e^{-ir}\right)\left(\frac{1}{2}w(i)e^{ir}\right) = \left|\frac{w(i)}{2}\right|^2.$$

It follows that both $U(r) \neq 0$ and $U'(r) \neq 0$ for all $r \geq 0$. Additionally, we know by ODE-theory that $U(r)$ is is bounded on each compact interval. By Theorem 4.4.4, it remains bounded on the whole semi-infinite interval. Hence it follows that

$$\mathrm{Im}\left(\frac{U_j'(r)}{U_j(r)}\right) = \left|\frac{w(i)}{2U_j(r)\overline{U_j(r)}}\right|^2 > 0.$$

□

Next we consider the spectral properties (4.4.6) and (4.4.7).

4.4.5. High-Frequency Limit of the Cut Functions. We look at the limit process $\lambda_j \to -\infty$. Since the eigenvalues of the Laplace-Beltrami operator are all non-positive, we may equivalently write $\lambda_j = -\nu^2$ with $\nu \geq 0$ and let $\nu \to \infty$. Let us consider our integral equation (74) with the polynomial $P(\cdot)$ for the constant potential, see (4.2.12), chosen according to our high-frequency consideration

$$\begin{aligned} \psi_+(t) + \int_0^t \mathrm{d}t_1 \, \frac{P\left(t-t_1\right)}{t(t-2i)} \psi_+\left(t_1\right) &= \pi \frac{P(t)}{t(t-2i)} \\ P(t) &:= -\nu^2 e^{-at} t. \end{aligned}$$

We introduce a new variable τ by $t := \nu\tau$ and normalize ψ_+ to $\widetilde{\psi}_\nu := c_\nu \psi_+$ via

$$\int_0^\infty \mathrm{d}\tau \, \widetilde{\psi}_\nu(\tau) = 1. \tag{4.4.9}$$

By Lemma 4.3.6, this is always possible. Next, we reformulate the integral equation using the new quantities

$$\widetilde{\psi}_\nu(\tau) + \int_0^\tau \mathrm{d}\tau_1 \, \frac{\widetilde{P}\left(\tau-\tau_1\right)}{\nu\tau\left(\tau - \frac{2i}{\nu}\right)} \widetilde{\psi}_\nu\left(\tau_1\right) = \pi \frac{c_\nu \widetilde{P}(\tau)}{\nu^2 \tau\left(\tau - \frac{2i}{\nu}\right)} \tag{4.4.10}$$

$$\widetilde{P}(\tau) = -\nu^3 e^{-a\nu\tau}\tau. \tag{4.4.11}$$

Now the idea is to approximate the integral for functions $\widetilde{\psi}_\nu(\tau)$ which vary at each point τ (with respect to the new scale) much slower than $\exp(-a\nu\tau)$ by

$$\begin{aligned} \int_0^\tau \mathrm{d}\tau_1 \, \frac{\widetilde{P}\left(\tau-\tau_1\right)}{\nu\tau\left(\tau - \frac{2i}{\nu}\right)} \widetilde{\psi}_\nu\left(\tau_1\right) &= -\int_0^\tau \mathrm{d}\tau_1 \, \frac{\nu^2 e^{-a\nu(\tau-\tau_1)}\left(\tau-\tau_1\right)}{\tau\left(\tau - \frac{2i}{\nu}\right)} \widetilde{\psi}_\nu\left(\tau_1\right) \\ &\approx -\frac{\widetilde{\psi}_\nu(\tau)}{(a\tau)^2} \qquad \text{for } \nu \text{ large.} \end{aligned}$$

Hence, the integral operator, which is compact, converges to a multiplication operator, which is not compact. In order to motivate our approach further, let us assume for the moment, that this integral approximation is exact and let us replace the integral operator in (4.4.10) by the multiplication operator and multiply the whole equation by $(a\tau)^2$

$$\widetilde{\psi}_\nu(\tau)\left((a\tau)^2 - 1\right) = -\pi a^2 c_\nu \nu e^{-a\nu\tau}. \tag{4.4.12}$$

The right-hand side tends to zero exponentially fast for any $\tau \geq \epsilon > 0$ *if* c_ν is bounded. Hence $\widetilde{\psi}_\nu$ tends to zero on the same interval except at the position $\tau = 1/a$. Consequently we expect $\widetilde{\psi}_\nu(\cdot) \to \delta(\cdot - 1/a)$ as $\nu \to \infty$. If we integrate both sides of (**??**) over the entire real axis, we find that the left-hand side tends to zero for $\nu \to \infty$ due to the delta property. The integration of the right-hand side yields $\pi a c_\nu$. Hence we expect $c_\nu \xrightarrow{\nu\to\infty} 0$. By the foregoing it seems reasonable to conjecture the following:

(1) $\widetilde{\psi}_\nu(\cdot) \to \delta\left(\cdot - 1/a\right)$ as $\nu \to \infty$, or equivalently,
(2) $\psi_\nu(\cdot) \to 1/c_\nu \, \delta\left(\cdot/\nu - 1/a\right)$ as $\nu \to \infty$, and
(3) $c_\nu \to 0$ as $\nu \to \infty$.

Using these assumptions we are able to derive the spectral properties (4.4.6) and (4.4.7). Based on the delta-property we obtain directly from (4.4.5) and our basic representation Theorem 4.4.1

$$\begin{aligned}\lim_{\nu\to\infty}\frac{\mathrm{DtN}\left(-\nu^2,a\right)}{\nu} &= -\lim_{\nu\to\infty}\tfrac{d-1}{2a\nu}+\lim_{\nu\to\infty}\frac{\frac{1}{2}-\frac{\nu}{2\pi i c_\nu}\int_0^\infty \mathrm{d}\tau\,(i-\nu\tau)\widetilde{\psi}_\nu(\tau)}{\frac{\nu}{2i}-\frac{\nu^2}{2\pi i c_\nu}\int_0^\infty \mathrm{d}\tau\,\widetilde{\psi}_\nu(\tau)}\\ &= -\frac{1}{a}.\end{aligned}$$

This implies both (4.4.6) and (4.4.7). Thus we have shown, up to the yet incomplete part in the proof of the spectral property, the unique solvability of the scattering problem interior to the artificial boundary Γ_a. We remember again that a complete, but technical proof is given in **[58]**.

4.4.6. Relation to the PML Method. There is a very close relation between the pole condition and the PML method, as indicated already in Section 2.5, p. 35. This is surprising, since the motivations behind both methods are very different. We want to discuss this relation in a brief, formal manner; nevertheless an extension of the above theory to the infinite dimensional case can be used to prove rigorously the convergence of the PML method even for problems with position dependent potentials. This work is part of the ongoing research and will be published together with Hohage and Zschiedrich in **[57]**. The proof, which we will present there, combines the technique of Lassas and Somersalo **[67]** with our representation formula (4.4.1).

To get the point, however, the more qualitative discussion given below is sufficient. First we specialize the modified Helmholtz equation (4.2.8) from page 65 to the most simple case, namely $d = 1$, $\lambda_j = 0$, $k = 1$, and $p(\cdot) \equiv 0$. This is the the normalized Helmholtz equation in 1D. The incoming field $U(r) := \exp(-ir)$ is a solution of this equation and its Laplace transform $\widehat{U}(s) := 1/(s+i)$ has a pole at $-i$. If we shift the pole by a distance σ, $\sigma \in \mathbb{R}_+$, to the right, cf. Fig. 4.4.3, we obtain the modified Laplace transform $\widehat{U}_{\mathrm{B}}(s) := 1/(s-[-i+\sigma])$. The corresponding solution in the spatial domain is $U_{\mathrm{B}}(r) := \exp([-i+\sigma]\,r) = \exp(-i\,[1+i\sigma]\,r)$, where the subscript B in $\widehat{U}_{\mathrm{B}}$ and U_{B} indicates the modification due to Bérenger. We can give this a new interpretation as continuation of the real distance variable r into the upper half of the complex plane. The desired effect of the stretching into the complex plane is that undamped oscillating, incoming functions are mapped to exponentially increasing functions and undamped oscillating outgoing functions are mapped to exponentially decreasing functions.

We want to extend this to the full transformed Helmholtz equation (4.2.8) in arbitrary space dimensions. We define $\alpha = 1 + i\sigma$, $0 \le \arg\alpha < \pi/2$ and discuss the complexified distance variable αr, where r denotes as before the real, positive physical distance. To analyze the effect of the complexification, we study the representation formula (4.4.1) (p. 80), with the cut functions $\psi_\pm$ replaced by their original definition (4.3.12)

$$\begin{aligned}U_{\mathrm{B}}(\alpha r) &= e^{i\alpha r}\left(\frac{1}{2i}-\frac{1}{2\pi i}\int_0^\infty \mathrm{d}t\, e^{-t\alpha r}\left[\widehat{U}_+\right](i-t)\right)\\ &\quad +e^{-i\alpha r}\left(-\frac{1}{2i}+\frac{1}{2\pi i}\int_0^\infty \mathrm{d}t\, e^{-t\alpha r}\left[\widehat{U}_-\right](-i-t)\right).\end{aligned}$$

By Lemma 4.3.5, p. 75, the cut functions are $C^\infty(\mathbb{R}_+)$-functions, and by Lemma 4.3.6, p. 77, they decay exponentially for r sufficiently large. Hence the integrals in the representation formula tend to 0 for r sufficiently large. Consequently, the behavior

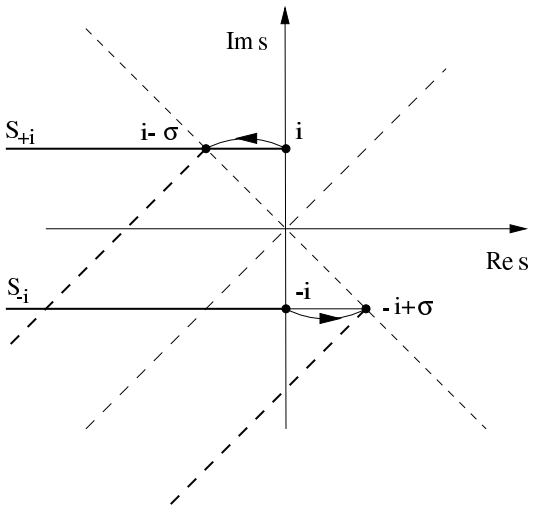


FIGURE 4.4.3. Original and rotated complex plane

of $U_{\mathrm{B}}(\alpha r)$ at sufficient large distances r is determined by the exponential factors before the parenthesis. The first term belongs to an exponentially decreasing, the second one to an exponentially increasing function. Exact transparent boundary conditions drop *incoming* functions. Now the central idea of PML is to drop the incoming functions, which at the same time are the exponential *increasing* functions by fixing zero-Dirichlet boundary conditions at the outer artificial boundary of the PML layer. Obviously, this forces $\widehat{U}_-$ to get small. In fact, Lassas and Somersalo proved in **[67]** that the error of the interior domain decreases like $\mathcal{O}\left(\exp\left(-Cr\right)\right)$ with C a positive constant and r the thickness of the PML-layer.

There is an even closer relation to our analysis based on the pole condition. Since the singularities at $\pm i$ of our original equations are mapped to $\pm\alpha i$, we may consider the complexified problem in a rotated complex plane $\mathbb{C}_\alpha := \alpha\mathbb{C}$, compare Fig. 4.4.3. This is the natural transform, since the complexification results in a transform

$$\alpha\widehat{U}_{\mathrm{B},\alpha}(\alpha s) = \widehat{U}_{\mathrm{B}}(s), \tag{4.4.13}$$

see Lemma A.1.9 of the Appendix A.1.

We introduce rotated cuts, see Fig. 4.4.3, define cut functions along these cuts, and carry out the whole analysis as for the original system. Finally we obtain the representation formula

$$\begin{aligned} U_{B,\alpha}(r) \;=\; & e^{i\alpha r}\left(\frac{1}{2i}-\frac{1}{2\pi i}\int_0^\infty \mathrm{d}t\, e^{-t\alpha r}\left[\widehat{U}_{\mathrm{B},\alpha,+}\right](\alpha(i-t))\right) \\ & +e^{-i\alpha r}\left(-\frac{1}{2i}+\frac{1}{2\pi i}\int_0^\infty \mathrm{d}t\, e^{-t\alpha r}\left[\widehat{U}_{\mathrm{B},\alpha,-}\right](\alpha(-i-t))\right). \end{aligned}$$

This coincides, of course, exactly with the representation formula derived for the original system, setting $U_{\mathrm{B},\alpha}(r) = U_{\mathrm{B}}(\alpha r)$. Thus we may express the relation between the pole condition approach and the PML method as follows:

(1) The complexification with a constant factor α corresponds to a rotation and rescaling in the Laplace domain.

(2) Singularities on the negative imaginary axis remain singularities in the lower half of the complex plane, hence the corresponding incoming functions remain incoming functions.
(3) Incoming functions asymptotically increase exponentially.
(4) The stretching into the complex plane offers the possibility to remove the singularities in the lower half of the complex plane by imposing zero-Dirichlet boundary conditions at infinity, which may be approximated by Dirichlet boundary conditions at sufficient large distances.

CHAPTER 5

Numerical Treatment of Helmholtz-Type Scattering Problems

We present two qualitative different approaches to solve scattering problems based on the pole condition:

(1) The factorization method.
(2) The Laplace domain method.

The factorization technique eliminates the poles in the lower half of the complex plane in a direct manner. That is, these poles are computed by a factorization of matrices and then eliminated by a proper choice of the Dirichlet and Neumann data on the boundary. In contrast, the Laplace domain method computes the Laplace transform either on the real axis (real axis approach) or on the cuts and eliminates the undesired poles by an additional integral condition (real axis approach) or simply by selecting of only one of the two possible cuts (cut function approach). In rather simple cases the factorization approach is preferable, since no additional discretization in the spectral domain is necessary. In more involved situations or non-separable coordinates, the Laplace domain method becomes advantageous, since no additional effort for matrix factorizations is needed.

5.1. Factorization Approach

We use the factorization approach to compute the reflection from two different types infinite obstacles: (1) The classical reflection from an infinite plane, (2) Reflection and diffraction from the half-plane already discussed in Section 2.2. This allows a direct comparison with the boundary integral method.

5.1.1. Reflection at an Infinite Plane. The reflection of a plane wave on an infinite plane with homogeneous Dirichlet or Neumann boundary conditions is the most simple example of a scattering problem, and analytical solutions are obtained immediately. However, the very closely related example of reflection and diffraction on a semi-infinite plane causes much more difficulties, and special tools from complex function theory like the Wiener-Hopf technique have to be applied. The numerical solution of the simple reflection on the infinite plane based on our Laplace technique with the pole condition in its core seems to be more difficult than the common technique based on an ansatz of plane waves. However, there is no much additional effort if we step over to the semi-infinite scattering problem, neither in the difficulty of the problem solution nor in the computational complexity.

Continuous Case. To prepare the semi-discrete case, which yields the desired algorithm, we look first at the continuous case. We wish to solve the Helmholtz equation

$$\begin{aligned} \partial_x^2 u + \partial_y^2 u + k^2 u &= 0 \quad \text{for } y \le 0 \\ \text{with } \partial_y u(x,y) &= u'(x,0) = 0 \qquad \text{at} \quad y = 0 \quad \text{for all} \quad x \in \mathbb{R} \end{aligned}$$

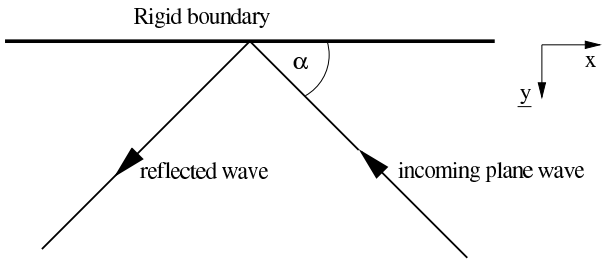


FIGURE 5.1.1. Reflection problem

in the space of twice differentiable, bounded functions on $\{(x,y) \in \mathbb{R}^2 : y \leq 0\}$, where an incident plane wave

$$u_{\text{in}}(x,y) = \mathrm{e}^{ik_x x + ik_y y}, \qquad \text{with} \quad k_x^2 + k_y^2 = k^2, \qquad k_x, k_y \in \mathbb{R},\ |k_x| < k \tag{5.1.1}$$

excites the reflected field. Note that in this special case the general source field u_{src} of the scattering problem (4.0.6) is equal to an incoming field, i. e. $u_{\text{in}} = u_{\text{src}}$, in the sense of our definition of incoming and outgoing fields. Further the above bounds on k_x, k_y exclude explicitely the case $|k_x| = k$, that is we exclude phase fronts perpendicular to the reflecting plane. The situation is depicted in Fig. 5.1.1. We define $\underline{y} := -y$ and use the $\underline{y}-$coordinate as unique distance variable; this prepares the reflection problem of the semi-infinite plane. The transverse variable x describes the position on the reflecting plane. We will solve the reflection problem as follows:

(1) Fourier transform of the field in x-direction, viewing the distance coordinate $\underline{y}$ as real parameter. Thus each Fourier component depends on $\underline{y}$.
(2) Laplace transform of the result with respect to $\underline{y}$.
(3) Imposing the pole condition to each Laplace transformed Fourier component. This supplies directly the field on the axis $\underline{y} = 0$.
(4) Inverse Laplace and Fourier transform to obtain the complete field.

Let $\mathcal{S}$ be the space of rapidly decreasing functions and $\mathcal{S}'$ the space of tempered generalized functions (tempered distributions). $\mathcal{S}'$ includes both the Dirac delta distribution and the undamped exponential function. By Fourier transform $F : \mathcal{S}' \to \mathcal{S}'$ along the boundary $\underline{y} = 0$

$$\widehat{v}\,(K_x) = \int_{-\infty}^{\infty} e^{-\mathrm{i}K_x x} v(x)\,\mathrm{d}x,$$

where the functions v are associated with the corresponding functionals $v_f : \mathcal{S} \to \mathbb{C}$, hence $v_f \in \mathcal{S}'$, via $v_f(w) = \int_{-\infty}^{\infty} \mathrm{d}x\, vw$ for all $w \in \mathcal{S}$, the Helmholtz equation transforms into an ODE with respect to the distance variable $\underline{y}$

$$-K_x^2 \widehat{u}\,(K_x, \underline{y}) + \partial_{\underline{y}}^2 \widehat{u}\,(K_x, \underline{y}) + k^2 \widehat{u}\,(K_x, \underline{y}) = 0.$$

Thus each Fourier component is indicated by a Fourier frequency K_x. Note that we did not impose any boundary condition so far.

Both incident and reflected waves have to satisfy this equation. Next we perform the Laplace transform with respect to $\underline{y}$, taking s as dual variable,

$$(s^2 + k^2 - K_x^2)\,\widehat{\widehat{u}}\,(K_x, s) = s\widehat{u}\,(K_x, 0) + \widehat{u}'\,(K_x, 0)\,, \tag{5.1.2}$$

which yields the corresponding solution

$$\widehat{\widehat{u}}(K_x,s)=\frac{s\widehat{u}(K_x,0)+\widehat{u}'(K_x,0)}{s^2+k^2-K_x^2} \tag{5.1.3}$$

in terms of the transversal wavenumber K_x and the Laplace variable s. We consider K_x as real parameter. The poles of this equation are given by

$$s_\pm=\pm i\sqrt{k^2-K_x^2}. \tag{5.1.4}$$

The poles s_- belong to incoming or increasing (in positive $\underline{y}$-direction) fields. Thus (5.1.3) satisfies the pole condition for each Fourier component if the numerator vanishes at the critical points $s=s_-$. Since we require that the outgoing solution $\widehat{\widehat{u}}_{\text{out}}(K_x,s)$ with boundary data $\widehat{u}_{\text{out}}(K_x,0)$ and $\widehat{u}'_{\text{out}}(K_x,0)$ possesses no pole s_- the following condition has to be satisfied:

Pole condition for the simple reflection problem

$$s_-\widehat{u}_{\text{out}}(K_x,0)+\widehat{u}'_{\text{out}}(K_x,0)=0. \tag{5.1.5}$$

Using the superposition $\widehat{u}=\widehat{u}_{\text{in}}+\widehat{u}_{\text{out}}$, we reformulate this condition in terms of the incoming and the complete field

$$s_-\left[\widehat{u}(K_x,0)-\widehat{u}_{\text{in}}(K_x,0)\right]+\widehat{u}'(K_x,0)-\widehat{u}'_{\text{in}}(K_x,0)=0.$$

This yields, taking the imposed Neumann condition $\widehat{u}'(K_x,0)=0$ into account,

$$\widehat{u}(K_x,0)=\widehat{u}_{\text{in}}(K_x,0)+\frac{\widehat{u}'_{\text{in}}(K_x,0)}{s_-},\ s_-\neq 0. \tag{5.1.6}$$

By definition of the incoming field (5.1.1) we have $\widehat{u}_{\text{in}}(K_x,0)=\delta(K_x-k_x)$ and $\widehat{u}'_{\text{in}}(K_x,0)=-ik_y\delta(K_x-k_x)$[1], hence $\widehat{u}_{\text{in}}(K_x,0),\,\widehat{u}'_{\text{in}}(K_x,0)\in\mathcal{S}'$. From (5.1.4) we see that $1/s_-$ becomes singular at the Fourier frequencies $|K_x|=k$. On the other hand, by definition of the incoming field and the bound $|k_x|<k$ we find $\lim_{|K_x|\to k}\widehat{u}'_{\text{in}}(K_x,0)/s_-(K_x)=0$. Thus $\widehat{u}(K_x,0)$ computed using (5.1.6) is in $\mathcal{S}'$ and the inverse Fourier transform exists. Computing the inverse Fourier transform of (5.1.6) yields

$$\begin{aligned}u(x,0) &= u_{\text{in}}(x,0)+F^{-1}\left(\frac{-ik_y\delta(K_x-k_x)}{-i\sqrt{k^2-K_x^2}}\right)\\ &= u_{\text{in}}(x,0)+\frac{k_y}{\sqrt{k^2-k_x^2}}e^{ik_xx}\\ &= u_{\text{in}}(x,0)+u_{\text{in}}(x,0).\end{aligned}$$

Thus we obtain the boundary condition for the superposed field

$$u(x,0)=2u_{\text{in}}(x,y),\qquad \text{or equivalently,}\quad \widehat{u}(K_x,0)=2\widehat{u}_{\text{in}}(K_x,0).$$

Once this boundary condition is known, the corresponding field can be computed directly. Using the solution representation (5.1.3), the imposed boundary condition $\widehat{u}'(K_x,0)=0$ and the just obtained boundary condition we find

[1]The minus sign occurs since the derivative refers to $\underline{y}$.

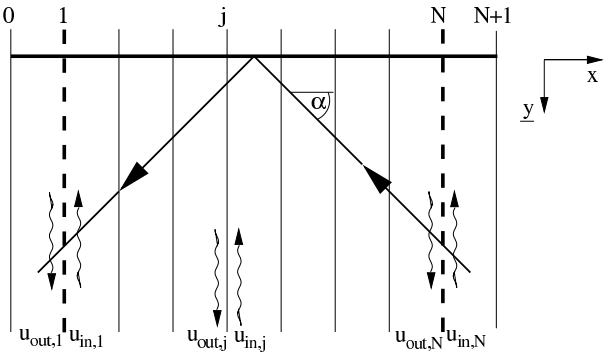


FIGURE 5.1.2. Reflection from an infinite plane: semi-discrete approximation

$$\widehat{\widehat{u}}\,(K_x,s)=\frac{2s\widehat{u}_{\rm in}\,(K_x,0)}{s^2+k^2-K_x^2}.$$

Partial fraction decomposition results in

$$\widehat{\widehat{u}}\,(K_x,s)=\left(\frac{1}{s-s_-}+\frac{1}{s+s_-}\right)\widehat{u}_{\rm in}\,(K_x,0)$$

which supplies, after inverse Laplace and Fourier transform,

$$u\,(x,\underline{y})=\mathrm{e}^{\mathrm{i}(k_x x+k_y\underline{y})}+\mathrm{e}^{\mathrm{i}(k_x x-k_y\underline{y})}. \tag{5.1.7}$$

REMARK. To derive (5.1.7) we have imposed an *angle condition* with respect to the incident field: the angle between the *wavevector* of the incident field and the reflecting *plane* must be larger than zero. In the limiting case $\alpha\to 0$, i. e.~$k_x\to k$ and $k_y\to 0$, we get from (5.1.7) $u(x,\underline{y})=2\exp(ikx)$. On the other hand, the incident plane wave $u_{\rm in}(x,\underline{y})=\exp(ikx)$ itself is a solution of the Helmholtz equation. Further *any* function $u(x,\underline{y})=\mathrm{const}\cdot\exp(ikx)$ solves the reflection problem with the imposed boundary condition. Hence the solution does not depend continuously on the incident angle if $\alpha\to 0$, which justifies the angle condition introduced above. In the semi-discrete case discussed in the next section we will find corresponding conditions which ensures the existence of the discrete solution.

Semi-discrete Case. We repeat the procedure above *without* using a Fourier transform. Instead, we restrict the infinite plane to a finite (but long) interval $[-a,a]$. The situation is schematically depicted in Fig. 5.1.2.

Our approach is as follows:

(1) Laplace transform of the semi-discrete problem with respect to $\underline{y}$.
(2) Discretization in x-direction.
(3) The pole condition with respect to the solution on each ray is traced back to a pole condition with respect to eigenmodes of the discrete operator.

Note that the first two points might be interchanged. We consider a finite element (FE) as well as a finite difference (FD) implementation. Let us assume for simplicity that the interval is divided into $N+1$ uniform subintervals. Further let us assume that the fields on the vertical boundaries $u_0\,(\underline{y})$ and $u_{N+1}\,(\underline{y})$ in case of finite differences or $\partial_x u_1\,(\underline{y})$ and $\partial_x u_N\,(\underline{y})$ in case of finite elements and the incident fields $u_{{\rm in},j}\,(\underline{y})$, $j=1,\ldots,N$, are given. First we perform a Laplace transform of the Helmholtz equation in positive $\underline{y}$-direction, using $\underline{y}$ as distance variable as before:

$$\partial_x^2\widehat{u}(x,s)+s^2\widehat{u}(x,s)+k^2\widehat{u}(x,s)=su(x,0)+u'(x,0). \tag{5.1.8}$$

Next we want to discretize this equation in x-direction by the standard finite difference and finite element methods. To this end we introduce the vector of unknown functions with elements $\widehat{u}_j(s) = \widehat{u}(x_j, s)$, $j = 0, \ldots, N+1$, and the vectors containing the Laplace transformed functions along the rays and the corresponding initial values at $\underline{y} = 0$

$$\widehat{\mathbf{u}}(s) = \begin{pmatrix} \widehat{u}_1(s) \\ \vdots \\ \widehat{u}_N(s) \end{pmatrix}, \quad \mathbf{u}_0 = \begin{pmatrix} u_1(0) \\ \vdots \\ u_N(0) \end{pmatrix}, \quad \mathbf{u}_0' = \begin{pmatrix} u_1'(0) \\ \vdots \\ u_N'(0) \end{pmatrix}.$$

A discretization results in the algebraic system[2], depending on the complex parameter s

$$(\mathrm{A} + s^2\mathrm{M})\,\widehat{\mathbf{u}}(s) = \widehat{\mathbf{u}}_\mathrm{b}(s) + s\mathrm{M}\mathbf{u}_0 + \mathrm{M}\mathbf{u}_0'. \tag{5.1.9}$$

In case of a finite element discretization the symmetric matrices $\mathrm{A} \in \mathbb{R}^{N\times N}$ and $\mathrm{M} \in \mathbb{R}^{N\times N}$ are computed based on the test and trial space $V_h = \operatorname{span}\{v_1, \ldots, v_N\}$ with continuous, piecewise linear functions v_j, where we set $v_1(-a) = 1$ and $v_N(a) = 1$. The finite element matrices A and M are computed by $(\mathrm{A})_{ij} = -(\partial_x v_i, \partial_x v_j) + k^2(v_i, v_j)$ and $(M)_{ij} = (v_i, v_j)$, the corresponding standard finite difference matrices are considered in Example 5.1.3. Note that M is symmetric positive definite and becomes the identity matrix in case of a finite difference discretization. The vertical boundary vector $\widehat{\mathbf{u}}_\mathrm{b} \in \mathbb{C}^N$ results from the boundary terms $v(\pm a)\partial_x\widehat{u}(\pm a, s)$ of the integration by parts in case of the finite element formulation or from the leftmost and the rightmost contribution of the three-point finite difference stencil in case of the finite difference formulation

$$\text{FE}: \widehat{\mathbf{u}}_\mathrm{b}(s) := \begin{pmatrix} -\partial x\widehat{u}_1(s) \\ \mathbf{0} \\ \partial x\widehat{u}_N(s) \end{pmatrix}, \quad \text{FD}: \widehat{\mathbf{u}}_\mathrm{b}(s) := -\frac{1}{h^2}\begin{pmatrix} \widehat{u}_0(s) \\ \mathbf{0} \\ \widehat{u}_{N+1}(s) \end{pmatrix}. \tag{5.1.10}$$

Eq. (5.1.9) is the semi-discrete analogue to (5.1.2), completed with the influence of the prescribed vertical boundaries.

We will formulate the boundary condition based on the spectral properties of the discrete operator $\mathrm{A} + s^2\mathrm{M}$ which appears on the left-hand side of (5.1.9). Therefore we consider, in preparation of the derivation of the boundary condition, the generalized, real symmetric eigenvalue problem

$$\mathrm{A}\mathbf{u}_j = \lambda_j^2\mathrm{M}\mathbf{u}_j \tag{5.1.11}$$

with eigenvalues λ_j^2 and eigenvectors $\mathbf{u}_j$. The eigenvalues are real and satisfy $\lambda_j^2 \leq k^2$. The eigenvectors $\mathbf{u}_j$ can be normalized to vectors $\mathbf{u}_j^0$ such that

$$\mathrm{U}^\mathrm{H}\mathrm{M}\mathrm{U} = \mathrm{I}, \qquad \text{with} \quad \mathrm{U} = (\mathbf{u}_1^0, \ldots, \mathbf{u}_N^0).$$

By means of U we can transform the matrix A to diagonal form

$$\mathrm{U}^\mathrm{H}\mathrm{A}\mathrm{U} = \Lambda^2, \;\; \Lambda^2 = \operatorname{diag}(\lambda_j^2). \tag{5.1.12}$$

Let the square root of λ_j^2 be defined as usual

[2] In this section we denote vectors arising from the discretization of continuous functions by bold roman letters. This simplifies, e. g., to distinguish between u as a function and $\mathbf{u}$ as a vector.

$$\lambda_j = \begin{cases} \sqrt{\lambda^2} & \text{if } \lambda^2 > 0 \\ i\sqrt{-\lambda^2} & \text{if } \lambda^2 < 0 \\ 0 & \text{if } \lambda^2 = 0. \end{cases}$$

Further, let $\mathrm{S}_\pm = \pm i\Lambda$ with $\Lambda = \operatorname{diag}(\lambda_j)$. Accordingly, we define the sets $\sigma_\pm = \{\pm i\lambda_1, \ldots, \pm i\lambda_N\}$. Note that $\sigma_- \cap \sigma_+$ is either empty or contains the zero eigenvalue.

Now we are able to derive the desired boundary condition. By the transform (5.1.12) we rewrite the semi-discrete, Laplace transformed Helmholtz equation (5.1.9) to

$$\left(\mathrm{U}^{\mathrm{H}}\mathrm{A}\mathrm{U} + s^2\mathrm{U}^{\mathrm{H}}\mathrm{M}\mathrm{U}\right)\mathrm{U}^{-1}\widehat{\mathbf{u}}(s) = \mathrm{U}^{\mathrm{H}}\left(\widehat{\mathbf{u}}_{\mathrm{b}}(s) + s\mathrm{M}\mathbf{u}_0 + \mathrm{M}\mathbf{u}_0'\right)$$

which yields

$$\left(\Lambda^2 + s^2\mathrm{I}\right)\mathrm{U}^{\mathrm{H}}\mathrm{M}\widehat{\mathbf{u}}(s) = \mathrm{U}^{\mathrm{H}}\left(\widehat{\mathbf{u}}_{\mathrm{b}}(s) + s\mathrm{M}\mathbf{u}_0 + \mathrm{M}\mathbf{u}_0'\right). \tag{5.1.13}$$

We factorize the operator on the left-hand side

$$\Lambda^2 + s^2\mathrm{I} = (\mathrm{I}s - S_-)\,(\mathrm{I}s + S_-).$$

Next we consider, as in the continuous case, only the outgoing (reflected) field resulting from data $\widehat{\mathbf{u}}_{\mathrm{b,out}}(s)$ of the vertical boundary and the corresponding initial data on the plane $\mathbf{u}_{0,\mathrm{out}}$, $\mathbf{u}'_{0,\mathrm{out}}$. The outgoing field has to satisfy the basic semi-discrete equation (5.1.9). To avoid a singularity of the expression $(\mathrm{I}s + S_-)\,\mathrm{U}^{\mathrm{H}}\mathrm{M}\widehat{\mathbf{u}}_{\mathrm{out}}(s)$ at one of the diagonal entries of S_-, i.e. for $s \in \sigma_-$, and equivalently of $\widehat{\mathbf{u}}_{\mathrm{out}}(s)$ at these points, we require that the right-hand side of (5.1.13) lies in the orthogonal complement of the nullspace of the transposed operator $\mathrm{I}s - S_-$:

$$\mathrm{U}^{\mathrm{H}}\left(\widehat{\mathbf{u}}_{\mathrm{b,out}}(s) + s\mathrm{M}\mathbf{u}_{0,\mathrm{out}} + \mathrm{M}\mathbf{u}'_{0,\mathrm{out}}\right) \in \mathcal{N}_{\mathrm{E}}^{\perp}\left(\mathrm{I}s - S_-\right) \qquad \text{for } s \in \sigma_-.$$

Here, and in the following we use the notation: $N^{\perp}(\mathrm{A})$ denotes the orthogonal complement of the null space of a matrix A with respect to $\langle\cdot,\mathrm{M}\cdot\rangle$ and $\mathcal{N}_{\mathrm{E}}^{\perp}(\mathrm{A})$ denotes the orthogonal complement of the null space of a matrix A with respect to the Euclidean metric $\langle\cdot,\cdot\rangle$. Since S_- is a diagonal matrix, the null space of $(\mathrm{I}s_j - S_-)$, with $s_j := -i\lambda_j \in \sigma_-$, is spanned by a single Euclidean unit vector $(\mathbf{e}_j)_{i=1,\ldots,N} = \delta_{ij}$. Thus, with $\mathbf{u}_j^0$ the jth normalized eigenvector from (5.1.11), the above pole condition may be given alternatively as

$$\left\langle \mathbf{u}_j^0, \widehat{\mathbf{u}}_{\mathrm{b,out}}(s_j) + s_j\mathrm{M}\mathbf{u}_{0,\mathrm{out}} + \mathrm{M}\mathbf{u}'_{0,\mathrm{out}}\right\rangle = 0, \qquad s_j \in \sigma_- \quad \text{and} \quad j = 1,\ldots,N. \tag{5.1.14}$$

This equation supplies N conditions. Defining the vector $\mathbf{b}_{\mathrm{v,out}}$ with elements

$$\left(\mathbf{b}_{\mathrm{v,out}}\right)_j := \left\langle \mathbf{u}_j^0, \widehat{\mathbf{u}}_{\mathrm{b,out}}(s_j)\right\rangle, \qquad j = 1,\ldots,N \tag{5.1.15}$$

and taking into account $\mathrm{U}\mathrm{U}^H\mathrm{M} = \mathrm{I}$ we can rewrite this in compact form

Pole condition for the semi-discrete reflection problem

$$\mathbf{0} = \mathrm{U}\mathbf{b}_{\mathrm{v,out}} + \mathrm{U}S_-\mathrm{U}^{\mathrm{H}}\mathrm{M}\mathbf{u}_{0,\mathrm{out}} + \mathbf{u}'_{0,\mathrm{out}}. \tag{5.1.16}$$

Eq. (5.1.16) is the semi-discrete analogue to condition (5.1.5) for the continuous problem. It avoids that the outgoing (reflected) field with boundary data $\mathbf{u}_{0,\mathrm{out}}, \mathbf{u}'_{0,\mathrm{out}}$ possesses poles which belong to incoming or increasing functions.

Next we compute the field caused by an incident field $\mathbf{u}_{\mathrm{in}}$. Let $\widehat{u}_{\mathrm{in},j}(s) = \widehat{u}_{\mathrm{in}}(x_j, s)$ be given. This means, we require the knowledge of the incident field along the vertical rays, cf. Fig. 5.1.2. As we will see, it does not play a role whether the incident field is a true incoming field in the sense of our definition of incoming and outgoing fields or a general source field containing both incoming and outgoing parts. We define the vector of Laplace transformed incoming functions along the rays, and the corresponding Dirichlet and Neumann data on the reflecting plane by

$$\widehat{\mathbf{u}}_{\mathrm{in}}(s) := \begin{pmatrix} \widehat{u}_{\mathrm{in},1}(s) \\ \vdots \\ \widehat{u}_{\mathrm{in},N}(s) \end{pmatrix}, \quad \mathbf{u}_{\mathrm{in},0} := \begin{pmatrix} u_{\mathrm{in},1}(0) \\ \vdots \\ u_{\mathrm{in},N}(0) \end{pmatrix}, \quad \mathbf{u}'_{\mathrm{in},0} := \begin{pmatrix} u'_{\mathrm{in},1}(0) \\ \vdots \\ u'_{\mathrm{in},N}(0) \end{pmatrix}.$$

Additionally we introduce the vectors $\mathbf{b}_{\mathrm{v}}$, $\mathbf{b}_{\mathrm{v,in}}$ in analogy to (5.1.15) by

$$\mathbf{b}_{\mathrm{v}} := \begin{pmatrix} \langle \mathbf{u}_1^0, \widehat{\mathbf{u}}_{\mathrm{b}}(s_1) \rangle \\ \vdots \\ \langle \mathbf{u}_N^0, \widehat{\mathbf{u}}_{\mathrm{b}}(s_N) \rangle \end{pmatrix}, \qquad \mathbf{b}_{\mathrm{v,in}} := \begin{pmatrix} \langle \mathbf{u}_1^0, \widehat{\mathbf{u}}_{\mathrm{b,in}}(s_1) \rangle \\ \vdots \\ \langle \mathbf{u}_N^0, \widehat{\mathbf{u}}_{\mathrm{b,in}}(s_N) \rangle \end{pmatrix}. \tag{5.1.17}$$

Based on the superposition of incoming and outgoing fields we rewrite (5.1.16) in terms of the complete and the incoming field

$$\mathbf{0} = \mathrm{U}(\mathbf{b}_{\mathrm{v}} - \mathbf{b}_{\mathrm{v,in}}) + \mathrm{US\_U^HM}(\mathbf{u}_0 - \mathbf{u}_{0,\mathrm{in}}) + \mathrm{UU^HM}(\mathbf{u}'_0 - \mathbf{u}'_{0,\mathrm{in}}). \tag{5.1.18}$$

We want to discuss the unique solvability of our reflection problem (5.1.9) together with condition (5.1.18), where we consider as given:

(1) The data $\mathbf{u}_{0,\mathrm{in}}$, $\mathbf{u}'_{0,\mathrm{in}}$ on the reflecting plane determined by the incoming field.
(2) The Neumann boundary condition $\mathbf{u}'_0(x)$ on the reflecting plane.
(3) The incoming fields $\widehat{\mathbf{u}}_{\mathrm{b,in}}(s)$ on the vertical boundaries, see (5.1.10).
(4) The complete field $\widehat{\mathbf{u}}_{\mathrm{b}}(s)$ (superposed from the incoming and the outgoing fields) on the vertical boundaries.

First of all note that the semi-discrete scattering problem (5.1.9) has a unique solution in the spatial domain for all given data $\mathbf{u}_0$, $\mathbf{u}'_0$, and $\mathbf{u}_{\mathrm{b}}$, as long as their Laplace transform exists. Thus the remaining question is under which circumstances we can compute a unique function $\mathbf{u}_0$ such that the reflected wave is an outgoing wave. We summarize these conditions in the following

LEMMA 5.1.1. *Let the given data have the following properties:*

(1) *Horizontal boundary: Let* $(\mathbf{u}'_0 - \mathbf{u}'_{0,\mathrm{in}}) \in \mathcal{N}^{\perp}(\mathrm{A})$.
(2) *Vertical boundary: Let* $(\widehat{\mathbf{u}}_{\mathrm{b}}(0) - \widehat{\mathbf{u}}_{\mathrm{b,in}}(0)) \in \mathcal{N}^{\perp}_{\mathrm{E}}(\mathrm{A})$

Then, the following holds true :
The boundary condition (5.1.18) has a unique solution vector $\mathbf{u}_0$ *and the reflected wave is an outgoing wave.*

PROOF. Let $0 \in \sigma_{-}$, otherwise the statement is trivial. Obviously, (5.1.18) has a unique solution if both

$$\text{(a)} \quad \mathrm{U^HM}(\mathbf{u}'_0 - \mathbf{u}'_{0,\mathrm{in}}) \in \mathcal{N}^{\perp}_{\mathrm{E}}(S_{-}) \qquad \text{and} \qquad \text{(b)} \quad (\mathbf{b}_{\mathrm{v}} - \mathbf{b}_{\mathrm{v,in}}) \in \mathcal{N}^{\perp}_{\mathrm{E}}(S_{-}).$$

Let the subscript j_0 indicate elements of nullspaces. Let $\mathbf{e}_{j_0} \in \mathcal{N}(\mathrm{S}_-)$. Hence $\mathbf{u}_{j_0} := \mathrm{U}\mathbf{e}_{j_0} \in \mathcal{N}(\mathrm{A})$. It follows directly from our assumption (1) written as

$$0 = \left\langle u_{j_0}, \mathrm{M}\left(\mathbf{u}_0' - \mathbf{u}_{0,\mathrm{in}}'\right)\right\rangle$$

that condition (a) is met

$$\begin{aligned} 0 &= \left\langle \mathrm{U}\mathbf{e}_{j_0}, \mathrm{M}\left(\mathbf{u}_0' - \mathbf{u}_{0,\mathrm{in}}'\right)\right\rangle \\ &= \left\langle \mathbf{e}_{j_0}, \mathrm{U}^{\mathrm{H}}\mathrm{M}\left(\mathbf{u}_0' - \mathbf{u}_{0,\mathrm{in}}'\right)\right\rangle. \end{aligned}$$

Recall the definition of $\mathbf{b}_{\mathrm{v}}$ and $\mathbf{b}_{\mathrm{v,in}}$ in (5.1.17):

$$\left(\mathbf{b}_{\mathrm{v}}\right)_j := \left\langle \mathbf{u}_j^0, \widehat{\mathbf{u}}_{\mathrm{b}}\left(s_j\right)\right\rangle, \qquad \left(\mathbf{b}_{\mathrm{v,in}}\right)_j := \left\langle \mathbf{u}_j^0, \widehat{\mathbf{u}}_{\mathrm{b,in}}\left(s_j\right)\right\rangle, \qquad j = 1, \ldots, N.$$

By $s_{j_0} = -i\lambda_{j_0} = 0$ it follows

$$\begin{aligned} \left\langle \mathbf{u}_{j_0}^0, \widehat{\mathbf{u}}_{\mathrm{b}}\left(s_{j_0}\right)\right\rangle &= \left\langle \mathbf{u}_{j_0}^0, \widehat{\mathbf{u}}_{\mathrm{b}}\left(0\right)\right\rangle \\ &= \left(\mathbf{b}_{\mathrm{v}}\right)_{j_0}. \end{aligned}$$

Obviously we have also

$$\left\langle \mathbf{u}_{j_0}^0, \widehat{\mathbf{u}}_{\mathrm{b}}\left(0\right) - \widehat{\mathbf{u}}_{\mathrm{b,in}}\left(0\right)\right\rangle = \left(\mathbf{b}_{\mathrm{v}}\right)_{j_0} - \left(\mathbf{b}_{\mathrm{v,in}}\right)_{j_0}$$

Hence it follows from assumption (2), $(\widehat{\mathbf{u}}_{\mathrm{b}}(0) - \widehat{\mathbf{u}}_{\mathrm{b,in}}(0)) \in \mathcal{N}_{\mathrm{E}}^{\perp}(\mathrm{A})$, that $\left(\mathbf{b}_{\mathrm{v}}\right)_{j_0} - \left(\mathbf{b}_{\mathrm{v,in}}\right)_{j_0} = 0$. Thus $\langle \mathbf{b}_{\mathrm{v}} - \mathbf{b}_{\mathrm{v,in}}, \mathbf{e}_{j_0}\rangle = 0$. This is condition (b). The boundary condition (5.1.16) determines the properties of the reflected field. The condition excludes all modes with poles $\operatorname{Im} s_j < 0$. Thus the reflected field must be an outgoing field in the sense of our definition. □

The lemma tells us that the outgoing field has to satisfy two independent conditions, one concerning the horizontal and the other concerning the vertical boundaries. However, based the orthogonality condition (5.1.14) we can put these into a single condition using the notation of Lemma 5.1.1:

COROLLARY 5.1.2. *The boundary condition (5.1.18) has unique solution vector* $\mathbf{u}_0$ *if*

$$\left\langle \mathbf{u}_{j_0}^0, \widehat{\mathbf{u}}_{\mathrm{b}}\left(0\right) - \widehat{\mathbf{u}}_{\mathrm{b,in}}\left(0\right) + \mathrm{M}\left(\mathbf{u}_0' - \mathbf{u}_{0,\mathrm{in}}'\right)\right\rangle = 0 \qquad \textit{for all } j_0 \textit{ with } \lambda_{j_0}^2 = 0$$

and the corresponding reflected wave is an outgoing wave.

The boundary condition (5.1.16) is the general semi-discrete boundary condition for the reflection from the (restricted) infinite plane, which we want to keep in this form, in view of its application to the scattering from a semi-infinite plane. We get the semi-discrete analogue to the continuous condition (5.1.6) if we set $\mathbf{u}' = 0$, as required in the problem formulation.

EXAMPLE 5.1.3. Let us apply the standard uniform finite-difference method to derive (5.1.9) to the case of rigid scattering $u'(x,0) = 0$. We take x_j, $j = 0, \ldots, N+1$, such that $h = x_{k+1} - x_k$, $i = 0, \ldots, N$ and $x_1 = -a$, $x_N = a$ and obtain the real symmetric matrices $\mathrm{A}, \mathrm{M} \in \mathbb{R}^{N \times N}$

$$\mathrm{A} := \frac{1}{h^2}\begin{pmatrix} -2 & 1 & & \\ 1 & -2 & 1 & \\ & & \ddots & \\ & & 1 & -2 \end{pmatrix} + k^2\mathrm{I}, \qquad \mathrm{M} = \mathrm{I}. \tag{5.1.19}$$

Let the incoming field be a plane wave

$$u_{\mathrm{in}}\left(x_j, \underline{y}\right) = \mathrm{e}^{ik_{\mathrm{x}}\mathrm{x}_j + ik_{\mathrm{y}}\underline{y}},$$

with real wavenumbers k_x and k_y, $k_y^2 < k^2$. We set $z = \exp\left(ik_x h\right)$ such that the source field takes the form $u_{\mathrm{in}}\left(x_j, \underline{y}\right) = z^j \exp\left(ik_{\mathrm{y}}\underline{y}\right)$. Since k_x is real we find $|z| = 1$. We require that the incoming field satisfies the semi-discrete Helmholtz equation on the half-plane $\underline{y} \geq 0$ and $j = 1, \ldots, N$:

(5.1.20)
$$\frac{1}{h^2}\left(u_{\mathrm{in}}\left(x_{j-1}, \underline{y}\right) - 2u_{\mathrm{in}}\left(x_j, \underline{y}\right) + u_{\mathrm{in}}\left(x_{j+1}, \underline{y}\right)\right) + \partial_y^2 u_{\mathrm{in}}\left(x_j, \underline{y}\right) + k^2 u_{\mathrm{in}}\left(x_j, \underline{y}\right) = 0.$$

Replacing $u_{\mathrm{in}}\left(x_j, \underline{y}\right)$ by the above ansatz, the semi-discrete dispersion relation

$$z^2 + z\left(-2 + h^2\left(-k_y^2 + k^2\right)\right) + 1 = 0 \tag{5.1.21}$$

follows. We assume that the discretization step width h and the wavenumber k_y of the incident plane wave are given. Then, the two roots of the semi-discrete dispersion relation supply the two possible values for the transversal wavenumber k_x. We discuss this in brief. For

$$h < \frac{2}{\sqrt{k^2 - k_y^2}}$$

the zeros $z_{1,2}$ of (5.1.21) have an imaginary part. Since the polynomial is real, they must be complex conjugates. Then, by Vieta's theorem $z_1 z_2 = 1$, hence $|z_1| = |z_2| = 1$ as required. The expressions z_1^j, z_2^j are discrete periodic functions. The field

$$u_{\mathrm{in}}(x_j, \underline{y}) = \mathrm{e}^{\pm i\left(\arg \mathrm{z}_1\right) j + ik_{\mathrm{y}}\underline{y}} \tag{5.1.22}$$

obeys the semi-discrete dispersion relation and we find $k_x = \pm\left(\arg \mathrm{z}_1\right)/h$.

In order to follow the treatment of this section, namely the formulation of the boundary condition in terms of the eigenvalues and eigenvectors of the discrete operator, we compute first all eigenvectors $\mathbf{u}_j$ and corresponding eigenvalues λ_j^2 of the eigenvalue problem

$$\mathrm{A}\mathbf{u}_j = \lambda_j^2 \mathbf{u}_j, \qquad \text{with} \quad \mathbf{u}_j = e^{i\lambda_j \underline{y}}\left(u_j\left(x_1\right), \ldots, u_j\left(x_N\right)\right)^T \tag{5.1.23}$$

and set $\sigma(\mathrm{A}) = \left\{\lambda_1^2, \ldots, \lambda_N^2\right\}$. To solve the eigenvalue problem we insert the ansatz $\left(\mathbf{v}_j\right)_l = z^l \exp\left(i\lambda_j \underline{y}\right)$, $l = 1, \ldots, N$, into the semi-discrete Helmholtz equation (5.1.20) and compute the zeros $z_{1,2}$ of the semi-discrete dispersion relation

$$z^2 + z\left(-2 + h^2\left(k^2 - \lambda_j^2\right)\right) + 1 = 0. \tag{5.1.24}$$

We use these zeros to compose a vector $\left(\widetilde{\mathbf{u}_j}\right)_{l=0,\ldots,N+1} = bz_1^l + cz_2^l$, with complex coefficients b, c. The restricted vector $\left(\widetilde{\mathbf{u}_j}\right)_{l=1,\ldots,N}$ constructed this way solves the eigenvalue problem (5.1.23) with arbitrary coefficients b, c except for the first and the last row of the system. The additional conditions $\left(\widetilde{\mathbf{u}}\right)_0 = 0$ and $\left(\widetilde{\mathbf{u}}\right)_{N+1} = 0$ then guaranty that $\left(\widetilde{\mathbf{u}_j}\right)_{l=1,\ldots,N} = \left(\mathbf{u}_j\right)_{l=1,\ldots,N}$ where $\mathbf{u}_j$ is in fact the desired eigenvector since it satisfies the eigenvalue equation (5.1.23) . These additional conditions together with the properties $|z_1| = |z_2| = 1$ and $z_1 = \overline{z}_2$ yield

$$\begin{aligned} b &= -c \\ z_{1,2}^{2(N+1)} &= 1. \end{aligned}$$

The last equation has $N+1$ different zeros: $z_{1,2}^{(j)} = \exp(\pm\pi ij/(N+1))$, $j = 0,\ldots,N$. Ignoring the constant solution corresponding to $j=0$, which only yields the trivial eigenvector due to the vanishing boundary data, we find N periodic eigenfunctions and N corresponding eigenvalues

$$\begin{aligned} (\mathbf{u}_j)_{l=1,\ldots,N} &= \sin\left(j\frac{l\pi}{N+1}\right) \\ \lambda_j^2 &= k^2 + \frac{2}{h^2}\left(\cos\left(j\frac{\pi}{N+1}\right) - 1\right), \qquad j = 1,\ldots,N. \end{aligned} \tag{5.1.25}$$

From the eigenvalue formula (5.1.25) it becomes clear that the nullspace of A in general is empty. Then, the unique solvability of the boundary condition (5.1.18) is trivial and Lemma 5.1.1 tells us that the reflected wave is an outgoing wave.

However, we may construct a situation where the nullspace of the matrix A is non-empty. Let us analyze such a situation in the following. First we choose the parameter h such that there exists a zero-eigenvalue. We fix an integer j_0, $1 \le j_0 \le N$, and set in the eigenvalue equation (5.1.25) $\lambda_{j_0}^2 = 0$, which results in a condition with respect to h

$$h = \frac{1}{k}\sqrt{2\left(1 - \cos\frac{j_0\pi}{N+1}\right)}.$$

The corresponding eigenvector $\mathbf{u}_{j_0}$ meets

$$\mathrm{A}\mathbf{u}_{j_0} = \mathbf{0}.$$

A comparison of the semi-discrete dispersion relation (5.1.21) which yields wavenumbers k_y^2 such that the corresponding semi-discrete periodic functions satisfy the semi-discrete Helmholtz equation (5.1.20) and the dispersion relation (5.1.24) which yields the eigenvalues λ_j^2 of the eigenvalue problem (5.1.23) shows that only wavenumbers k_y^2 are admissible, which are in the spectrum of A. Otherwise the incident wave $u_{\mathrm{in}}(x_j, \underline{y})$ is *not* a discrete periodic function in x. Note that the vector $\mathbf{u}_{0,\mathrm{in}}$ defined by $(\mathbf{u}_{0,\mathrm{in}})_j = u_{\mathrm{in}}(x_j, 0)$, $j = 1,\ldots,N$, is not an eigenfunction of the eigenvalue problem (5.1.23), despite of $k_y^2 \in \sigma(\mathrm{A})$. This follows from the fact that the incidence wave has a representation z_1^j, whereas the eigenvector has a representation $z_1^j - z_1^{-j}$. Further we see from the eigenvalue formula (5.1.25) that $\lambda_j^2 < 0$ for $j > j_0$, hence $\mathrm{Im}(\lambda_j) \neq 0$ for $j > j_0$. Since $k_y^2 \in \sigma(\mathrm{A})$ we find a necessary restriction for the possible choices of k_y^2, namely $k_y^2 = \lambda_j^2$ for some j with $1 \le j < j_0$. Otherwise the incident wave would be an exponential increasing or decreasing function.

After this preparation, we can analyze the solvability of our example problem in analogy to the continuous problem. Let $k_y \in \mathbb{R}$, $k_y^2 \in \sigma(\mathrm{A})$, $0 < |k_y| < k$, let k_x be computed according to (5.1.22). Further let the vertical boundary functions are the exact functions known from the continuous case

$$\begin{aligned} u_{\mathrm{in}}(\pm a, \underline{y}) &= \mathrm{e}^{\mathrm{i}(\pm ak_x - k_y\underline{y})} \\ u(\pm a, \underline{y}) &= \mathrm{e}^{\mathrm{i}(\pm ak_x + k_y\underline{y})} + \mathrm{e}^{\mathrm{i}(\pm ak_x - k_y\underline{y})}. \end{aligned}$$

with Laplace transforms

$$\begin{aligned}\widehat{u}_{\text{in}}(\pm a, s) &= \frac{1}{s+ik_y}\mathrm{e}^{\pm \mathrm{i}k_x a}\\ \widehat{u}(\pm a, s) &= \frac{1}{s-ik_y}\mathrm{e}^{\pm \mathrm{i}k_x a}+\frac{1}{s+ik_y}\mathrm{e}^{\pm \mathrm{i}k_x a}.\end{aligned}$$

Further we compute

(5.1.26)
$$\widehat{\mathbf{u}}_{\text{b,in}}(0)=\frac{1}{ik_y}\begin{pmatrix}\mathrm{e}^{-\mathrm{i}k_x a}\\ \mathbf{0}\\ \mathrm{e}^{\mathrm{i}k_x a}\end{pmatrix},\qquad \widehat{\mathbf{u}}_{\text{b}}(0)=\frac{1}{ik_y}\begin{pmatrix}\mathrm{e}^{-\mathrm{i}k_x a}\\ \mathbf{0}\\ \mathrm{e}^{\mathrm{i}k_x a}\end{pmatrix}-\frac{1}{ik_y}\begin{pmatrix}\mathrm{e}^{-\mathrm{i}k_x a}\\ \mathbf{0}\\ \mathrm{e}^{\mathrm{i}k_x a}\end{pmatrix}=\mathbf{0}.$$

We have to show that these function meet the condition of Corollary 5.1.2. With (5.1.26) we can express this condition as

(5.1.27)
$$\left\langle \mathbf{u}^0_{j_0}, \widehat{\mathbf{u}}_{\text{b,in}}(0)+\mathbf{u}'_{0,\text{in}}\right\rangle = 0.$$

To prove the condition, we begin with the fact that the incoming field $\mathbf{u}_{\text{in}}(\underline{y})$ must satisfy the spatial counterpart of the semi-discrete Laplace transformed Helmholtz equation (5.1.9) for all $\underline{y}\geq 0$, see also this equation in component-wise form (5.1.20). We restrict this to $\underline{y}=0$, set again $\mathbf{u}_{0,\text{in}}:=\mathbf{u}_{\text{in}}(0)$ and obtain

(5.1.28)
$$\mathrm{A}\mathbf{u}_{0,\text{in}}-k_y^2\mathbf{u}_{0,\text{in}}=\mathbf{u}_{\text{b,in}}.$$

Further, it follows from (5.1.26)

$$\widehat{\mathbf{u}}_{\text{b,in}}(0)=\frac{1}{ik_y}\mathbf{u}_{\text{b,in}}.$$

Replacing $\mathbf{u}_{\text{b,in}}$ in (5.1.28) by $\widehat{\mathbf{u}}_{\text{b,in}}(0)$ using the last equation and taking the inner product with $\mathbf{u}_{j_0}$ it follows

$$\left\langle \mathbf{u}_{j_0}, \mathrm{A}\mathbf{u}_{0,\text{in}}\right\rangle - k_y^2\left\langle \mathbf{u}_{j_0}, \mathbf{u}_{0,\text{in}}\right\rangle = ik_y\left\langle \mathbf{u}_{j_0}, \widehat{\mathbf{u}}_{\text{b,in}}(0)\right\rangle.$$

Since $\mathbf{u}_{j_0}\in\mathcal{N}(\mathrm{A})$ and $\mathrm{A}=\mathrm{A}^T$ we get

$$ik_y\left\langle \mathbf{u}_{j_0}, \widehat{\mathbf{u}}_{\text{b,in}}(0)\right\rangle + k_y^2\left\langle \mathbf{u}_{j_0}, \mathbf{u}_{0,\text{in}}\right\rangle = 0.$$

Finally we use $\mathbf{u}'_{0,\text{in}}=-ik_y\mathbf{u}_{0,\text{in}}$, which holds true by the definition of the semi-discrete plane wave, to find

$$ik_y\left\langle \mathbf{u}_{j_0}, \widehat{\mathbf{u}}_{\text{b,in}}(0)\right\rangle + k_y^2\left\langle \mathbf{u}_{j_0}, \frac{1}{-ik_y}\mathbf{u}'_{0,\text{in}}\right\rangle = 0.$$

Since $k_y\neq 0$ this shows (5.1.27). Hence the discrete problem has exactly one solution even in the case when the nullspace of A is non-empty, if the incidence wave is a discrete periodic wave and $k_y\neq 0$. This coincides exactly with the angle condition in the continuous case.

We show by a direct computation that $\widehat{\mathbf{u}}(s)=\widehat{\mathbf{u}}_{\text{in}}(s)+\widehat{\mathbf{u}}_{\text{out}}(s)$ with

(5.1.29)
$$\begin{array}{cccccccc}(\widehat{\mathbf{u}}_{\text{in}})_j &=& \frac{1}{s+ik_y}\mathrm{e}^{\mathrm{i}k_x x_j}, & (\mathbf{u}_{0,\text{in}})_j &=& \mathrm{e}^{\mathrm{i}k_x x_j}, & (\mathbf{u}'_{0,\text{in}})_j &=& -k_y(\mathbf{u}_{\text{in},0})_j\\ (\widehat{\mathbf{u}}_{\text{out}})_j &=& \frac{1}{s-ik_y}\mathrm{e}^{\mathrm{i}k_x x_j}, & (\mathbf{u}_{0,\text{out}})_j &=& \mathrm{e}^{\mathrm{i}k_x x_j}, & (\mathbf{u}'_{0,\text{out}})_j &=& k_y(\mathbf{u}_{\text{out},0})_j\end{array}$$

($0\leq j\leq N+1$) satisfies in fact the semi-discrete, Laplace tranformed Helmholtz equation

$$(\mathrm{A}+s^2)\,\widehat{\mathbf{u}}(s)=\widehat{\mathbf{u}}_{\mathrm{b}}(s)+s\mathbf{u}_0+\mathbf{u}_0'. \tag{5.1.30}$$

Let us consider the jth row of the system. Using $z^j=\exp(\mathrm{i}k_x x_j)$ it follows

$$\begin{aligned}(\widehat{\mathbf{u}}_{\mathrm{in}})_j+(\widehat{\mathbf{u}}_{\mathrm{out}})_j &= \left(\frac{1}{s+ik_y}+\frac{1}{s-ik_y}\right)z^j\\ &= \frac{2s}{s^2+k_y^2}z^j.\end{aligned}$$

Further recall that by the definition of $\widehat{\mathbf{u}}_{\mathrm{b}}(s)$ in (5.1.10) and A in (5.1.19)

$$(\mathrm{A}\widehat{\mathbf{u}}(s)-\widehat{\mathbf{u}}(s))_j=\left[\frac{1}{h^2}\left(z^{j-1}-2z^j+z^{j+1}\right)+k^2z^j\right]\frac{2s}{s^2+k_y^2},\qquad 1\le j\le N.$$

Thus the left-hand side of (5.1.30) minus $\widehat{\mathbf{u}}(s)$ reads

$$\left(\mathrm{A}\widehat{\mathbf{u}}(s)-\widehat{\mathbf{u}}(s)+s^2\widehat{\mathbf{u}}(s)\right)_j=\left[\frac{1}{h^2}\left(z^{j-1}-2z^j+z^{j+1}\right)+k^2z^j+s^2\right]\frac{2s}{s^2+k_y^2}.$$

Taking into account that $\mathbf{u}_0'=0$ by the imposed boundary condition we rewrite (5.1.30) to

$$\begin{aligned}\left[\frac{1}{h^2}\left(z^{j-1}-2z^j+z^{j+1}\right)+k^2z^j+s^2z^j\right]\frac{2s}{s^2+k_y^2} &= s\,(\mathbf{u}_0)_j\\ &= 2sz^j.\end{aligned}$$

This results in

$$\frac{1}{h^2}\left(z^{-1}-2+z\right)+k^2+s^2=s^2+k_y^2$$

which is obviously identical to the semi-discrete dispersion relation (5.1.21), hence it is satisfied. Thus the functions (5.1.29) obey the semi-discrete scattering problem. An inverse Laplace transform supplies the corresponding spatial solutions:

$$u\,(x_j,\underline{y})=\mathrm{e}^{\mathrm{i}(k_xx_j+k_y\underline{y})}+\mathrm{e}^{\mathrm{i}(k_xx_j-k_y\underline{y})},\qquad j=1,\dots,N,$$

which in fact meets the boundary condition $\mathbf{u}_0'=\mathbf{u}_{0,\mathrm{in}}'+\mathbf{u}_{0,\mathrm{out}}'=\mathbf{0}$. Apparently, the obtained functions are the direct discrete counterpart to the continuous solution (5.1.7).

The following facts are remarkable:

(1) We obtained exactly the same solution as obtained for the continuous problem, simply sampled at discrete values on the x-axis.
(2) The rays needed to discretize the exterior domain do not need to follow the phase front of the incident waves.
(3) The incident waves may have both incoming and outgoing parts, nevertheless, the reflected wave is a pure outgoing wave.

5.1.2. Reflection and Diffraction by a Semi-Infinite Plane. We apply the results to the reflection/diffraction by a semi-infinite plane. We decompose the x-axis into a left and a right semi-infinite interval, see Fig. 5.1.3 and require

$$\lim_{\underline{y}\downarrow 0}\partial_{\underline{y}}u_{21}\,(x,\underline{y})=0\quad\text{and}\quad\lim_{y\downarrow 0}\partial_y u_{11}(x,y)=0\quad\text{for } x\le 0, \tag{5.1.31}$$

which are the imposed boundary conditions. The Laplace transformed Helmholtz equation (5.1.8) should hold both in the upper and in the lower half-plane. We denote the distance variable in the upper half-plane by y, the distance variable in

the lower half-plane with $\underline{y}$, as in the previous case. We discretize the axis $y = 0$ in the interval $[-a, a]$ and obtain, corresponding to (5.1.9), the two algebraic equations

$$(\mathrm{A} + s^2\mathrm{I})\,\widehat{\mathbf{u}}_1(s) = \widehat{\mathbf{u}}_{\mathrm{b},1}(s) + s\mathbf{u}_{10} + \mathbf{u}'_{10} \tag{5.1.32}$$

$$(\mathrm{A} + s^2\mathrm{I})\,\widehat{\mathbf{u}}_2(s) = \widehat{\mathbf{u}}_{\mathrm{b},2}(s) + s\mathbf{u}_{20} + \mathbf{u}'_{20}, \tag{5.1.33}$$

where we want to consider only the finite difference case with $\mathrm{M} = \mathrm{I}$. The finite element approach can be treated in complete analogy. Let the incident wave be as in Example 5.1.3

$$u_{\mathrm{in}}\left(x_j, \underline{y}\right) = e^{\mathrm{i}k_x x_j} e^{-ik_y \underline{y}}, \quad k_y^2 \in \sigma(\mathrm{A}), \quad 0 < |k_y| < k, \tag{5.1.34}$$

where as before $\left(\mathbf{u}_{0,\mathrm{in}}\right)_j = u_{\mathrm{in}}\left(x_j, 0\right)$, $j = 0, \ldots, N+1$. We make the interval width $2a$ so large that the following vertical boundary conditions become reasonable

$$\begin{aligned} u_{11}(-a, y) &= 0 \\ u_{12}(a, y) &= e^{\mathrm{i}k_x a} e^{ik_y y} \\ u_{21}(-a, \underline{y}) &= e^{-\mathrm{i}k_x a} e^{ik_y \underline{y}} + e^{-\mathrm{i}k_x a} e^{-ik_y \underline{y}} \\ u_{22}(a, \underline{y}) &= e^{\mathrm{i}k_x a} e^{-ik_y \underline{y}}. \end{aligned} \tag{5.1.35}$$

Consequently, we define the vertical boundary data needed to compose the vertical boundary functions $\widehat{\mathbf{u}}_{\mathrm{b}}$, as they has to be used by the finite difference (FD) method:

$$\begin{aligned} &\mathrm{FD}: \\ u_{11}(-a-h, y) &= 0 \\ u_{12}(a+h, y) &= e^{\mathrm{i}k_x(a+h)} e^{ik_y y} \\ u_{21}(-a-h, \underline{y}) &= e^{\mathrm{i}k_x(-a-h)} e^{ik_y \underline{y}} + e^{\mathrm{i}k_x(-a-h)} e^{-ik_y \underline{y}} \\ u_{22}(a+h, \underline{y}) &= e^{\mathrm{i}k_x(a+h)} e^{-ik_y \underline{y}} \end{aligned}$$

Then, the boundary functions $\widehat{\mathbf{u}}_{\mathrm{b},1}$ and $\widehat{\mathbf{u}}_{\mathrm{b},2}$ are computed based on the definition (5.1.10), where again the subscript 1 refers to the upper, and the subscript 2 to the lower half-plane:

$$\begin{aligned} \widehat{\mathbf{u}}_{\mathrm{b},1}(s) &:= -\frac{1}{h^2} \begin{pmatrix} \widehat{u}_{11}(s) \\ \mathbf{0} \\ \widehat{u}_{12}(s) \end{pmatrix} \\ &= -\frac{1}{h^2} \begin{pmatrix} 0 \\ \mathbf{0} \\ \frac{1}{s - ik_y} e^{\mathrm{i}k_x(a+h)} \end{pmatrix} \end{aligned} \tag{5.1.36}$$

$$\begin{aligned} \widehat{\mathbf{u}}_{\mathrm{b},2}(s) &:= -\frac{1}{h^2} \begin{pmatrix} \widehat{u}_{21}(s) \\ \mathbf{0} \\ \widehat{u}_{22}(s) \end{pmatrix} = \\ &-\frac{1}{h^2} \begin{pmatrix} \frac{1}{s - ik_y} e^{\mathrm{i}k_x(-a-h)} + \frac{1}{s + ik_y} e^{\mathrm{i}k_x(-a-h)} \\ \mathbf{0} \\ \frac{1}{s + ik_y} e^{\mathrm{i}k_x(a+h)} \end{pmatrix} \end{aligned} \tag{5.1.37}$$

The corresponding vectors $\mathbf{b}_{\mathrm{v}}$ are computed by (5.1.15).

Next, we write simply the general boundary condition (5.1.16) twice, each for the upper and the lower half plane. We set $\mathrm{B} := \mathrm{US\_U^H}$ and drop the index 0 at the Dirichlet and Neumann data on $y = \underline{y} = 0$ and set $\delta\mathbf{b}_{\mathrm{v},k} := \mathbf{b}_{\mathrm{v},k} - \mathbf{b}_{\mathrm{in,v},k}$, $k \in \{1, 2\}$. To distinguish between the lower and the upper half-plane, and the

right and left half-plane, respectively, we use indices given in Fig. 5.1.3. Thus the boundary condition (5.1.16) can be written as

$$\begin{aligned} \mathrm{B}\begin{pmatrix} \mathbf{u}_{11} \\ \mathbf{u}_{12} \end{pmatrix} &= -\left(\mathrm{U}\delta\mathbf{b}_{\mathrm{v},1} + \begin{pmatrix} \mathbf{u}'_{11} - \mathbf{u}'_{\mathrm{in},11} \\ \mathbf{u}'_{12} - \mathbf{u}'_{\mathrm{in},12} \end{pmatrix}\right) + \mathrm{B}\mathbf{u}_{\mathrm{in},1} \\ \mathrm{B}\begin{pmatrix} \mathbf{u}_{21} \\ \mathbf{u}_{22} \end{pmatrix} &= -\left(\mathrm{U}\delta\mathbf{b}_{\mathrm{v},2} + \begin{pmatrix} \mathbf{u}'_{21} - \mathbf{u}'_{\mathrm{in},21} \\ \mathbf{u}'_{22} - \mathbf{u}'_{\mathrm{in},22} \end{pmatrix}\right) + \mathrm{B}\mathbf{u}_{\mathrm{in},2}. \end{aligned}$$

We specify the reflection problem by: $\mathbf{u}_{\mathrm{in},1} = 0$, $\mathbf{u}'_{\mathrm{in},11} = \mathbf{u}'_{\mathrm{in},12} = 0$, $\mathbf{u}'_{11} = \mathbf{u}'_{21} = 0$. Obviously, we have four unknown vectors $\mathbf{u}_{11}, \mathbf{u}_{12}, \mathbf{u}_{22}, \mathbf{u}'_{12}$, since $\mathbf{u}'_{12} = -\mathbf{u}'_{21}$ by definition of the coordinate systems and $\mathbf{u}_{12} = \mathbf{u}_{22}$. Taking this into account, the problem simplifies to

$$\mathrm{B}\begin{pmatrix} \mathbf{u}_{11} \\ \mathbf{u}_{12} \end{pmatrix} = -\left(\mathrm{U}\delta\mathbf{b}_{\mathrm{v},1} + \begin{pmatrix} \mathbf{0} \\ \mathbf{u}'_{12} \end{pmatrix}\right) \tag{5.1.38}$$

$$\mathrm{B}\begin{pmatrix} \mathbf{u}_{21} \\ \mathbf{u}_{12} \end{pmatrix} = -\left(\mathrm{U}\delta\mathbf{b}_{\mathrm{v},2} + \begin{pmatrix} -\mathbf{u}'_{\mathrm{in},21} \\ -\mathbf{u}'_{12} - \mathbf{u}'_{\mathrm{in},22} \end{pmatrix}\right) + \mathrm{B}\mathbf{u}_{\mathrm{in},2}. \tag{5.1.39}$$

We want to show that the discrete half-plane scattering problem has exactly one solution under the similar conditions as the more simple discrete reflection problem of Section 5.1.1, pp. 91.

LEMMA 5.1.4. *Let the discrete scattering problem of a semi-infinite plane defined by the boundary conditions (5.1.34), the discretizations of the upper and lower half-planes (5.1.32), (5.1.33), the incident plane wave (5.1.34) and the vertical boundary conditions (5.1.35).*

Further assume that the nullspace of the matrix

$$\mathrm{C} := \begin{bmatrix} \left(\mathrm{U}S\_\mathrm{U}^{\mathrm{H}}\mathrm{M}\right)_{1,\ldots,N_1;\,1,\ldots,N_1} & \mathbf{0} \\ \left(\mathrm{U}S\_\mathrm{U}^{\mathrm{H}}\mathrm{M}\right)_{N_1+1,\ldots,N;\,1,\ldots,N_1} & -\mathrm{I}_{N_2,N_2} \end{bmatrix}$$

with N_1 the number of points on the rigid boundary and $N_1 + N_2 = N$ is empty.

Then, the discrete problem has a unique solution and the reflected wave is an outgoing wave.

PROOF. We have to show that the two equations (5.1.38), (5.1.39) have a unique solution. First we take the sum and the difference of (5.1.32) and (5.1.39)

$$\begin{aligned} \mathrm{B}\begin{pmatrix} \mathbf{u}_{11} + \mathbf{u}_{21} \\ 2\mathbf{u}_{12} \end{pmatrix} &= -\left(\mathrm{U}\left(\delta\mathbf{b}_{\mathrm{v},1} + \delta\mathbf{b}_{\mathrm{v},2}\right) + \begin{pmatrix} -\mathbf{u}'_{\mathrm{in},21} \\ -\mathbf{u}'_{\mathrm{in},22} \end{pmatrix}\right) \\ &\quad + \mathrm{B}\mathbf{u}_{\mathrm{in},2} \end{aligned} \tag{5.1.40}$$

$$\begin{aligned} \mathrm{B}\begin{pmatrix} \mathbf{u}_{11} - \mathbf{u}_{21} \\ 0 \end{pmatrix} &= -\left(\mathrm{U}\left(\delta\mathbf{b}_{\mathrm{v},1} - \delta\mathbf{b}_{\mathrm{v},2}\right) + \begin{pmatrix} \mathbf{u}'_{\mathrm{in},21} \\ 2\mathbf{u}'_{12} + \mathbf{u}'_{\mathrm{in},22} \end{pmatrix}\right) \\ &\quad - \mathrm{B}\mathbf{u}_{\mathrm{in},2} \end{aligned} \tag{5.1.41}$$

We consider (5.1.40) first. We aim to establish an orthogonality condition similar to that in Corollary 5.1.2. Let j_0 indicate elements of nullspaces, that is $\mathbf{u}_{j_0} \in \mathcal{N}(\mathrm{A})$ and and the corresponding eigenvalue $\lambda^2_{j_0} \in \sigma(\mathrm{A})$ is zero. Further let $\mathbf{u}_{j_0} = \mathrm{U}\,\mathbf{e}_{\mathrm{j}_0}$. The system (5.1.40) has a unique solution if

$$\begin{aligned}
&\left\langle \mathbf{e}_{j_0}, \delta\mathbf{b}_{\mathrm{v},1} + \delta\mathbf{b}_{\mathrm{v},2} + \mathrm{U}^{\mathrm{H}} \begin{pmatrix} -\mathbf{u}'_{\mathrm{in},21} \\ -\mathbf{u}'_{\mathrm{in},22} \end{pmatrix} \right\rangle \\
&= (\delta\mathbf{b}_{\mathrm{v},1} + \delta\mathbf{b}_{\mathrm{v},2})_{j_0} + \left\langle \mathbf{u}_{j_0}, \begin{pmatrix} -\mathbf{u}'_{\mathrm{in},21} \\ -\mathbf{u}'_{\mathrm{in},22} \end{pmatrix} \right\rangle \\
&= \langle \mathbf{u}_{j_0}, \widehat{\mathbf{u}}_{\mathrm{b},1}(0) - \widehat{\mathbf{u}}_{\mathrm{b,in},1}(0) + \widehat{\mathbf{u}}_{\mathrm{b},2}(0) - \widehat{\mathbf{u}}_{\mathrm{b,in},2}(0) \rangle - \left\langle \mathbf{u}_{j_0}, \begin{pmatrix} \mathbf{u}'_{\mathrm{in},21} \\ \mathbf{u}'_{\mathrm{in},22} \end{pmatrix} \right\rangle \\
&= \left\langle \mathbf{u}_{j_0}, \widehat{\mathbf{u}}_{\mathrm{b},1}(0) - \widehat{\mathbf{u}}_{\mathrm{b,in},1}(0) + \widehat{\mathbf{u}}_{\mathrm{b},2}(0) - \widehat{\mathbf{u}}_{\mathrm{b,in},2}(0) - \begin{pmatrix} \mathbf{u}'_{\mathrm{in},21} \\ \mathbf{u}'_{\mathrm{in},22} \end{pmatrix} \right\rangle \\
&= 0.
\end{aligned}$$

The evaluation of (5.1.36) and (5.1.37) shows $\widehat{\mathbf{u}}_{\mathrm{b},1}(0) + \widehat{\mathbf{u}}_{\mathrm{b},2}(0) = \mathbf{0}$. Hence the condition reduces to

$$\left\langle \mathbf{u}_{j_0}, \widehat{\mathbf{u}}_{\mathrm{b,in},1}(0) + \widehat{\mathbf{u}}_{\mathrm{b,in},2}(0) + \begin{pmatrix} \mathbf{u}'_{\mathrm{in},21} \\ \mathbf{u}'_{\mathrm{in},22} \end{pmatrix} \right\rangle = 0.$$

This, however coincides with condition (5.1.27), p. 101, whose validity has been proved. Thus (5.1.40) has a unique solution.

Now we consider (5.1.41), which we want to rewrite, setting $\mathbf{u} := \mathbf{u}_{11} - \mathbf{u}_{21}$, $\mathbf{v} = 2\mathbf{u}'_{12}$,

$$\mathbf{r} := -\left(\mathrm{U}\,(\delta\mathbf{b}_{\mathrm{v},1} - \delta\mathbf{b}_{\mathrm{v},2}) + \begin{pmatrix} \mathbf{u}'_{\mathrm{in},21} \\ \mathbf{u}'_{\mathrm{in},22} \end{pmatrix} \right) - \mathrm{US\_U}^{\mathrm{H}} \mathbf{u}_{\mathrm{in},2},$$

and using the definition of B above to

$$\mathrm{US\_U}^{\mathrm{H}} \begin{pmatrix} \mathbf{u} \\ \mathbf{0} \end{pmatrix} - \begin{pmatrix} \mathbf{0} \\ \mathbf{v} \end{pmatrix} = \mathbf{r}.$$

Using the matric C this linear problem takes the form.

$$\mathrm{C} \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} = \mathbf{r}. \tag{5.1.42}$$

By assumption, (5.1.42) has a unique solution. Consequently, the whole problem has a unique solution. □

5.2. Laplace domain method

The Laplace domain methods in their two forms supply a technique to satisfy the pole condition without any explicit factorization. The price to pay is an additional discretization along axes in the spectral domain. We based the real axis approach on a conventional spline-collocation technique and the cut-function approach on a Runge-Kutta-technique. Both realizations are the very first implementations and far-away from being optimal. For example, it seems to be quite natural to replace the spline technique by a *wavelet* realization. Most of all questions are open in this direction. Nevertheless, the numerical results obtained so far demonstrate clearly the high potential of the Laplace domain methods. We describe both realizations in great detail, such that they may repeated immediately.

Model equation. We consider the numerical solution of our basic 1D model problem, the Bessel equation

$$\partial_r^2 u + \frac{d-1}{r} \partial_r u + k^2 u - \frac{\nu^2}{r^2} u = 0, \tag{5.2.1}$$

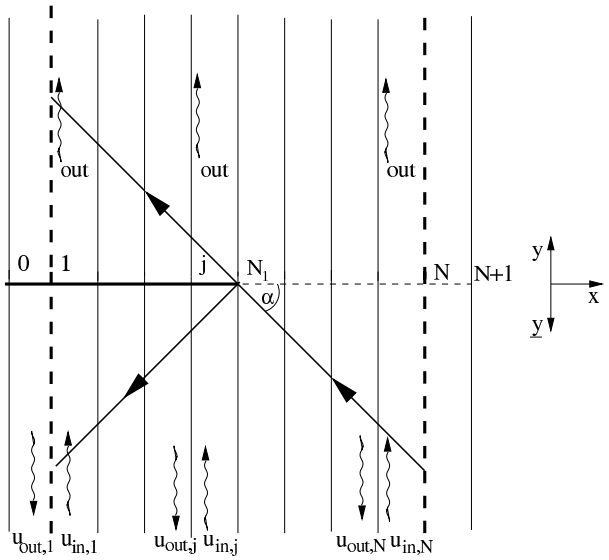


FIGURE 5.1.3. Diffraction problem: discretization by vertical rays

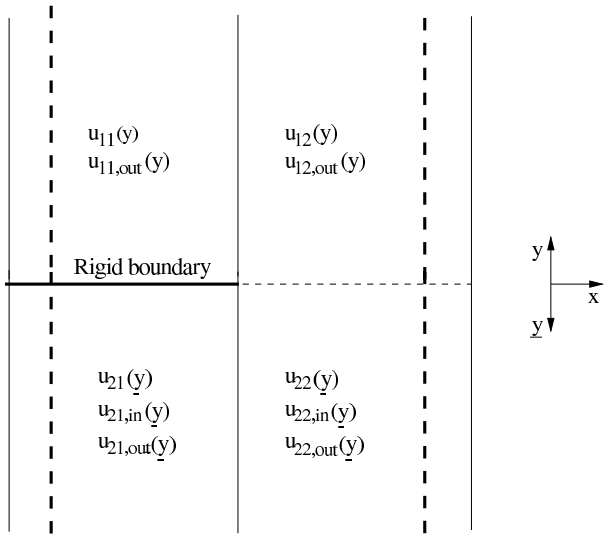


FIGURE 5.1.4. Notation of the solution on different sub-domains

where again d is the space dimension and ν an arbitrary real separation constant. We define the interior and the exterior domain to $\Omega_{\text{int}} = (r_{\text{i}}, r_0)$ and $\Omega_{\text{ext}} = (r_0, \infty)$, $r_0 > r_i > 0$ and impose the Dirichlet condition $u(r_{\text{i}}) = u_{\text{i}}$ on the interior boundary.

The map $u \mapsto \widetilde{u} = r^{-(d-1)/2}u$ transforms (5.2.1) into

$$\partial_r^2 u + \left[k^2 - \frac{\frac{(d-1)^2}{4} - \frac{d-1}{2} + \nu^2}{r^2}\right] u = 0, \tag{5.2.2}$$

where we dropped the tilde on top of u for notational simplicity. Thus we arrive at the same equation on which our theoretical discussion in Section 4.2 is based.

5.2.1. Real Axis Approach. The solution of (5.2.2) includes three subproblems supplying the three parts of the final algebraic system: First, the variational

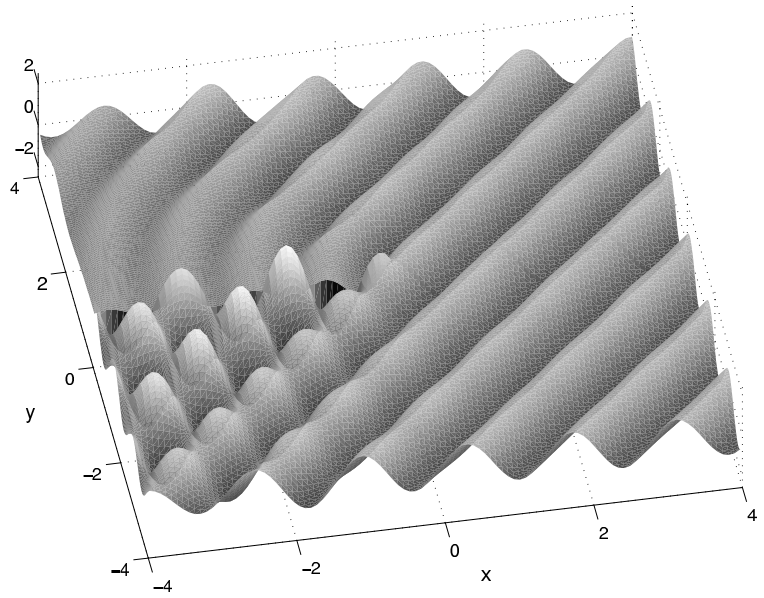


FIGURE 5.1.5. Diffraction of a plane wave with an incident angle of $3/2\pi$ by a rigid semi-infinite plane sheet ($y = 0, x \leq 0$).

formulation of (5.2.2) on Ω_{int}, second, the pole condition in integral form on the real axis linking the exterior and the interior problem such that asymptotically only outgoing waves exist, and, third the Laplace transform of (5.2.2) on Ω_{ext}.

Based on the function space

$$\begin{aligned} V &= \left\{ v \in H^1(\Omega_{\text{int}}) : v(r_{\text{i}}) = u_{\text{i}} \right\} \\ V_0 &= H_0^1(\Omega_{\text{int}}) \end{aligned}$$

(note that $H^1(\Omega_{\text{int}}) \subset C(\Omega_{\text{int}})$) and the sesquilinear form

$$\begin{aligned} a &: V_0 \times V \to \mathbb{C} \\ a(v,u) &= -(\nabla v, \nabla u) + \left(v, \left[k^2 - \frac{\frac{(d-1)^2}{4} - \frac{d-1}{2} + \nu^2}{r^2} + k^2 \right] u \right), \end{aligned}$$

we rewrite (5.2.2) into variational form yielding the first part of our problem.

Part I: (Continuous interior problem): Find $u \in V$ such that for all $v \in V_0$

$$a(v,u) + (\overline{v}u')(r_0) = 0.$$

The interior and the exterior problem are linked together by the pol-condition, which gives the second part of the problem.

Part II: (Continuous pole condition):

$$0 = \int_{\infty}^{-\infty} \mathrm{d}\tau \, \frac{\widehat{u}(\tau)}{\tau - s_+}, \quad \tau \in \mathbb{R}, \operatorname{Im}(s_+) > 0.$$

On the exterior domain we consider the Laplace transform of (5.2.2), where the term u/r^2 is given in its integral operator representation. This equation supplies the third part of the problem.

Part III: (Continuous Laplace transformed exterior problem):

$$\widehat{u}(s) - \frac{su(r_0) + u'(r_0)}{s^2+k^2} - \frac{\frac{(d-1)^2}{4} - \frac{d-1}{2} + \nu^2}{s^2+k^2} \int_s^\infty \mathrm{d}s_1 \, \mathrm{e}^{(s-s_1)r_0} (s-s_1) \widehat{u}(s_1) = 0. \tag{5.2.3}$$

The discretization of the sesquilinear form $a(v,u)$ results in standard manner by the approximation of V by a finite-dimensional space $V_h \subset V$. Let $\dim V_h = n$. Thus we approximate $a(v,u)$ by the symmetric matrix $A_{\mathrm{int}} \in \mathbb{R}^{n\times n}$ and arrive at the discrete interior problem .

Part I: Discrete Interior Problem.

Standard finite element technology supplies the system

$$\mathrm{A}_{\mathrm{int}}\mathbf{u} + \begin{bmatrix} \mathbf{0} \\ (\overline{v}u')(r_0) \end{bmatrix} = \mathbf{r}$$

where $\mathbf{u} \in \mathbb{C}^n$ is the vector of interior unknowns and the right-hand side $\mathbf{r} \in \mathbb{C}^n$ vanishes identically except for the entries arising from the given interior Dirichlet condition.

Part II: Discrete Pole Condition.

To approximate $\widehat{u}$ on the real axis, we define a set of nodes $\{x_i\}_{i=0}^m$ such that

$$-\infty < x_m < x_{m-1} < \ldots < x_1 < x_0 < \infty.$$

On the interval $[x_0, x_m]$ we use the cubic C^1–splines ($i = 1, \ldots, m-1$)

$$\phi_i(x) = \begin{cases} \phi_{i,1}(x) = ((x-x_i)/h_i + 1)^2 (1 - 2(x-x_i)/h_i) & \text{for } x \in [x_{i-1}, x_i] \\ \phi_{i,2}(x) = ((x-x_i)/h_{i+1} - 1)^2 (1 + 2(x-x_i)/h_{i+1}) & \text{for } x \in [x_i, x_{i+1}] \\ 0 & \text{elsewhere,} \end{cases}$$

$$\psi_i(x) = \begin{cases} \psi_{i,1}(x) = ((x-x_i)/h_i + 1)^2 (x-x_i) & \text{for } x \in [x_{i-1}, x_i] \\ \psi_{i,2}(x) = ((x-x_i)/h_{i+1} - 1)^2 (x-x_i) & \text{for } x \in [x_i, x_{i+1}] \\ 0 & \text{elsewhere,} \end{cases}$$

with $h_i = x_i - x_{i-1}$, as basis functions, see Fig. 5.2.1 for a graphical representation.

The support of these basis functions is contained completely in the interval $[x_0, x_m]$. Because we need a representation on the full axis, we define additionally

$$\phi_0(x) = \begin{cases} \phi_{0,1}(x) & \text{for } x \geq x_0 \\ \phi_{0,2}(x) = ((x-x_0)/h_1 - 1)^2 (1 + 2(x-x_0)/h_1) & \text{for } x \in [x_0, x_1] \\ 0 & \text{elsewhere,} \end{cases}$$

$$\psi_0(x) = \begin{cases} \psi_{0,2}(x) = ((x-x_0)/h_1 - 1)^2 (x-x_0) & \text{for } x \in [x_0, x_1] \\ 0 & \text{elsewhere,} \end{cases}$$

for the basis functions crossing x_0 and

$$\phi_m(x) = \begin{cases} \phi_{m,1}(x) = ((x-x_m)/h_m + 1)^2 (1 - 2(x-x_m)/h_m) & \text{for } x \in [x_{m-1}, x_m] \\ \phi_{m,2}(x) & \text{for } x \leq x_m \\ 0 & \text{elsewhere,} \end{cases}$$

$$\psi_m(x) = \begin{cases} \psi_{m,1}(x) = ((x-x_m)/h_m + 1)^2 (x-x_m) & \text{for } x \in [x_{m-1}, x_m] \\ 0 & \text{elsewhere,} \end{cases}$$

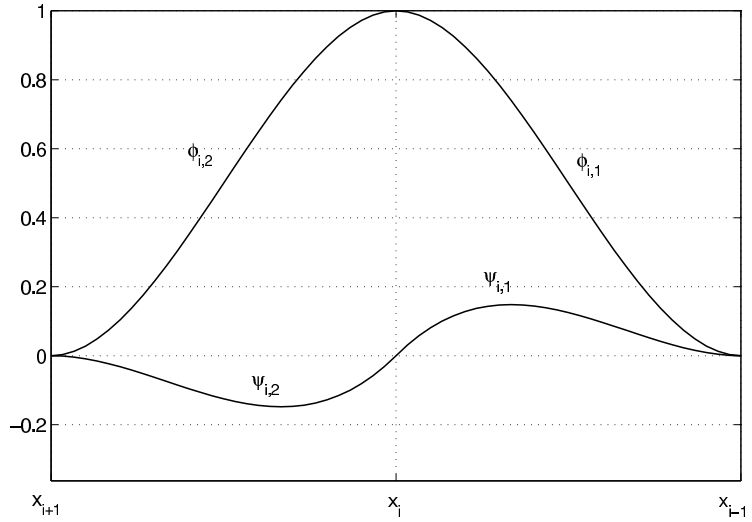


FIGURE 5.2.1. The function ϕ_i is the Lagrange basis function with respect to $\widehat{u}\,(x_i)$, ψ_i is the derivative of the of the Lagrange basis function with respect to $\widehat{u}'\,(x_i)$.

for the basis functions crossing x_m. The functions $\phi_{0,1}(x)$ and $\phi_{m,2}(x)$ are not given explicitely, here. They can be chosen from the asymptotic expressions of the solution of the basic integral equation or can be any other reasonable functions ensuring the continuity conditions $\phi_{0,1}(x_0) = 1$ and $\phi_{m,2}(x_m) = 1$. For simplicity of the presentation we will use simply the constant functions $\phi_{0,1}(x) = 1$ and $\phi_{m,2}(x) = 1$ for x outside the interval $[x_0, x_m]$.

Thus we have *two* independent conditions per node or, equivalently, we may choose two collocation points per interval. This classic spline representation has the feature that the coefficients of $\phi_i\,(x)$ have an immediate interpretation as nodal values of the function at $x = x_i$ and the coefficients of $\psi_i\,(x)$ as the derivatives at these positions. Note, that the interval lengths h_i in our definition are negative. We approximate a complex-valued function $\widehat{u}$ on the real axis by

$$\widehat{u}_h(x) = \sum_{i=0}^{m} \widehat{u}_i \phi_i(x) + \sum_{i=0}^{m} \widehat{u}'_i \psi_i(x). \tag{5.2.4}$$

This requires $2m+2$ coefficients for m intervals. Along with the set of nodes $\{x_i\}_{i=0}^{m}$ we define the set of Gaussian collocation points $\{x_{\mathrm{g}i}\}_{i=1}^{2m}$ by

$$\begin{aligned} x_{\mathrm{g},2i-1} &= \frac{x_i + x_{i-1}}{2} - \frac{h_i}{2\sqrt{3}} \\ x_{\mathrm{g},2i} &= \frac{x_i + x_{i-1}}{2} + \frac{h_i}{2\sqrt{3}}, \; h_i = x_i - x_{i-1}. \end{aligned}$$

Let $\widehat{\mathbf{u}}_{\mathrm{g}} \in \mathbb{C}^{2m}$ be the vector of $\widehat{u}_h$ evaluated at the collocation points. Further, let $\widehat{\mathbf{u}}_{\mathrm{c}} = (\widehat{u}_0, \widehat{u}_1, \widehat{u}'_1, \widehat{u}_2, \widehat{u}'_2, \ldots, \widehat{u}_{m-1}, \widehat{u}'_{m-1}, \widehat{u}_m) \in \mathbb{C}^{2m}$ the vector of spline coefficients. Then, the evaluation of (5.2.4) at all collocation points $x_{\mathrm{g}i}, i = 1, \ldots, 2m$ defines the transform

$$\widehat{\mathbf{u}}_{\mathrm{g}} = \mathrm{M}_{\mathrm{c}} \widehat{\mathbf{u}}_{\mathrm{c}}. \tag{5.2.5}$$

Next we approximate the integral condition by piecewise Gaussian quadratures. We replace the integral condition by a discrete approximation

$$0 = \frac{1}{2}\left(\sum_{i=1}^{m} \frac{h_i}{x_{\mathrm{g},2i-1} - s_+}\widehat{u}_{\mathrm{g},2i-1} + \sum_{i=1}^{m} \frac{h_i}{x_{\mathrm{g},2i} - s_+}\widehat{u}_{\mathrm{g},2i}\right)$$

with $\widehat{u}_{\mathrm{g},j}$, $1 \le j \le 2m$ the jth component of $\widehat{\mathbf{u}}_{\mathrm{g}}$. It is convenient to rewrite this in form of an Euclidian inner product based on the $m \times m$ identity matrix $\mathrm{I}_{m,m}$ and the vector $\mathbf{h} = (h_1, \ldots, h_m)$

$$0 = \frac{1}{2}\left\langle \mathrm{diag}\left(\frac{1}{x_{\mathrm{g}i} - s_+}\right)_{i=1,\ldots,2m} \mathrm{I} \otimes \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mathbf{h},\, \widehat{\mathbf{u}}_{\mathrm{g}} \right\rangle.$$

Based on (5.2.5), we obtain the discrete pole condition in terms of the spline coefficients

$$\frac{1}{2}\left\langle \mathrm{M}_{\mathrm{c}}^{T}\, \mathrm{diag}\left(\frac{1}{x_{\mathrm{g}i} - s_+}\right)_{i=1,\ldots,2m} \mathrm{I}_{m,m} \otimes \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mathbf{h},\, \widehat{\mathbf{u}}_{c} \right\rangle = 0.$$

Part III: Discrete Laplace Transformed Exterior Problem.

The spectral equation (5.2.3) is solved by collocation, i.e., it is rewritten in terms of the collocation coefficients and solved pointwise for any collocation point x_{gi}. The only non-standard problem here is the representation of the integral operator.

We define the complex number $\widehat{I}_i$ by

$$\widehat{I}_i \widehat{u}_h = \int_{x_{\mathrm{g}i}}^{\infty} \mathrm{d}x_1\, \mathrm{e}^{(x_{gi} - x_1)r_0}\, (x_{gi} - x_1)\, \widehat{u}_h(x_1).$$

The evaluation of the expression for each $x_{\mathrm{g}i}$ will supply one row of the full discrete integral operator. Using (5.2.4) we obtain

$$\begin{aligned} \widehat{I}_i \widehat{u}_h = &\int_{x_{gi}}^{\infty} \mathrm{d}x_1\, \mathrm{e}^{(x_{gi} - x_1)r_0}\, (x_{gi} - x_1) \sum_{j=0}^{m} \widehat{u}_j \phi_j(x_1) \\ &+ \int_{x_{gi}}^{\infty} \mathrm{d}x_1\, \mathrm{e}^{(x_{gi} - x_1)r_0}\, (x_{gi} - x_1) \sum_{j=0}^{m} \widehat{u}'_j \psi_j(x_1). \end{aligned}$$

This integral operator might be assembled in a FEM-like manner, i.e., the contributions of each segment is computed separately and added to the global operator. We describe this approach briefly. First we rewrite the representation (5.2.4) such that trial functions with support on the same interval are grouped together

$$\begin{aligned} \widehat{u}_h(x) \quad &= \quad \sum_{i=1}^{m} \widehat{u}_{i-1}\phi_{i-1,2}(x) + \widehat{u}'_{i-1}\psi_{i-1,2}(x) + \widehat{u}_i\phi_{i,1}(x) + \widehat{u}'_i\psi_{i,1}(x) \\ &\quad\; + \widehat{u}_0\phi_{0,1}(x) + \widehat{u}_m\phi_{m,2}(x) \\ &=: \quad \sum_{i=1}^{m} \left[\widehat{u}_{i-1}, \widehat{u}'_{i-1}, \widehat{u}_i, \widehat{u}'_i\right]_{\Sigma}(x) + \widehat{u}_0\phi_{0,1}(x) + \widehat{u}_m\phi_{m,2}(x) \end{aligned}$$

We use the expression in brackets as short-hand for the full expression and remark that $\left[\widehat{u}_{i-1}, \widehat{u}'_{i-1}, \widehat{u}_i, \widehat{u}'_i\right]_{\Sigma}$ refers to the interval $[x_i, x_{i-1}]$. Note, that the boundary conditions $\psi_{0,2}(x) = \psi_{m,1}(x) \equiv 0$ are understood implicitly within this notation. Let x_k be the smallest element of $\{x_i\}_{i=0}^{m}$ larger than a given x_{gi}, from the set of Gaussian collocation points. Then we may write

$$\begin{aligned}\widehat{I}_i\widehat{u}_h &= \int_{x_0}^{\infty} \mathrm{d}x_1\, \mathrm{e}^{(x_{gi}-x_1)r_0}\,(x_{gi}-x_1)\,\widehat{u}_0\phi_{0,1}(x_1)\\ &\quad+\int_{x_k}^{x_0} \mathrm{d}x_1\, \mathrm{e}^{(x_{gi}-x_1)r_0}\,(x_{gi}-x_1)\sum_{j=1}^{k}\left[\widehat{u}_{j-1},\widehat{u}'_{j-1},\widehat{u}_j,\widehat{u}'_j\right]_\Sigma(x_1)\\ &\quad+\int_{x_{gi}}^{x_k} \mathrm{d}x_1\, \mathrm{e}^{(x_{gi}-x_1)r_0}\,(x_{gi}-x_1)\left[\widehat{u}_k,\widehat{u}'_k,\widehat{u}_{k+1},\widehat{u}'_{k+1}\right]_\Sigma(x_1).\end{aligned}$$

The integral expression can be computed element-wise. To this end we introduce

$$\begin{aligned}\widehat{I}_{i0}\widehat{u}_h &:= \int_{x_0}^{\infty} \mathrm{d}x_1\, \mathrm{e}^{(x_{gi}-x_1)r_0}\,(x_{gi}-x_1)\,\widehat{u}_0\phi_{0,1}(x_1)\\ \widehat{I}_{ij}\widehat{u}_h &:= \int_{x_j}^{x_{j-1}} \mathrm{d}x_1\, \mathrm{e}^{(x_{gi}-x_1)r_0}\,(x_{gi}-x_1)\left[\widehat{u}_{j-1},\widehat{u}'_{j-1},\widehat{u}_j,\widehat{u}'_j\right]_\Sigma(x_1)\\ &\quad\text{for } 1\le j\le k\\ \widehat{I}_{i,k+1}\widehat{u}_h &:= \int_{x_{gi}}^{x_k} \mathrm{d}x_1\, \mathrm{e}^{(x_{gi}-x_1)r_0}\,(x_{gi}-x_1)\left[\widehat{u}_k,\widehat{u}'_k,\widehat{u}_{k+1},\widehat{u}'_{k+1}\right]_\Sigma(x_1),\end{aligned}$$

which allows us to compute the contribution of the ith Gauss-point as

$$\widehat{I}_i\widehat{u}_h = \widehat{I}_{i0}\widehat{u}_h + \sum_{j=1}^{k}\widehat{I}_{ij}\widehat{u}_h + \widehat{I}_{i,k+1}\widehat{u}_h.$$

The indices of $\widehat{I}_{ij}$ becomes easier readable if we remember that the first index, i, counts the number of the Gauss-point ($1\le i\le 2m$) and the second index, j, denotes the *end* point of the corresponding interval $[x_{j-1},x_j]$, $1\le j\le m$. The element-wise contributions can be computed exactly or by numerical quadrature. On the interval $x\ge x_0$ we continue $\widehat{u}_h$ by a constant function – as discussed above – and set $\phi_{0,1}=1$. Since $\widehat{u}=\mathcal{O}\left(1/|x|\right)$ for $x\to\infty$ on the real axis, this introduces an error of the order $\mathcal{O}\left(1/|x_0|^2\right)$. Of course this may be improved essentially by using the known asymptotic approximations of $\widehat{u}$. Further, as it is seen, the entries decay *exponentially* away from the ith segment. We associate to each number $\widehat{I}_{i0}\widehat{u}_h$ a complex number $\widehat{\mathrm{I}}_{i0}$ and to each number $\widehat{I}_{ij}\widehat{u}_h$ a local 1×4 matrix $\widehat{\mathrm{I}}_{ij}$ through

$$\begin{aligned}\widehat{\mathrm{I}}_{i0} &:= \int_{x_0}^{\infty} \mathrm{d}x_1\, \mathrm{e}^{(x_{gi}-x_1)r_0}\,(x_{gi}-x_1)\,\phi_{0,1}(x_1)\\ \widehat{\mathrm{I}}_{ij} &:= \int_{x_j}^{x_{j-1}} \mathrm{d}x_1\, \mathrm{e}^{(x_{gi}-x_1)r_0}\,(x_{gi}-x_1)\left[\phi_{j-1,2}(x_1),\psi_{j-1,2}(x_1),\phi_{j,1}(x_1),\psi_{j,1}(x_1)\right]\\ &\quad\text{for } 1\le j\le k\\ \widehat{\mathrm{I}}_{i,k+1} &:= \int_{x_{gi}}^{x_k} \mathrm{d}x_1\, \mathrm{e}^{(x_{gi}-x_1)r_0}\,(x_{gi}-x_1)\left[\phi_{k,2}(x_1),\psi_{k,2}(x_1),\phi_{k+1,1}(x_1),\psi_{k+1,1}(x_1)\right],\end{aligned}$$

The resulting discrete integral operator $\widehat{\mathrm{I}}$ is composed from the local matrices $\widehat{\mathrm{I}}_{ij}$. The corresponding procedure is given in Algorithm 1. This way we assemble a matrix $\widehat{\mathrm{I}}\in\mathbb{C}^{2m\times(2m+2)}$. We arrive at the desired matrix if we delete the columns corresponding to ψ_0 (the column 2) and to ψ_m (the last column). Further we replace the vector of nodal values at the Gaussian points $\widehat{\mathbf{u}}_\mathrm{g}$ by the vector of unknown spline coefficients $\widehat{\mathbf{u}}_\mathrm{c}$ according to $\widehat{\mathbf{u}}_\mathrm{g}=\mathrm{M}_\mathrm{c}\widehat{\mathbf{u}}_\mathrm{c}$,(5.2.5), and abbreviate $c_\nu=-\left(d-1\right)^2/4+(d-1)/2-\nu^2$ to obtain

Algorithm 1 FEM-like assembling of the integral operator

for $k = 1$ to m **do**
  for $j = k$ to m **do**
    $l = 2j - 1$; $n = 2j - 2k + 1$
    $\widehat{\mathrm{I}}(l, n : n+3) = \widehat{\mathrm{I}}(l, n : n+3) + \widehat{\mathrm{I}}_{l,n-1}$
    $l = l + 1$
    $\widehat{\mathrm{I}}(l, n : n+3) = \widehat{\mathrm{I}}(l, n : n+3) + \widehat{\mathrm{I}}_{l,n-1}$
  end for
end for
for $l = 1$ to $2m$ **do**
  $\widehat{\mathrm{I}}(l, 1) = \widehat{\mathrm{I}}(l, 1) + \widehat{\mathrm{I}}_{l,0}$
end for

$$\mathrm{M_c}\widehat{\mathbf{u}}_\mathrm{c} - \left[\frac{x_{\mathrm{g}i}}{x_{\mathrm{g}i}^2 + k^2}\right]_{i=1,\ldots,2m} u(r_0) - \left[\frac{1}{x_{\mathrm{g}i}^2 + k^2}\right]_{i=1,\ldots,2m} u'(r_0) + c_\nu \mathrm{diag}\left(\frac{1}{x_{\mathrm{g}i}^2 + k^2}\right) \widehat{\mathbf{I}}\widehat{\mathbf{u}}_\mathrm{c} \quad = \quad \mathbf{0}.$$

All three contributions of the composite problem are assembled into an algebraic system of the form $\mathrm{A}\mathbf{x} = \mathbf{b}$ and solved simultaneously. This is summarized in Algorithm 2, and the structure of the composite system matrix is displayed in Fig. 5.2.2.

Algorithm 2 Discrete operator with pole condition in spectral form

1. FEM discretization of the interior problem

$$A_{11} = A_{\mathrm{int}} + \begin{bmatrix} \mathbf{0} \\ (\overline{v}u')(r_0) \end{bmatrix}$$

2. Discretization of the pole condition

$$A_{22} = \begin{bmatrix} \frac{1}{2} & \mathrm{M_c}^T \left(\mathrm{diag}\left(\frac{1}{x_{\mathrm{g}i} - s_+}\right)_{i=1,\ldots,m} I_{m,m} \otimes \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mathbf{h}\right)^T \end{bmatrix}$$

3. Discretization of the Laplace transformed exterior equation

$$A_{32} = \begin{bmatrix} -\mathbf{g}_0, & -\mathbf{g}', & \mathrm{M_c} + \mathrm{c}_\nu \mathrm{diag}(\mathbf{g}')\widehat{\mathbf{I}} \end{bmatrix}$$

with $\mathbf{g}_0 = \left[\frac{x_{\mathrm{g}i}}{x_{\mathrm{g}i}^2 + k^2}\right]_{i=1,\ldots,2m}$ and $\mathbf{g}' = \left[\frac{1}{x_{\mathrm{g}i}^2 + k^2}\right]_{i=1,\ldots,2m}$

4. Assemble the system matrix A

$$A = \begin{bmatrix} A_{11} & \mathbf{0} \\ \mathbf{0} & A_{22} \\ \mathbf{0} & A_{32} \end{bmatrix}$$

Numerical studies. To test the numerical performance of the procedure, we define the following test problem: $r_\mathrm{i} = 3$, $r_0 = 3 + 2\pi$, dimension $d = 2$, $k = 1$. The interior problem is discretized by n linear finite elements, the real axis for the discretization of the Laplace-transformed equation by N collocation points, symmetrically placed around the origin. The real axis is discretized uniformly in the interval $[-2, 2]$ by $N/4$ of all points and with geometrically increasing segment lengths outside this interval. The full interval on the real axis were chosen to be $[-5/8\,N, 5/8\,N]$, hence its size increases if the number of collocation points increases. Fig. 5.2.3 shows the relative error of the computed DtN-number as a

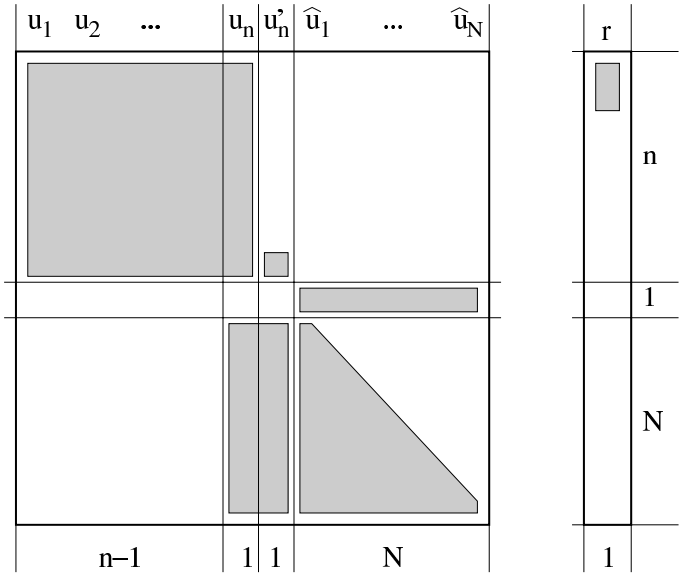


FIGURE 5.2.2. Occupation pattern of the composite matrix. The upper left block matrix is the sparse, real and symmetric FE-matrix, the lower right block matrix is empty in its $(N-3)\times(N-3)$ upper triangular part. The number of interior degrees of freedom is n, the number of collocation points is $N = 2m$, with m the number of segments on the real axis.

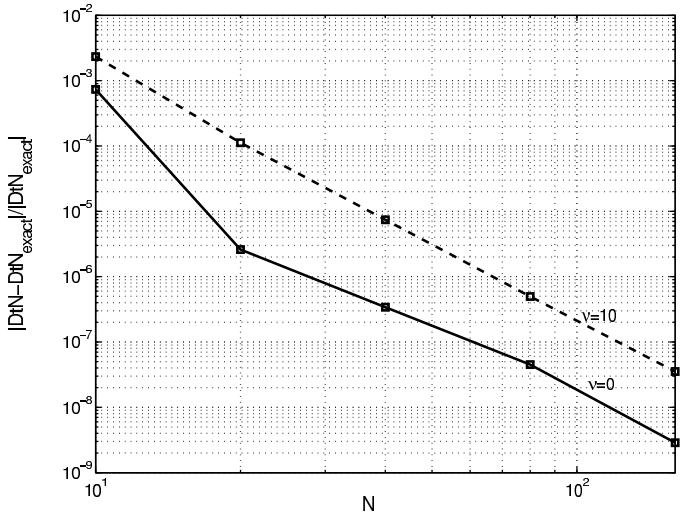


FIGURE 5.2.3. Convergence of the algorithm with respect to the relative error of the DtN-number as a function of the number of collocation points N on the real axis.

function of the number of collocation points N on the real axis for the frequencies $\nu = 0$ and $\nu = 10$. The relative error has been computed by setting $u(r_0) = 1$, i.e. without to solve the interior problem, in order to obtain the result independent from the interior problem. The result shows that even with a relative small number of spectral collocation points ($N \leq 160$) a good accuracy can be achieved. It

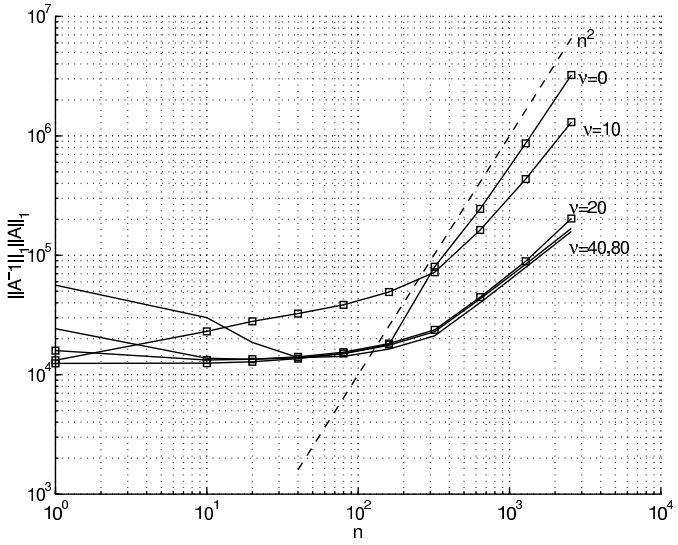


FIGURE 5.2.4. Condition number with respect to the 1-norm of the resulting matrix vs. the number of interior degrees of freedom n.The number of spectral points is kept fix with $N = 160$. The dashed line displays the function n^2 indicating the condition number of the discrete Laplacian.

should be noted that this result was obtained with an explicit restriction of the triangular integral operator matrix to a bandwidth of $N/4$. A far better compression might be obtained using adaptive discretization techniques and more elaborated compression techniques. Fig. 5.2.4 shows the evolution of the relative condition number of the composite system matrix A. The important results are that, first, higher frequencies ν do not cause an explosion of the condition number and, second, that the condition number is asymptotically dominated by the condition number of the corresponding discrete Laplacian of the interior problem. Hence the stability properties of the discrete Helmholtz scattering problem and the discrete Laplace problem are comparable.

This is underlined by the computation of the spectral norm of the inverse of the system matrix A. To ensure discrete solvability we want that the norm of the inverse remains bounded. Fig. 5.2.5 shows the computed results as a function of the number of interior nodes, where the number of spectral collocation points is kept to $N = 160$. The figure shows that the computed norms are even getting smaller as the number of unknown increases until it approaches a certain constant.

Final remarks. So far, we have not specified, how to carry out the integration of the spline elements. The numerical experiments demonstrated above used closed formulas obtained via a symbolic computation. This becomes possible, since all the involved integrals are of elementary nature (exponential functions times polynomials). For an efficient implementation, however, this should be replaced by numerical quadrature formulas. There is a variety of possible choices. The next section which deals with the implementation of the cut-function approach, presents one of the possibilities: the realization based on Runge-Kutta type integrators. This is a very efficient approach for the present case, too. Since we will discuss the details in the next section, we drop it here. However, in order to compare the structure of

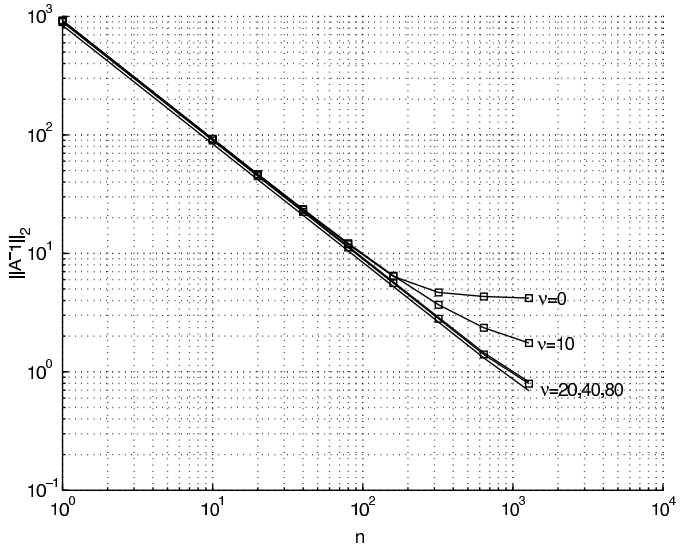


FIGURE 5.2.5. Spectral norm of the inverse of the resulting matrix vs. the number of interior degrees of freedom.

both methods, we give the pole condition and the integral equation using $\widetilde{\tau} = -s$, $\widehat{u}(s) = \widehat{u}(-\widetilde{\tau}) =: \widetilde{u}(\widetilde{\tau})$ in a slightly different form

$$\begin{aligned} \widetilde{u}(\widetilde{\tau}) + \frac{\widetilde{\tau} u(r_0) - u'(r_0)}{\widetilde{\tau}^2 + k^2} + c_\nu \int_{-\infty}^{\widetilde{\tau}} \mathrm{d}\tau_1 \, e^{(\tau_1 - \widetilde{\tau}) r_0} (\tau_1 - \widetilde{\tau}) \, \widetilde{u}(\tau_1) &= 0 \\ \int_{-\infty}^{\infty} \mathrm{d}\widetilde{\tau} \, \frac{\widetilde{u}(\widetilde{\tau})}{\widetilde{\tau} + s_+} &= 0 \end{aligned}$$

We add that in general the Runge-Kutta type realization of the Volterra integral equation –unlike the ODE-case–is stable only for *explicit* schemes if large step sizes are used. This results from the fact that implicit methods have to compute $\exp(a(\tau - \widetilde{\tau}))$ as part of the kernel for $\tau > \widetilde{\tau}$, which causes large entries in the discrete approximation and leads finally to extremely large condition numbers. Finally we remark that we implemented successfully the real axis approach to the Laplace domain method using the explicit Runge-Kutta methods from the explicit Euler scheme ($p = 1$) to the Dormand-Prince RK5(4) scheme ($p = 5$).

5.2.2. Cut Function Approach. We derive an algorithm, which applies the results of Chapter 4, especially of Sections 4.3, 4.4.2 (pp. 74) in a direct manner. This approach possesses the essential advantage to supply the far field without any extra numerical costs. We use the same model problem as in the foregoing section. For convenience, we repeat it here. We wish to solve Bessel's equation (5.2.1)

$$\partial_r^2 u + \frac{d-1}{r} \partial_r u + k^2 u - \frac{\nu^2}{r^2} u = 0$$

on the real half-axis $\{r \in \mathbb{R}_+ : r > r_i\}$. At r_i a Dirichlet condition is given. Bessel's equation is transformed to (5.2.2)

$$\partial_r^2 u + \left[k^2 + \frac{c_\nu}{r^2} \right] u = 0$$

with $c_\nu = -(d-1)^2/4 + (d-1)/2 - \nu^2$ and a new function u. We solve the transformed equation simultaneously on the domains $\Omega_{\mathrm{int}} = (r_{\mathrm{i}}, r_0)$ and $\Omega_{\mathrm{ext}} = (r_0, \infty)$,

$r_0 > r_i > 0$. Our strategy is the same as in Section 5.2, i. e., we decompose the problem into three subproblems: The interior problem (Part I), the linking conditions between the interior end the exterior problem (Part II), and the exterior problem in spectral form (Part III). First we give the three subproblems in continuous form.

Part I: (Continuous interior problem): Find $u \in V$ such that for all $v \in V_0$

$$a(v,u) + (\overline{v}u')(r_0) = 0,$$

Part II: (Continuous linking conditions).

$$\begin{aligned} u(r_0) - \frac{w(i)}{2i} + \frac{1}{2\pi i}\int_0^\infty dt\,\widehat{\psi} &= 0 \\ u'(r_0) - \frac{w(i)}{2} + \frac{1}{2\pi i}\int_0^\infty dt\,(i-t)\widehat{\psi} &= 0. \end{aligned}$$

These are special forms of our representation formulas, evaluated at r_0. Note, that we used the map $w(i)\psi \mapsto \widehat{\psi}$ to define a new function $\widehat{\psi}$ in spectral domain. Further, these representations suggest to use the pole strength $w(i)$ as additional unknown for the numerical implementation. Finally, the restriction $\widehat{\psi} = \widehat{\psi}_+$, which means the identification of the spectral solution representation with the pure outgoing one, forces the whole solution to be purely outgoing. This plays the same role as the direct implementation of the pole condition as extra condition in our real axis approach, compare Section 5.2.

Part III: (Continuous Laplace transformed exterior problem):

$$\begin{aligned} \widehat{\psi}(t) - \frac{\pi c_\nu}{t-2i} e^{-r_0 t} w(i) \\ + \frac{c_\nu}{t(t-2i)} \int_0^t \mathrm{dt}_1\, \mathrm{e}^{(t-t_1)r_0}\,(t-t_1)\,\widehat{\psi}(t_1) = 0. \end{aligned} \tag{5.2.6}$$

In this form, the similarity between both approaches, the direct and the cut function approach, becomes apparent. Note the close similarity between the spectral equation (5.2.3) and (5.2.6).

Next we derive the corresponding discrete equations.

Part I: Discrete Interior Problem.

The discrete interior problem is given again canonically :

$$A_{\mathrm{int}}\mathbf{u} + \begin{bmatrix} \mathbf{0} \\ (\overline{v}u')(r_0) \end{bmatrix} = \mathbf{r}$$

where $\mathbf{u} \in \mathbb{C}^n$ is the vector of interior unknowns and the right-hand side $\mathbf{r} \in \mathbb{C}^n$ vanishes identically except for the entries arising from the given interior Dirichlet condition. Next we discuss the discretization of $\widehat{\psi}$ on the real axis $t \geq 0$.

Part II: Discrete Linking Conditions.

Our implementation of the cut function formulation follows a completely different approach than used in the real axis method. Whereas we applied a collocation method based on cubic C^1-splines in a FEM-like manner in the latter one, we develop the cut function method based on quadrature formulas for ordinary differential equations (ODE's). This is motivated by the observed strong stability of the integral equation (5.2.6) when solved as ODE in forward ($+\infty$) direction. We extend the ODE approach to the integral formulation based on explicit and implicit Runge-Kutta-type collocation schemes. Following the motivation given in **[27]** and the presentation of Runge-Kutta-type collocation schemes given in **[25]**,

we derive in the following the Runge-Kutta schemes for our basic linear Volterra integral equation.

Let us denote (5.2.6) by

$$\widehat{\psi}(t) = f(t) + \int_0^t dt_1\, K(t,t_1)\widehat{\psi}(t_1)$$

with the kernel $K(t,t_1)$ and $f(t)$ as given above

$$\begin{aligned} K(t,t_1) &= -\frac{c_\nu}{t(t-2i)}\mathrm{e}^{(t-t_1)r_0}\,(t-t_1) \\ f(t) &= \frac{\pi c_\nu}{t-2i}e^{-r_0 t}w(i). \end{aligned}$$

We define the set of nodes $\{t_i\}_{i=0}^m$

$$0 = t_0 < t_1 < \ldots < t_m < \infty$$

which decompose the positive real axis into m intervals between t_0 and t_m. Let us denote $\mathbf{t} = (t_0,\ldots,t_{m-1})^T$. First, we consider a solution between the points t_l and t_{l+1}, $0 \le l < m$ and write

$$\begin{aligned} \widehat{\psi}(t_l+\tau) &= f(t_l+\tau) + \int_0^{t+\tau} dt_1\, K(t_l+\tau,t_1)\widehat{\psi}(t_1) \\ &= \sum_{n=0}^{l-1}\int_{t_n}^{t_{n+1}} dt_1\, K(t_l+\tau,t_1)\widehat{\psi}(t_1) \\ &\quad + \int_{t_l}^{t_l+\tau} dt_1\, K(t_l+\tau,t_1)\widehat{\psi}(t_1) + f(t_l+\tau). \end{aligned} \tag{5.2.7}$$

The relation to the integration of ODE's becomes apparent, if we consider the simplified sub-problem

$$\widetilde{\psi}(t_l+\tau) = \widetilde{\psi}(t_l) + \int_{t_l}^{t_l+\tau} dt_1\, \widetilde{K}(t_1)\widetilde{\psi}(t_1).$$

This is equivalent to the ODE

$$\frac{d}{d\tau}\widetilde{\psi}(t+\tau) = \widetilde{K}(t+\tau)\widetilde{\psi}(t+\tau), \qquad \widetilde{\psi}(t) = \psi_0.$$

Now, a general $s_{RK}-$stage Runge-Kutta scheme with consistency order p has the form (cf. e. g. **[25**, Sec. 6.2, Eq. 6.9])

$$\begin{aligned} g_i &= \widetilde{\psi}(t_l) + \tau_l \sum_{j=1}^{s_{RK}} a_{ij} K\left(t_l + c_j\tau_l\right) g_j, \qquad i = 1,\ldots,s \\ \widetilde{\psi}(t_l+\tau) &= \widetilde{\psi}(t_l) + \tau \sum_{j=1}^{s_{RK}} b_j K\left(t_l + c_j\tau_l\right) g_j + \mathcal{O}\left(\tau_l^{p+1}\right) \qquad \tau_l = t_{l+1} - t_l. \end{aligned}$$

As it is common practice, we use the table (Butcher scheme)

$$\begin{array}{c|c} \mathbf{c} & \mathrm{A} \\ \hline & \mathbf{b} \end{array}$$

with $(\mathrm{A})_{ij} = a_{ij}$, $\mathbf{c} = (c_1,\ldots,c_{s_{RK}})$, $\mathbf{b} = (b_1,\ldots,b_{s_{RK}})$ to characterize a Runge-Kutta method. Following Engl **[27]**, we adopt the method to solve our integral

equation. The form of the Runge-Kutta scheme in the form given above, suggests a direct generalization: taking the dependence of the kernel K from a second variable into account as well as the fact that f is given explicitely, we obtain the discretization of (5.2.7)

$$\begin{aligned}\widehat{\psi}(t_l+\tau_l c_i) \quad \approx \quad & \sum_{n=0}^{l-1}\tau_n\sum_{j=1}^{s_{RK}} b_j K(t_l+\tau_l c_i, t_n+\tau_n c_j)\,\widehat{\psi}(t_n+\tau_n c_j)\\ & +\tau_l\sum_{j=1}^{s_{RK}} a_{ij} K(t_l+\tau_l c_i, t_l+\tau_l c_j)\,\widehat{\psi}(t_l+\tau_l c_j)\\ & +f(t_l+\tau_l c_i)\,.\end{aligned}$$

For the numerical implementation it is useful to define a vector τ containing all intermediate time-steps

$$\tau = \mathrm{I}_{\mathrm{m},m}\otimes \mathrm{I}_{\mathrm{sRK},1}\mathbf{t} + \Delta t\otimes \mathbf{c}$$

with Δt the vector of interval lengths $\Delta t = (t_1-t_0, t_2-t_1, \ldots, t_m-t_{m-1})$ and the kernel-matrix

$$\begin{aligned}(\mathrm{K})_{pq} \quad = \quad & K(\tau_p,\tau_q)\,, \qquad \tau_p=(\tau)_p\,,\ \tau_q=(\tau)_q\\ & K_{pq}=0 \qquad \text{for } p+s_{\mathrm{RK}}-2+p \bmod s_{\mathrm{RK}} < q.\end{aligned}$$

Then, the integral operator component of the discrete system can be assembled by the sum of a sub-diagonal block matrix representing the quadrature formulas and a block-diagonal matrix corresponding to the Runge-Kutta scheme. Let T be a lower triangular matrix, whose sub-diagonal part consists only of unit entries

$$\mathrm{T} = \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ \vdots & \ddots & \ddots & \\ 1 & \cdots & 1 & 0 \end{bmatrix}, \qquad \mathrm{T}\in\mathbb{R}^{m,m}.$$

The discretized integral operator can be realized in vectorized form:

$$\mathrm{M} = \operatorname{diag}\Delta t \ \otimes \mathrm{A}\,.*\,\mathrm{K} + (\mathrm{T}\operatorname{diag}\Delta t)\otimes(\mathrm{I}_{\mathrm{sRK},1}\otimes\mathbf{b})\,.*\,\mathrm{K}.$$

The MatLab-type dot-multiplication $\mathrm{A}.*\mathrm{B}$ denotes element-wise multiplication of the elements of matrices of equal dimension: $(\mathrm{A}.*\mathrm{B})_{ij} = (\mathrm{A})_{ij}\cdot(\mathrm{B})_{ij}$. This makes an implementation both simple and effective. For the study of the numerical performance we always used the implicit Runge-Kutta scheme of Gaussian type with $p=4$, in order to have the possibility to compare the cut function approach with our real axis method. The corresponding Butcher scheme is given by

$$\begin{array}{c|cc} \frac{1}{2}-\frac{\sqrt{3}}{6} & \frac{1}{4} & \frac{1}{4}-\frac{\sqrt{3}}{6} \\ \frac{1}{2}+\frac{\sqrt{3}}{6} & \frac{1}{4}+\frac{\sqrt{3}}{6} & \frac{1}{4} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}$$

However, all the other common Gauss and Radau schemes from $p=1$ (implicit Euler method) to the implicit Gauss-method with $p=6$ have been implemented

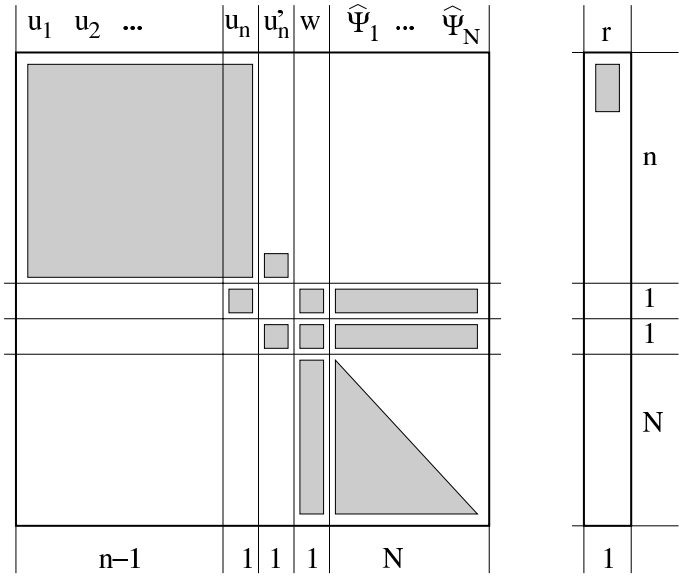


FIGURE 5.2.6. Structure of the algebraic system corresponding to the cut-function approach. Compare Fig. 5.2.2.

and tested successfully. Based on these components, we can express the linking conditions in discrete form

$$\begin{aligned} u(r_0) - \frac{w(i)}{2i} + \frac{1}{2\pi i}(\Delta \mathbf{t})^T \otimes \mathbf{b}\,\widehat{\psi} &= 0 \\ u'(r_0) - \frac{w(i)}{2} + \frac{1}{2\pi i}(i-\tau)^T .* \left[(\Delta \mathbf{t})^T \otimes \mathbf{b}\right]\,\widehat{\psi} &= 0. \end{aligned}$$

Part III: Discrete Laplace Transformed Exterior Problem.
Finally, we obtain the equation for the cut-function along the real axis

$$\widehat{\psi} - \frac{\pi c_\nu}{\tau - 2i} e^{-r_0 \tau} w(i) - \mathrm{M}\widehat{\psi} = 0.$$

The second term in this equation is a vector whose entries has to be evaluated according to the entries of τ.

Numerical studies. Fig. 5.2.6 shows the occupation pattern of the composite block-matrix system. Its structure is very similar to the one obtained for the real axis approach.

Fig. 5.2.7 presents the convergence result of the cut-function approach. The relative error of the DtN-number for the frequencies $\nu = 0$ and $\nu = 10$ are displayed versus the number of steps of the Runge-Kutta method in the spectral domain. This and all the following results refer to implicit Gauss-method with $p = 4$. It is seen that the accuracy of the high-frequency mode is lower than the of the zero-frequency mode. Compared with the real axis approach, Fig. 5.2.3, we note the accuracy of the zero-frequency mode in the cut-function approach at the same step numbers is better than in the real axis approach, whereas the situation is reversed for the high-frequency modes.

The essential advantage of the cut-function approach is that it supplies the far-field without additional costs. The corresponding convergence results are displayed

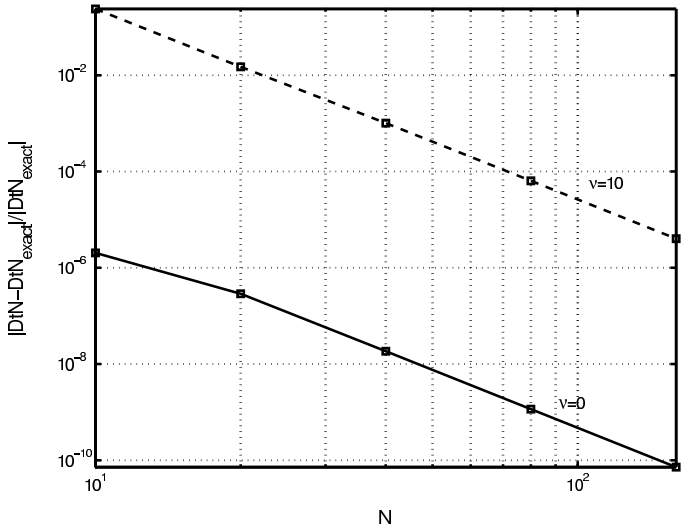


FIGURE 5.2.7. Convergence of the DtN-number in the cut-function approach vs. the number of intervals in the spectral domain.

in Fig. 5.2.8. As far-field we understand here the coefficient of first term of the asymptotic expansion of the Hankel function, normalized such that it matches the value $u(r_i) = 1$. Again we consider the modes $\nu = 0$ and $\nu = 10$. Obviously, there is no essential improvement in the accuracy if we decrease the step length (increase the number of intervals). The reason is that the accuracy of the far-field is bounded by the accuracy of the interior problem. To show this, we refined the interior problem uniformly, ending up with a far-field accuracy improved by a factor 1/4. The accuracy of the far-field of the high-frequency mode is much worse than for the low-frequency mode, as expected. Nevertheless, it converges linearly in the log-log plot. The pole strength of the mode, hence the far field amplitude, depends exponentially on the frequency as $\nu \to \infty$ if the amplitude on the inner radius of the interior problem is fixed. This result is verified numerically, as shown in Fig. 5.2.9.

The stability of the discrete problem is analyzed as in the real axis approach. First, Fig. 5.2.10 supplies the condition numbers of the discrete problem for different frequencies. Unlike the real axis approach, the condition number grows rapidly with the frequency, until a saturation is reached. If the number of interior nodes is large enough, the condition number of the interior problem dominates the condition number of the composite problem.

The norm of the inverse discrete operator remains bounded as the number of degrees of freedom in the spectral domain, as seen in Fig. 5.2.11. However, for a small number of discretization points, that is far away from the continuous solution, the norm is larger than for a finer discretization. One possibility to reduce the size of the condition numbers for high frequencies is to normalize the pole strength $w(i)$ by the frequency. In Fig. 5.2.12 we used the normalization $w(i) \mapsto \widetilde{w}(i) := w(i)/\left(1+\nu^2\right)$. This results in a very effective reduction of the condition number and shows a way how to handle this property in higher dimensional cases.

5.3. Non-Separable, Discrete Problems in 2D

In the following we describe a canonical method to derive a finite set of equations approximating the exterior Helmholtz equation by a semi-discrete ansatz.

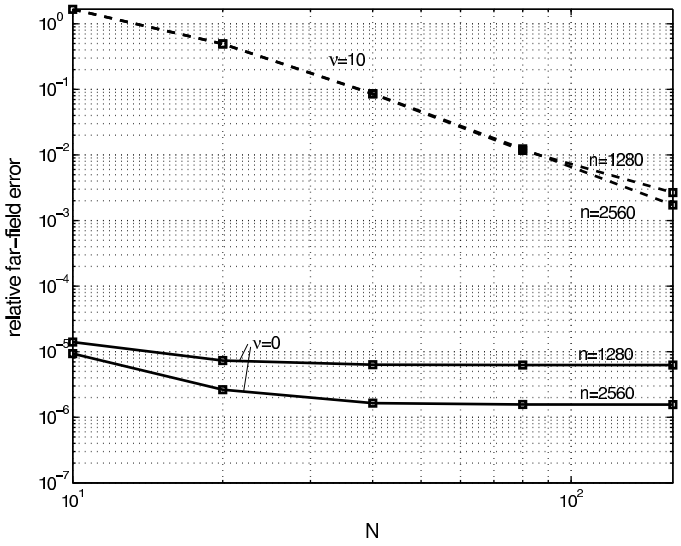


FIGURE 5.2.8. Convergence of the far-field in the cut-function approach vs. the number of intervals in the spectral domain for two different discretizations of the interior problem.

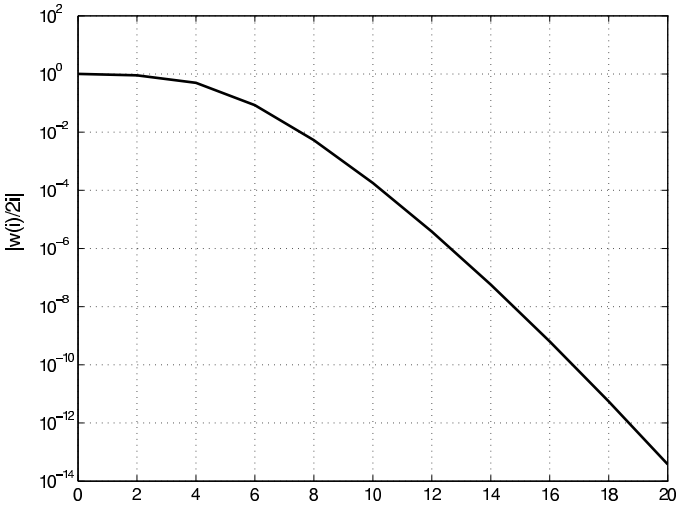


FIGURE 5.2.9. Pole strength as a function of the frequency ν

The technique is comparable to the method of (vertical) lines for the solution of parabolic problems , cf. the book of Johnson **[62**, Chapt. 8.3].

Basic Idea.

First, we decompose the exterior domain into a finite number of segments, as it is shown in the left-hand side of Fig. 5.3.1. The decomposition is based on straight rays connecting each vertex of the polygonal boundary $\partial\Omega_{\text{int}}$ with infinity. The rays do not intersect each other. In each semi-infinite quadrilateral obtained this way we approximate the continuous function $u(x, y)$ by a function $u_h(x, y)$, which is discrete in in y, using a local ansatz

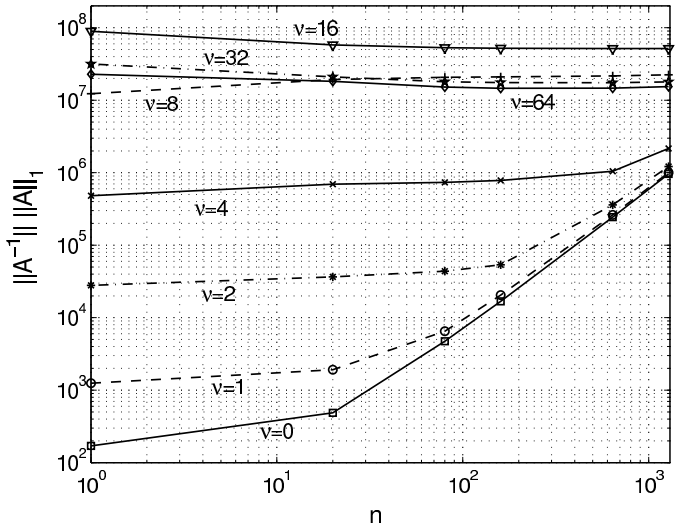


FIGURE 5.2.10. Condition of the composite discrete problem for different frequencies ν.

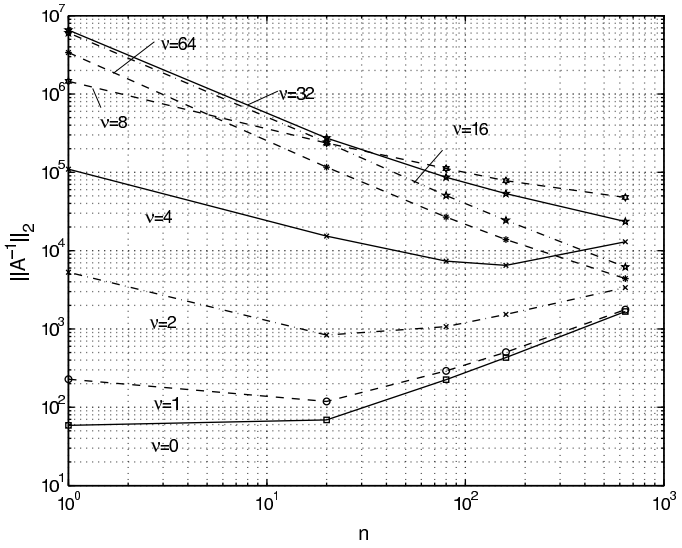


FIGURE 5.2.11. Spectral norm of the inverse discrete operator

$$u_h(x,y) = v_1(x,y)u_1(x) + v_2(x,y)u_2(x),$$

where $u_1(x)$ and $u_2(x)$ correspond to the rays $y_1(x)$ and $y_2(x)$, see Fig. 5.3.2. This generalizes the *global* separability used in Sections 5.1–5.2.2 to a *local* separability inside the considered segment. The gray marked quadrilateral in the left-hand side of Fig. 5.3.1 shows that, in our example, the variable x plays the role of a distance variable and the variable y behaves as an angular-like variable. The functions $v_1(x,y)$ and $v_2(x,y)$ are given linear trial functions defined by

$$v_1(x,y) = \frac{y_2(x)-y}{h(x)} \qquad v_2(x,y) = \frac{y-y_1(x)}{h(x)} \qquad h(x) = y_2(x) - y_1(x).$$

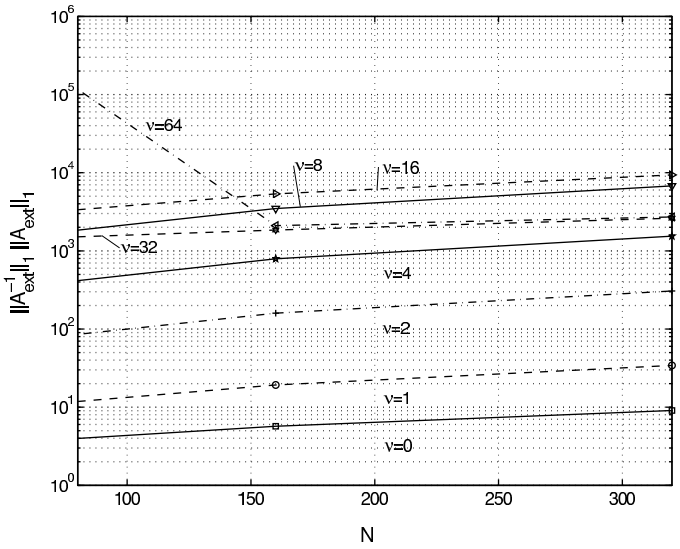


FIGURE 5.2.12. Condition number of the normalized exterior sub-problem.

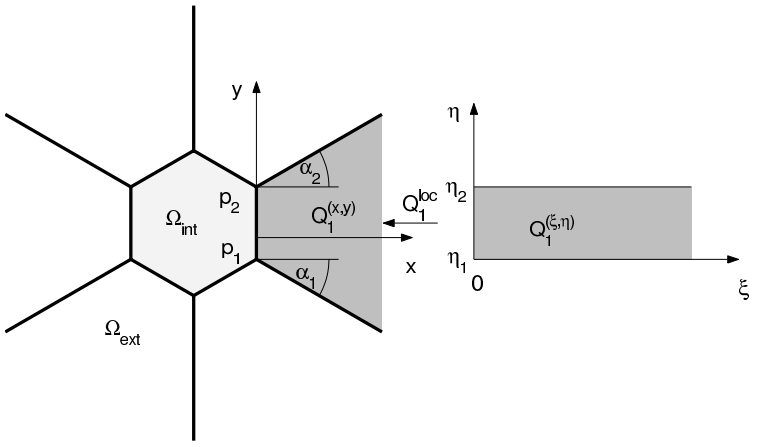


FIGURE 5.3.1. Discretization of the exterior domain by straight rays and mapping from the semi-infinite rectangular domain

This discretization can be extended to curved rays and higher order approximations in the angular-like variable in an obvious way. To obtain a global approximation we have to fit the *local* trial-functions defined on different quadrilaterals such that the resulting *global* trial function is continuous, hence exists at least in $H^1_{\mathrm{loc}}(\Omega_{\mathrm{ext}})$. The whole construction can be done in a FEM-like manner based on the transform of semi-infinite elements. In the following we make this idea more precise.

Exterior Domain in Transformed Coordinates. We proceed similarly as in Section 4.1. Let Ω_{int} be a polygonally interior domain. Let $\Omega_{\mathrm{ext}} = \mathrm{ext}\,\Omega_{\mathrm{int}}$ the corresponding exterior domain and $\partial\Omega_{\mathrm{int}}$ its polygonal, simple and closed boundary

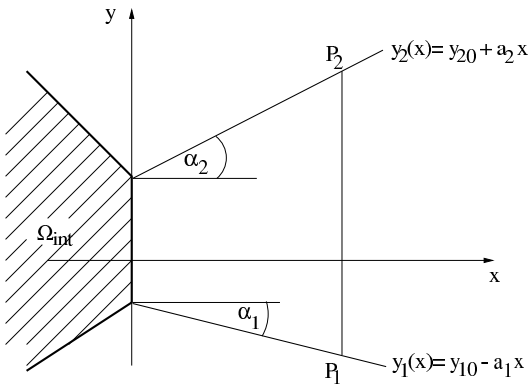


FIGURE 5.3.2. Linear rays belonging to nodes on the boundary ($a_{1,2} = \tan a_{1,2}$).

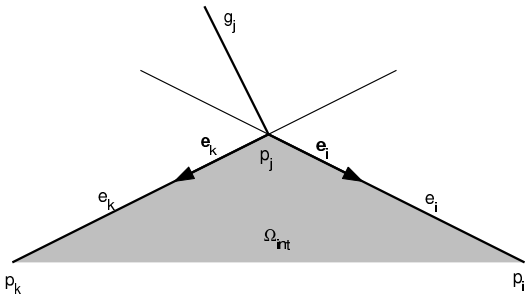


FIGURE 5.3.3. Construction of the ray g_j associated to vertex p_j if the interior angle $\angle(e_i, e_k) < \pi$

with vertices $\mathcal{P} = \{p_1, p_2, \ldots, p_N\}$. Each vertex p_j is associated with a straight ray g_j.

DEFINITION 5.3.1. The set $G = \{g_1, g_2, \ldots, g_N\}$ is an admissible set of straight rays, if it meets the following conditions for any $j \in 1, \ldots, N$:

(1) *(Linearity.)* Let p_j be given. Let the edges e_i and e_k , $i, k \in 1, \ldots, N$, $i \neq k$, have p_j as common vertex. Let the second vertex of e_i be p_i and the second vertex of e_k be p_k. Define two unit vectors $\mathbf{e}_i = (p_i - p_j) / |p_i - p_j|$ and $\mathbf{e}_k = (p_k - p_j) / |p_k - p_j|$ such that they are directed away from p_j.
   (a) If the segment $\overline{p_i p_k}$ intersects Ω_{int} (convexity), then the ray g_j must have a representation $g_j(\tau) = p_j + \tau (c_i \mathbf{e}_i + c_k \mathbf{e}_k)$ with $\tau \in \mathbb{R}_+$ and both c_i and c_j strictly negative, compare Fig. 5.3.3.
   (b) If the segment $\overline{p_i p_k}$ intersects $\operatorname{ext} \Omega_{\text{int}}$ (concavity), then the same representation must hold with both c_i and c_j strictly positive, compare Fig. 5.3.4.
   (c) If p_i, p_j, p_k are placed on a common straight line, then the ray g_j must have a representation $g_j(\tau) = p_j + \tau (c_i \mathbf{e}_i + c_n \mathbf{n})$ with $\mathbf{n}$ the unit normal vector directed outwards and c_n is strictly positive.

(2) *(Intersection condition.)* The rays neither intersect each other in the exterior domain nor they have common points with Ω_{int}. The only common points of the rays and $\overline{\Omega_{\text{int}}}$ are the vertices of the polygonal boundary.

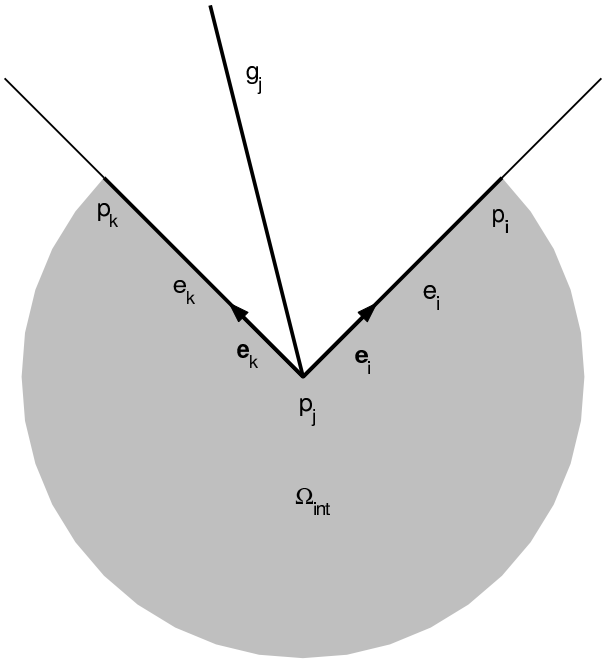


FIGURE 5.3.4. Construction of the ray g_j associated to vertex p_j if the interior angle $\angle(e_i, e_k) > \pi$

(3) *(Geometric similarity.)* There exist positive constants $\gamma_1, \ldots, \gamma_N$ such that for all $\tau > 0$ the points $p'_1 = g_1(\gamma_1\tau), \ldots, p'_N = g_N(\gamma_N\tau)$ are the vertices of a new simple polygon, where the vertices are connected in the same order as the corresponding vertices of the original polygon, and the edges of the new polygon and the corresponding edges of the original polygon are *parallel* to each other, cf. Fig. 5.3.5.

Condition 3 is the crucial condition. It enables a direct application of the finite element technology to the exterior domain. We give two schemes which allow the construction of sets of admissible rays for convex and for certain types of non-convex domains. These non-convex domains are star-shaped domains which we want to define as follows:

DEFINITION 5.3.2. A domain Ω is star-shaped with respect to a point $p \in \Omega$, if all rays starting at p hit the boundary $\partial\Omega$ in exactly one point.

1. Radial rays. Let a nonempty convex domain be given, compare Fig. 5.3.5. Fix an arbitrary interior point. Connect this interior point by line segments with each of the vertices of the boundary. Extend these line segments to linear rays. This scheme can be extended to star-shaped non-convex domains as follows. Let a star-shaped non-convex domain be given, as it is shown in Fig. 5.3.6. Hence it exists an interior point such that any line segment which connects this interior point with a vertex of the boundary hits the boundary only at this vertex, and we can continue this line segment to the desired ray. By geometrical similarity, the radial ray constructions supplies an admissible set of rays.

2. Generalized normal rays. Let a nonempty convex domain be given, compare Fig. 5.3.7. Mark all vertices of the boundary whose aligned edges enclose an interior angle less than π. These are all given vertices $p_1, \ldots, p_4$ in Fig. 5.3.7. Construct the rays successively corresponding to all but the last marked vertex according to

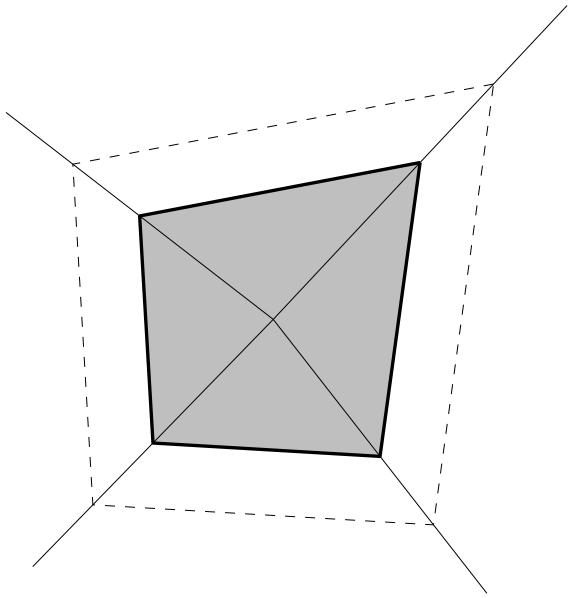


FIGURE 5.3.5. Radial ray construction for convex domains. The dashed line indicates the linearly scaled polygonal boundary. The scaling is based on a single scaling parameter, compare Definition 5.3.1, 3.

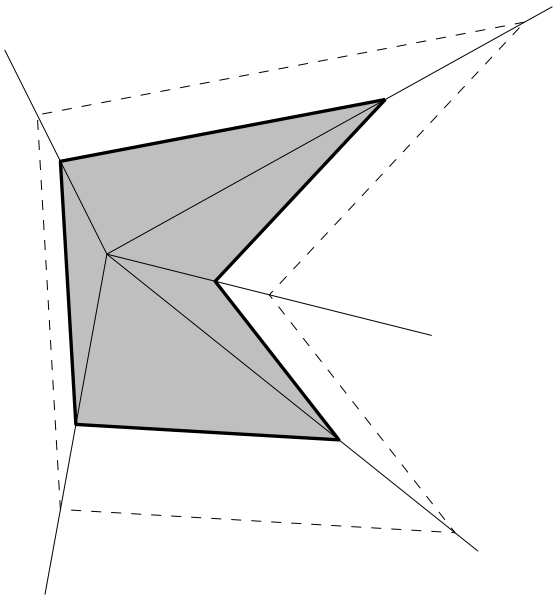


FIGURE 5.3.6. Radial ray construction for a star-shaped concave domain. Compare Fig. 5.3.5.

Definition 5.3.1, 1a). This corresponds to the rays $g_1, \ldots, g_3$ in our example. The last ray g_N (corresponding to g_4 in Fig. 5.3.7) must be constructed in a special manner:

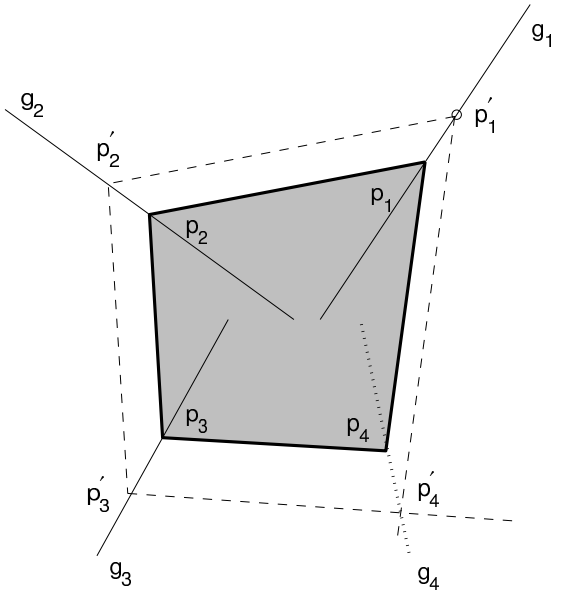


FIGURE 5.3.7. Normal ray construction for convex domains. They rays do not have a common intersection point. The last ray constructed is marked by the dotted line. It must ensure that the scaled polygonal boundary becomes a closed curve.

(1) Fix an arbitrary point on the ray through the first marked vertex (in the example point p_1').
(2) Construct the two lines which go trough this point, and which are parallel to their corresponding boundary segments (the two dashed lines intersecting in p_1').
(3) Determine the intersection point between this line and the next ray (the point p_2').
(4) Continue the construction, starting from the intersection point and moving in positive direction.
(5) The last ray must be constructed such that it goes through the last marked vertex and the intersection of the first and the last line constructed this way (the dotted ray in our example).

Construct the remaining rays according to Definition 5.3.1, 1c) and 2). The notion "normal" arises from the fact that the generalized normal rays tend to be scaled normal vectors of the boundary if the boundary tends to be smooth curve *and* the rays are constructed such that they halve the exterior angle between the aligned edges. Other schemes, e.g. a combination of the normal and radial ray scheme, are possible. A direct consequence from the geometric construction is the following

LEMMA. *Both normal and radial ray construction supply an admissible set of straight rays.*

A decomposition $\mathcal{Q} = \{Q_1, Q_2, \ldots, Q_N\}$ of $\overline{\Omega_{\text{ext}}}$ based on N semi-infinite quadrilaterals $Q_{1,\ldots,N}$, where each quadrilateral is bounded by two rays and one edge of $\partial\Omega_{\text{int}}$, provides the starting point for both the variational formulation and the finite element discretization. Fig. 5.3.1 illustrates the basic scheme.

DEFINITION 5.3.3. A decomposition $\mathcal{Q}$ based on straight lines (rays) is called admissible, if the following properties are met:

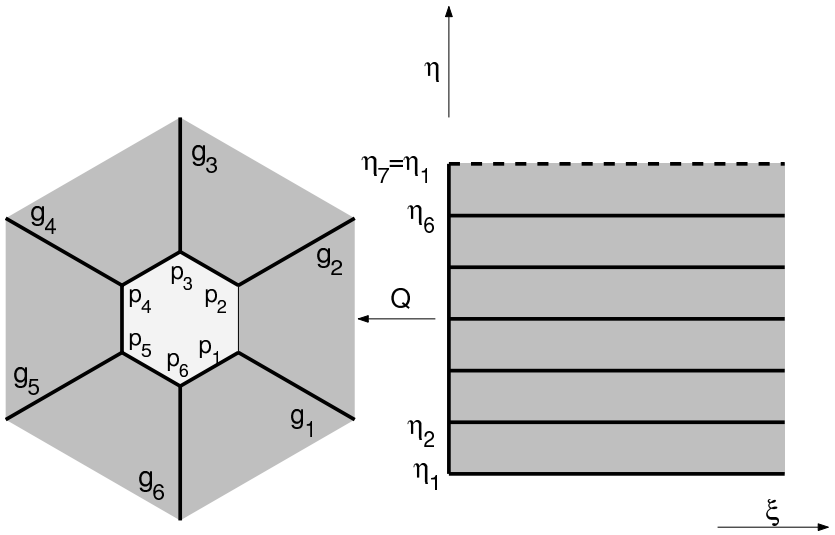


FIGURE 5.3.8. Mapping from the semi-infinite rectangular reference domain onto the exterior of the given problem. The dashed line indicates that the η-interval continuous periodically.

(1) Two nodes of each Q_j are also nodes of $\partial\Omega_{\text{int}}$.
(2) $\bigcup_{j=1}^{N} Q_j = \overline{\Omega_{\text{ext}}}$.
(3) If $i \neq j$, then the intersection $Q_i \cap Q_j$ is either empty or $Q_i \cap Q_j = \mathrm{g}_j(x,y)$, where $g_j(x,y)$ is a straight ray connecting the vertex p_j of $\partial\Omega_{\text{int}}$ with infinity. This boundary vertex is a common point of Q_i and Q_j.

We consider each semi-infinite quadrilateral as image of a semi-infinite rectangle under a appropriate diffeomorphism. Let us denote the jth semi-infinite rectangle by $Q_j^{(\xi,\eta)}$, where ξ and η are the Cartesian coordinates of the pre-image – ξ playing the role of a generalized distance variable and η a role as a generalized angular variable – and by $Q_j^{(x,y)}$the jth semi-infinite quadrilateral defined above. For clarity, we added the superscript (x,y) to indicate that the image is defined in the original x,y-coordinates of the Cartesian reference systems. We denote the jth diffeomorphism by

$$Q_j^{(\text{loc})} \,:\, Q_j^{(\xi,\eta)} \to Q_j^{(x,y)}.$$

These local mappings are connected with each other such that a global homeomorphism

$$Q \,:\, \overline{\Omega_{\text{ext}}^{(\xi,\eta)}} \to \overline{\Omega_{\text{ext}}^{(x,y)}}$$

results. Moreover, we connect the local domains and mappings in a way that the global mapping Q meets the following crucial conditions

(1) Q is at least twice continuously differentiable in ξ.
(2) Q is continuous and piecewise continuously differentiable in η.

The latter fact is important since it allows a standard variational formulation in η with finite-dimensional test and trial spaces based on continuous C^0- elements. Fig. 5.3.8 displays this global mapping in a schematic diagram.

Variational Formulation in Transformed Coordinates.

First we restate the variational formulation of the interior problem (4.0.9) in a slightly modified form. Without interior boundaries and sources, but with a given sufficiently regular normal derivative $\mathbf{n}(s)\nabla u(s)$ on $\partial\Omega_{\text{int}}$, the variational form of the interior problem reads: Find $u \in H^1(\Omega_{\text{int}})$ such that for all $v \in H^1(\Omega_{\text{int}})$

$$\text{(5.3.1)} \qquad \textbf{Interior problem}: \qquad a(v,u) = \int_{\partial\Omega_{\text{int}}} \mathrm{d}s\overline{v}(s)\mathbf{n}(s)\nabla u(s).$$

Again, $a(v,u)$ denotes our standard sesquilinear form

$$\begin{aligned} a &: H^1(\Omega_{\text{int}}) \times H^1(\Omega_{\text{int}}) \to \mathbb{C} \\ a(v,u) &= (\nabla v, \nabla u) - \left(v, k^2(\mathbf{x})\, u\right). \end{aligned}$$

In the following, we derive a corresponding variational formulation of the exterior problem formulated in transformed coordinates. The essential difference as compared to the interior problem is that the variational (weak) form is used in η-direction only, whereas in ξ-direction the original (strong) form of the Helmholtz equation is used. In its core, this is a method of lines type formulation, which *formally* represents the exterior problem as elliptic initial value problem with initial data given on $\partial\Omega_{\text{ext}}^{(\xi,\eta)}$. Given three sesquilinear forms

$$a_i(v,u) : H_\pi^1[\eta_{\min}, \eta_{\max}] \times W^i\left(\Omega_{\text{ext}}^{(\xi,\eta)}\right) \to \mathbb{C}, \quad i \in \{0,1,2\},$$

which will be construct below, the variational form of the exterior problem reads: Find $u \in W^2\left(\Omega_{\text{ext}}^{(\xi,\eta)}\right)$ such that for all $v \in H_\pi^1[\eta_{\min}, \eta_{\max}]$ and $\xi \in \mathbb{R}_+$

$$\text{(5.3.2)} \quad \textbf{Exterior problem}: \qquad \begin{aligned} a_0(v,u) + a_1(v, \partial_\xi u) + a_2\left(v, \partial_\xi^2 u\right) &= 0 \\ u(0) &= u_D \\ u'(0) &= u_N. \end{aligned}$$

The function spaces $W^i\left(\Omega_{\text{ext}}^{(\xi,\eta)}\right)$ might be defined as

$$W^i\left(\Omega_{\text{ext}}^{(\xi,\eta)}\right) := \left\{ \begin{array}{l} w(\xi_0,\eta) \in H_\pi^1[\eta_{\min}, \eta_{\max}] : \xi_0 \in \mathbb{R}, \text{ fixed and} \\ w(\xi,\eta_0) \in C^i(\mathbb{R}_+) : \eta_0 \in [\eta_{\min}, \eta_{\max}], \text{ fixed} \end{array} \right\}, \quad i \in \{0,1,2\}.$$

The construction of $a_i(\cdot,\cdot)$ proceeds as follows:

(1) The local mappings $Q_j^{(\text{loc})}$, $j = 1, \ldots, N$ from the ξ,η-system to the x,y-system are derived.
(2) The local mappings are glued together to obtain the global mapping Q : $\overline{\Omega_{\text{ext}}^{(\xi,\eta)}} \to \overline{\Omega_{\text{ext}}^{(x,y)}}$.
(3) The Helmholtz equation is represented in ξ,η-coordinates.
(4) The weak form with respect to η is derived.

The following geometrical consideration supplies the jth local mapping $Q_j^{(\text{loc})}$, compare Fig. 5.3.1. Let the set of vertices $\mathcal{P}$ be assigned to the set $\{\eta_1, \ldots, \eta_N\}$ of generalized angular coordinates. The values of η_j might be derived from the arc length of the boundary (for 2D problems) or from the angle coordinate of a cylindric coordinate system. Let us consider the semi-infinite quadrilateral $Q_j^{(x,y)}$ as marked in Fig 5.3.1 by the gray area. It is bounded by the segment $\overline{p_1p_2}$, $p_1, p_2 \in \mathcal{P}$, and the rays g_1, g_2. Let us denote the unit vectors defining the direction of the rays g_1, g_2 by $\mathbf{g}_1$ and $\mathbf{g}_2$. Since $g_1, g_2 \in \mathcal{G}$, they possess a parameter representation

$$\begin{aligned} g_1(\tau) &= p_1 + \tau \mathbf{g}_1 \\ g_2(\tau) &= p_2 + \gamma\tau \mathbf{g}_2,\ \tau \in \mathbb{R}_+ \end{aligned}$$

and it must hold $\overline{g_1(\tau)\, g_2(\tau)} \parallel \overline{p_1 p_2}$, compare the definition of the admissible rays Definition 5.3.1. Let the unit vector $\mathbf{n}$ be normal to $\overline{p_1 p_2}$. According to Definition 5.3.1 , we have to choose γ such that $\overline{g_1(\tau)\, g_2(\tau)} \perp \mathbf{n}$, it follows $\gamma = \cos\alpha_1 / \cos\alpha_2$. So far, τ is an arbitrary real parameter. We want to replace it by a distance variable ξ with a direct geometrical meaning. We define ξ as the distance between the straight line through p_1 and p_2 and the straight line through $g_1(\tau)$ and $g_2(\tau)$. It can be seen from Fig. 5.3.1 that $\xi = \tau \cos\alpha_1$. This way we obtain obtain a symmetric form of the parameter representation

$$\begin{aligned} g_1(\xi) &= p_1 + \frac{\xi}{\zeta \cos\alpha_1}\mathbf{g}_1 \\ g_2(\xi) &= p_2 + \frac{\xi}{\zeta \cos\alpha_2}\mathbf{g}_2,\ \xi \in \mathbb{R}_+, \end{aligned}$$

where the new scaling parameter is needed to fit the different local coordinate systems continuously to each other. We define the signs of the angles $\alpha_{1,2}$ such that they are positive if the corresponding rays and the lengthened normal through the mid-point of boundary segment diverge in the semi-plane right from the lengthened boundary segment.

So far, a point given on g_1 (or on g_2) is uniquely related to the generalized distance ξ. Next we assign a given point *between* the rays g_1 and g_2 to a generalized distance ξ *and* a generalized angle η. To this end let η on g_1 be equal to η_1 and η on g_2 equal to η_2. We perform a linear interpolation between these values inside $Q_j^{(x,y)}$. Consequently, a point $p(x,y) \in Q_j^{(x,y)}$can be expressed by

$$p(x,y) = g_1(\xi) + \frac{\eta - \eta_1}{\eta_2 - \eta_1}\left(g_2(\xi) - g_1(\xi)\right).$$

Further, with unit vectors $\mathbf{g}_1$ and $\mathbf{g}_2$ given in Cartesian coordinates

$$\mathbf{g}_1 = \begin{pmatrix} \cos\beta_1 \\ \sin\beta_1 \end{pmatrix} \quad \mathbf{g}_2 = \begin{pmatrix} \cos\beta_2 \\ \sin\beta_2 \end{pmatrix},$$

we can express $p(x,y)$ in Cartesian coordinates too, which at the same time supplies the desired mapping

$$\begin{aligned} \text{(5.3.3)} \qquad \begin{pmatrix} x \\ y \end{pmatrix} &=: Q_j^{(\mathrm{loc})}(\xi,\eta) \quad \text{with} \\ Q_j^{(\mathrm{loc})}(\xi,\eta) &= \left(1 - \frac{\eta - \eta_1}{\eta_2 - \eta_1}\right)\left[\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \frac{\xi}{\zeta\cos\alpha_1}\begin{pmatrix} \cos\beta_1 \\ \sin\beta_1 \end{pmatrix}\right] \\ &\quad + \frac{\eta - \eta_1}{\eta_2 - \eta_1}\left[\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} + \frac{\xi}{\zeta\cos\alpha_2}\begin{pmatrix} \cos\beta_2 \\ \sin\beta_2 \end{pmatrix}\right]. \end{aligned}$$

LEMMA 5.3.4. *The mapping $Q_j^{(\mathrm{loc})} : Q_j^{(\xi,\eta)} \to Q_j^{(x,y)}$ is a local C^∞-diffeomorphism.*

PROOF. We refer to the local inverse mapping theorem, cf. [**96**, Theorem 4.F, pp. 259]. We have to show

(i) Q_j^{loc} is a C^∞-map on some neighborhood of all points $(\xi_0,\eta_0) \in Q_j^{(\xi,\eta)}$.

(ii) $\left(Q_j^{(\mathrm{loc})}\right)'(\xi_0,\eta_0) : \mathbb{R}^2 \to \mathbb{R}^2$ is bijective.

Ad (i). The map $Q_j^{(\mathrm{loc})}(\xi,\eta)$ is bilinear in ξ and η. Hence the property is obvious.

Ad (ii). The property is invariant with respect to translation and rotation. Hence, without loss of generalization, we may choose $x_1 = x_2 = 0$, $y_1 = 0$, $y_2 = h$, $\eta_1 = 0$, $\eta_2 = h$, $\beta_1 = -\alpha_1$ and $\beta_2 = \alpha_2$. Then the jth local mapping simplifies to

$$\begin{aligned} Q_j^{(\mathrm{loc})}(\xi,\eta) &= \left(1-\frac{\eta}{h}\right)\frac{\xi}{\zeta}\begin{pmatrix} 1 \\ -\tan\alpha_1 \end{pmatrix} \\ &\quad + \frac{\eta}{h}\left[\begin{pmatrix} 0 \\ h \end{pmatrix} + \frac{\xi}{\zeta}\begin{pmatrix} 1 \\ \tan\alpha_2 \end{pmatrix}\right]. \end{aligned}$$

Introducing $\mathbf{h} := (d\xi, d\eta)^T$ we compute

$$\left(Q_j^{(\mathrm{loc})}\right)' \mathbf{h} = Q_j^{(\mathrm{loc})}(d\xi,\eta_0) + Q_j^{(\mathrm{loc})}(\xi_0,d\eta)$$

and it follows,

$$J_j := \left(Q_j^{(\mathrm{loc})}\right)' = \begin{pmatrix} \frac{1}{\zeta} & 0 \\ -\frac{1-\eta_0}{\zeta}\tan\alpha_1 + \frac{\eta_0}{\zeta}\tan\alpha_2 & 1 + \frac{\xi_0}{h\zeta}(\tan\alpha_1 + \tan\alpha_2) \end{pmatrix}.$$

The determinant of the jth Jacobian J_j becomes

$$|J_j| = \frac{h\zeta + \xi_0(\tan\alpha_1 + \tan\alpha_2)}{h\zeta^2}.$$

By construction, $\alpha_1, \alpha_2 \geq 0$, and both h and ζ are strictly positive. Hence $|J_j|$ is always positive, and the map is bijective. □

Next we glue two neighbored domains $\overline{Q}_j^{(\xi,\eta)}$ and $\overline{Q}_{j+1}^{(\xi,\eta)}$ together and set the free scaling parameters ζ_j and ζ_{j+1} such that the resulting mapping from the interior of the combined ξ,η-domains onto the corresponding combined x,y-domains remains C^∞ with respect to ξ and becomes globally C^0 with respect to η. We denote the interior of the composite domains by $Q_{j,j+1}^{(\xi,\eta)}$ and $Q_{j,j+1}^{(x,y)}$,respectively, and label the quantities α, β and ζ occurring in different subdomains by the index of the subdomain.

LEMMA 5.3.5. *Let $\overline{Q}_j^{(\xi,\eta)}$ and $\overline{Q}_{j+1}^{(\xi,\eta)}$ be two neighbored domains, and $\overline{Q}_j^{(x,y)}$ and $\overline{Q}_{j+1}^{(x,y)}$ their counterparts, separated by the ray g_{j+1}. The composite mapping*

$$Q_{j,j+1}^{(\mathrm{loc})} : \mathrm{int}\left(\overline{Q}_j^{(\xi,\eta)} \cup \overline{Q}_{j+1}^{(\xi,\eta)}\right) \to \mathrm{int}\left(\overline{Q}_j^{(x,y)} \cup \overline{Q}_{j+1}^{(x,y)}\right)$$

composed of the local mappings (5.3.3) for domains j and $j+1$ is a homeomorphism if and only if

$$\zeta_{j+1}\cos\alpha_{1,j+1} = \zeta_j \cos\alpha_{2,j}. \tag{5.3.4}$$

Moreover $Q_{j,j+1}^{(\mathrm{loc})}$ is C^∞ with respect to ξ.

PROOF. We consider a neighborhood of the ray g_{j+1}. The ray possesses the representations

$$\begin{aligned} g_{j+1}(\xi) &= p_{j+1} + \frac{\xi}{\zeta_j\cos\alpha_{2,j}}\mathbf{e}_{j+1}, \text{ with } \xi \in \mathbb{R}_+ \\ \text{and } g_{j+1}(\xi) &= p_{j+1} + \frac{\xi}{\zeta_{j+1}\cos\alpha_{1,j+1}}\mathbf{e}_{j+1}. \end{aligned}$$

Thus the condition $\zeta_{j+1}\cos\alpha_{1,j+1} = \zeta_j \cos\alpha_{2,j}$ ensures the continuity

$$\lim_{\eta\uparrow\eta_{j+1}} Q_j^{(\mathrm{loc})}(\xi,\eta) = \lim_{\eta\downarrow\eta_{j+1}} Q_{j+1}^{(\mathrm{loc})}(\xi,\eta) \text{ for all } \xi \in \mathbb{R}_+.$$

Furthermore, it holds

$$\lim_{\eta\uparrow\eta_{j+1}} \partial_\xi Q_j^{(\mathrm{loc})}(\xi,\eta) = \frac{1}{\zeta_j \cos\alpha_{2,j}} \begin{pmatrix} \cos\beta_{j+1} \\ \sin\beta_{j+1} \end{pmatrix}$$

as well as

$$\lim_{\eta\downarrow\eta_{j+1}} \partial_\xi Q_{j+1}^{(\mathrm{loc})}(\xi,\eta_j) = \frac{1}{\zeta_{j+1} \cos\alpha_{1,j+1}} \begin{pmatrix} \cos\beta_{j+1} \\ \sin\beta_{j+1} \end{pmatrix}.$$

This shows that in the limit $\eta \to \eta_{j+1}$ the derivatives of the local maps with respect to the generalized distance variable ξ coincide at both sides of the ray $g_{j+1}$to any order, if the angle condition (5.3.4) is met. □

Having established the local mapping properties between the domains $Q_j^{(\xi,\eta)}$ and $Q_j^{(x,y)}$ and the mapping properties of the composite mapping between two neighboring domains, we obtain the desired global mapping property.

COROLLARY 5.3.6. *Let* $\overline{Q}_{\mathrm{ext}}^{(\xi,\eta)} = \cup_{j=1}^N \overline{Q}_j^{(\xi,\eta)}$ *and* $\overline{Q}_{\mathrm{ext}}^{(x,y)} = \cup_{j=1}^N \overline{Q}_j^{(x,y)}$. *The composite mapping*

$$Q : \overline{\Omega_{\mathrm{ext}}^{(\xi,\eta)}} \to \overline{\Omega_{\mathrm{ext}}^{(x,y)}}$$

composed of the local mappings (5.3.3) for domains $j = 1,\ldots,N$ *is a homeomorphism if one scaling parameter of the set* $\{\zeta_1,\ldots,\zeta_N\}$ *is given and the other satisfy the angle condition of Lemma 5.3.5.*

Sesquilinear Forms in Transformed Coordinates.

The point of departure is the Laplacian in ξ,η-coordinates (4.1.2), which allows us to rewrite the Helmholtz equation in ξ,η-coordinates. To simplify notation, we do not introduce a new symbol for the solution u represented in ξ,η-coordinates. The transformed Helmholtz equation reads

$$\nabla_{\xi\eta}^T \left(J^{-1}J^{-T}\,|J|\,\nabla_{\xi\eta}\right) u + |J|\,k^2 u = 0.$$

Let us abbreviate $F := J^{-1}J^{-T}\,|J|$. Let $V\,[\eta_{\min},\eta_{\max}]$ denote the trial space of periodic, real functions v on the interval $[\eta_{\min},\eta_{\max}]$. We multiply the transformed Helmholtz equation with $v \in V$, where V is a suitable space of test functions periodic on $[\eta_{\min},\eta_{\max}]$ and integrate over the interval $\Gamma_{\eta=}[\eta_{\min},\eta_{\max}]$

$$\int_{\Gamma_\eta} \mathrm{d}\eta\, v(\eta)\,(\partial_\xi,\,\partial_\eta)\,F\nabla_{\xi\eta}u + \int_{\Gamma_\eta} \mathrm{d}\eta\;|J|\,k^2 v(\eta)u = 0.$$

We rewrite this with a coordinate representation of the transposed gradient separating its ∂_ξ- and ∂_η-part

$$\begin{aligned}(5.3.5)\qquad &\int_{\Gamma_\eta} \mathrm{d}\eta\, v(\eta)\,(\partial_\xi,\,0)\,F \begin{pmatrix} \partial_\xi u \\ \partial_\eta u \end{pmatrix} \\ &\quad + \int_{\Gamma_\eta} \mathrm{d}\eta\, v(\eta)\,(0,\,\partial_\eta)\,F \begin{pmatrix} \partial_\xi u \\ \partial_\eta u \end{pmatrix} + \int_{\Gamma_\eta} \mathrm{d}\eta\, v(\eta)\,|J|\,k^2 u = 0.\end{aligned}$$

Denoting the elements of the matrix F by F_{ij}, $i,j = 1,2$ we compute, with respect to the first term of (5.3.5),

$$\begin{aligned}
&\int_{\Gamma_\eta} \mathrm{d}\eta\, v(\eta)\,(\partial_\xi,\, 0)\, F\,(\partial_\xi u,\, \partial_\eta u)^T \\
&\quad = \int_{\Gamma_\eta} \mathrm{d}\eta\, v(\eta)\,[F_{11}\partial_{\xi\xi}u + F_{12}\partial_{\xi\eta}u + (\partial_\xi F_{11})\,\partial_\xi u + (\partial_\xi F_{12})\,\partial_\eta u] \\
&\quad = \int_{\Gamma_\eta} \mathrm{d}\eta\, v(\eta)\,[F_{11}\partial_{\xi\xi}u + (\partial_\xi F_{11})\,\partial_\xi u + (\partial_\xi F_{12})\,\partial_\eta u] - \int_{\Gamma_\eta} \mathrm{d}\eta\,\partial_\eta\,(vF_{12})\,\partial_\xi u.
\end{aligned}$$

Integration by parts supplies for the second term of (5.3.5)

$$\int_{\Gamma_\eta} \mathrm{d}\eta\, v(\eta)\,(0,\, \partial_\eta)\, F\,(\partial_\xi u,\, \partial_\eta u)^T \;=\; -\int_{\Gamma_\eta} \mathrm{d}\eta\,\partial_\eta v\,[F_{21}\partial_\xi u + F_{22}\partial_\eta u]\,.$$

Here we used the fact that the elements of the space V are periodic functions, so no boundary terms occur. Collecting the terms which correspond to the same derivatives of u, we find the desired sesquilinear forms with $v = v(\eta)$, $u = u(\xi, \eta)$ as follows

$$\begin{aligned}
(5.3.6)\quad a_2(v,\, \partial_{\xi\xi}u) &= \int_{\Gamma_\eta} \mathrm{d}\eta\, vF_{11}\partial_{\xi\xi}u \\
(5.3.7)\quad a_1(v,\, \partial_\xi u) &= \int_{\Gamma_\eta} \mathrm{d}\eta\, v\partial_\xi F_{11}\partial_\xi u - \int_{\Gamma_\eta} \mathrm{d}\eta\,\left[\partial_\eta\,(vF_{12}) + (\partial_\eta v)\,F_{21}\right]\partial_\xi u \\
(5.3.8)\quad a_0(v,\, u) &= \int_{\Gamma_\eta} \mathrm{d}\eta\, v\partial_\xi F_{12}\partial_\eta u - \int_{\Gamma_\eta} \mathrm{d}\eta\,\partial_\eta v F_{22}\partial_\eta u \\
&\quad + \int_{\Gamma_\eta} \mathrm{d}\eta\, v\,|J|\,k^2 u.
\end{aligned}$$

Based on these forms, we derive a FEM-discretization along the η-coordinate in standard fashion.

Finite Element Discretization.

Let us denote the semi-discretized function $u(\xi, \eta)$ by $u_h(\xi, \eta)$. We introduce the discrete, *local* separation ansatz

$$u_h(\xi, \eta) = \sum_{j=1}^{N} v_j(\eta) u_j(\xi)$$

with functions $v_j \in V$, thus employing a discretization of Galerkin-type. We consider the sesquilinear forms (5.3.6)-(5.3.8) one after the other. It follows from the local separation ansatz that

$$\begin{aligned}
\mathrm{A}_{2,ij}\partial_{\xi\xi}u_j &:= a_2\,(v_i,\, v_j\partial_{\xi\xi}u_j) \\
&= \left(\int_{\Gamma_\eta} \mathrm{d}\eta\, v_i F_{11} v_j\right)\partial_{\xi\xi}u_j.
\end{aligned}$$

Hence we obtain

$$\mathrm{A}_{2,ij} = \int_{\Gamma_\eta} \mathrm{d}\eta\, v_i v_j F_{11}.$$

The same way we get

$$\begin{aligned}
\mathrm{A}_{1,ij}\partial_\xi u_j &:= a_1(v_i,\, v_j\partial_\xi u_j) \\
&= \left(\int_{\Gamma_\eta} \mathrm{d}\eta\, v_i\partial_\xi F_{11} v_j - \int_{\Gamma_\eta} \mathrm{d}\eta\, \partial_\eta\left(v_i F_{12} + v_i F_{21}\right) v_j\right)\partial_\xi u_j \\
&= \left(\int_{\Gamma_\eta} \mathrm{d}\eta\, v_i\left[\partial_\xi F_{11} v_j + F_{12}\partial_\eta v_j\right] - \int_{\Gamma_\eta} \mathrm{d}\eta\, \partial_\eta v_i F_{21} v_j\right)\partial_\xi u_j.
\end{aligned}$$

The last equation follows by an integration by parts with respect to the term containing the factor F_{12}. This supplies a convenient and easy to implement form of the FEM-matrix

$$\mathrm{A}_{1,ij} = \int_{\Gamma_\eta} \mathrm{d}\eta\, \left(v_i v_j\partial_\xi F_{11} + v_i\partial_\eta v_j F_{12} - \partial_\eta v_i\, v_j F_{21}\right).$$

Finally, we compute the FEM-matrix which corresponds to $u_j(\xi)$

$$\begin{aligned}
\mathrm{A}_{0,ij}u_j &:= a_0(v_i,\, v_j u_j) \\
&= \left(\int_{\Gamma_\eta} \mathrm{d}\eta\, v_i\,\partial_\xi F_{12}\,\partial_\eta v_j - \int_{\Gamma_\eta} \mathrm{d}\eta\,\partial_\eta v_i\, F_{22}\,\partial_\eta v_j + \int_{\Gamma_\eta} \mathrm{d}\eta\, |J| k^2 v_i v_j\right) u_j
\end{aligned}$$

supplying

$$\mathrm{A}_{0,ij} = \int_{\Gamma_\eta} \mathrm{d}\eta\, \left(v_i\partial_\eta v_j\,\partial_\xi F_{12} - \partial_\eta v_i\partial_\eta v_j\, F_{22} + v_i v_j\,|J|\,k^2\right).$$

Linear C^0- Elements.

We compute the discrete scheme based on simple linear C^0- elements. The generalization to higher order elements follows as usual. For simplicity, we assume a strict ordering of the set of vertices $\mathcal{P}$ obtained if one follows the boundary in positive direction. The linear nodal basis function $v_i(\eta)$ associated with vertex p_i is given by

$$v_i(\eta) = \begin{cases} (\eta - \eta_{i-1})/(\eta_i - \eta_{i-1}) & \text{for } \eta \in [\eta_{i-1},\, \eta_i] \\ (\eta_{i+1} - \eta)/(\eta_{i+1} - \eta_i) & \text{for } \eta \in [\eta_i,\, \eta_{i+1}] \\ 0 & \text{elsewhere} \end{cases}$$

We derive local FEM-matrices from the unit element

$$v^{(0)}(\eta) = \begin{cases} v_1^{(0)} & \text{for } \eta \in [-1,\, 0] \\ v_2^{(0)} & \text{for } \eta \in [0,\, 1] \\ 0 & \text{elsewhere} \end{cases}, v_1^{(0)} = \eta,\ v_2^{(0)} = 1 - \eta.$$

On the segment $\overline{p_i p_{i+1}}$ we introduce the normalization $\widetilde{\eta} = (\eta - \eta_i)/h, h = \eta_{i+1} - \eta_i$. Accordingly, we transform segment-wise $\widetilde{F}(\xi, \widetilde{\eta}) = F(\xi, (\eta - \eta_i)/h)$ and $\widetilde{J}(\xi, \widetilde{\eta}) = J(\xi, (\eta - \eta_i)/h)$. Further, let us renumber the vertices of the segment: $p_1 := p_i$, $p_2 = p_{i+1}$ and drop the vertices-counting indices at the local matrices M_0, M_1, M_2. Thus we can reformulate these matrices, with $(i, j) \in \{1, 2\}$ as

$$
\begin{aligned}
\mathrm{A}_2 &= h\int_0^1 \mathrm{d}\widetilde{\eta}\, v_i^{(0)} v_j^{(0)} \widetilde{F_{11}} \\
\mathrm{A}_1 &= h\int_0^1 \mathrm{d}\widetilde{\eta}\, \left[v_i^{(0)} v_j^{(0)} \partial_\xi \widetilde{F_{11}} + \frac{1}{h} v_i^{(0)} \partial_{\widetilde{\eta}}\, v_j^{(0)} \widetilde{F_{12}} - \frac{1}{h} \partial_{\widetilde{\eta}} v_i\, v_j \widetilde{F_{21}} \right] \\
\mathrm{A}_0 &= h\int_0^1 \mathrm{d}\widetilde{\eta}\, \left[\frac{1}{h} v_i^{(0)} \partial_{\widetilde{\eta}} v_j^{(0)}\, \partial_\xi \widetilde{F_{12}} - \frac{1}{h^2} \partial_{\widetilde{\eta}} v_i^{(0)}\, \partial_{\widetilde{\eta}} v_j^{(0)} \widetilde{F_{22}} + v_i^{(0)} v_j^{(0)} \left|\widetilde{J}\right| k^2 \right]
\end{aligned}
$$

From now, we drop all the waves indicating the normalized argument the dependence of the normalized argument $\widetilde{\eta}$. We introduce the 2×2 block matrix W

$$
\mathrm{W} = \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix}
$$

whose elements are the following 2×2 matrices

$$
W_{11} = v^{(0)} v^{(0)T},\ W_{12} = v^{(0)} \partial_\eta v^{(0)T},\ W_{21} = \partial_\eta v^{(0)}\, v^{(0)T},\ W_{22} = \partial_\eta v^{(0)}\, \partial_\eta v^{(0)T}.
$$

This allows a more compact notation of the elementary matrices:

$$
\begin{aligned}
\mathrm{A}_2 &= h\int_0^1 \mathrm{d}\eta\, W_{11} F_{11} \\
\mathrm{A}_1 &= \int_0^1 \mathrm{d}\eta\, \left[W_{12} F_{12} + h W_{11} \partial_\xi F_{11} - W_{21} F_{21} \right] \\
\mathrm{A}_0 &= \int_0^1 \mathrm{d}\eta\, \left[W_{12} \partial_\xi F_{12} - \frac{1}{h} W_{22} F_{22} + h W_{11} |J|\, k^2 \right]
\end{aligned}
$$

To compute these matrices, we must compute the transformation matrix F and the Jacobian $|J|$. Since neither F nor $|J|$ depend on the choice of the origins of the ξ,η- or the x,y-coordinate system, we set without loss of generality $x_1 = x_2 = 0$, $y_1 = 0$, $y_2 = h$, $\eta_1 = 0$, $\eta_2 = h$, $\beta_1 = -\alpha_1$ and $\beta_2 = \alpha_2$. Using the abbreviations $a_1 = \tan\alpha_1$, $a_2 = \tan\alpha_2$, and $a = \tan\alpha_1 + \tan\alpha_2$, we obtain

$$
J = \begin{pmatrix} \frac{1}{\zeta} & 0 \\ -\frac{1-\eta}{\zeta} a_1 + \frac{\eta}{\zeta} a_2 & 1 + \frac{\xi}{h\zeta} a \end{pmatrix} \quad \text{and} \quad J^{-1} = \begin{pmatrix} \zeta & -\frac{\zeta h(-a_1+\eta a)}{h\zeta + a\xi} \\ 0 & \frac{\zeta h}{h\zeta + a\xi} \end{pmatrix}.
$$

It follows

$$
|J| = \frac{h\zeta + \xi a}{h\zeta^2} \quad \text{and} \quad F = \begin{pmatrix} \zeta + \frac{\xi}{h} a & a_1 - \eta a \\ a_1 - \eta a & h\frac{(a_1-\eta a)^2+1}{h\zeta + \xi a} \end{pmatrix}.
$$

Based on these results we can compute the elementary FEM-matrices. Now we label the matrices and the quantities h, a_1, a_2, a, and ζ with a subscript j and the matrices with a superscript (j) to indicate that these quantities belong to $Q_j^{(\xi,\eta)}$ and may differ between different local domains.

$$
\mathrm{A}_2^{(j)} = (h_j\zeta_j + \xi a_j) \begin{pmatrix} \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} \end{pmatrix}
$$

$$
\mathrm{A}_1^{(j)} = \frac{1}{3} \begin{pmatrix} a & -a_2 + 2a_1 \\ 2a_2 - a_1 & a \end{pmatrix}^{(j)}
$$

$$\begin{aligned}\mathrm{A}_0^{(j)} &= \frac{1}{3}\frac{1}{h_j\zeta_j+\xi a_j}\begin{pmatrix} -a_1^2+a_1a_2-a_2^2-3 & a_1^2-a_1a_2+a_2^2+3 \\ a_1^2-a_1a_2+a_2^2+3 & -a_1^2+a_1a_2-a_2^2-3 \end{pmatrix}^{(j)} \\ &\quad + (h_j\zeta_j+\xi a_j)\frac{k_j^2}{\zeta_j^2}\begin{pmatrix} \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} \end{pmatrix}.\end{aligned}$$

The matrices depend on the quantity $(h_j\zeta_j+\xi a_j)$, which is the distance dependent width $h(\xi)$ of each semi-infinite element. In view of the Laplace transform, it is convenient to regroup the matrices $\mathrm{A}_{0,1,2}^{(j)}$ according to this factor. The local matrices $\mathrm{A}_2^{(j)}$ are associated to $a_2\left(v,\partial_\xi^2 u\right)$, the matrices $\mathrm{A}_1^{(j)}$ to $a_1\left(v,\partial_\xi u\right)$ and $\mathrm{A}_0^{(j)}$ to $a\left(v,u\right)$, according to the exterior variational problem (5.3.2). Thus we have, in a formal notation, the local approximation

$$a_0\left(v,u\right)+a_1\left(v,\partial_\xi u\right)+a_2\left(v,\partial_\xi^2 u\right) \xrightarrow{\text{local approx}} \mathrm{A}_0^{(j)}\mathbf{u}+\mathrm{A}_1^{(j)}\partial_\xi\mathbf{u}+\mathrm{A}_2^{(j)}\partial_\xi^2\mathbf{u}. \tag{5.3.9}$$

We aim to replace the expression on the right-hand side by sum over matrices which indicate besides the order of the associated derivative the power of $(h_j\zeta_j+\xi a_j)$. To this end we introduce the matrices $\mathrm{M}_{i,j}^{(k)}\in\mathbb{R}^{2\times 2}$

$$\sum_{j=0}^{2}\left(\sum_{i=-1}^{1}\left(h_k\zeta_k+\xi a_k\right)^i\mathrm{M}_{i,j}^{(k)}\right)\partial_\xi^j\mathbf{u} := \mathrm{A}_0^{(k)}\mathbf{u}+\mathrm{A}_1^{(k)}\partial_\xi\mathbf{u}+\mathrm{A}_2^{(k)}\partial_\xi^2\mathbf{u} \tag{5.3.10}$$

with

$$\begin{aligned}\mathrm{M}_{-1,0}^{(k)} &= \frac{1}{3}\begin{pmatrix} -a_1^2+a_1a_2-a_2^2-3 & a_1^2-a_1a_2+a_2^2+3 \\ a_1^2-a_1a_2+a_2^2+3 & -a_1^2+a_1a_2-a_2^2-3 \end{pmatrix}^{(k)} \\ \mathrm{M}_{1,0}^{(k)} &= \frac{k_k^2}{\zeta_k^2}\begin{pmatrix} \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} \end{pmatrix} \\ \mathrm{M}_{0,1}^{(k)} &= \frac{1}{3}\begin{pmatrix} a & -a_2+2a_1 \\ 2a_2-a_1 & a \end{pmatrix}^{(k)} \\ \mathrm{M}_{1,2}^{(k)} &= \begin{pmatrix} \frac{1}{3} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{3} \end{pmatrix},\end{aligned}$$

and all the other matrices $\mathrm{M}_{i,j}^{(k)}=0$. In this form it becomes transparent that the chosen method-of-lines type discretization yields a simple algebraic structure.

Coupling between interior and exterior domain.

Based on these matrices we are able to assemble the whole discrete system in ξ,η-coordinates. Exterior and interior domain interact via the boundary integral, which appears on the right-hand side of the variational formulation of the interior problem (in x,y-coordinates) (5.3.1)

$$\int_{\partial\Omega_{\mathrm{int}}} \mathrm{d}s\overline{\nu}(s)\mathbf{n}(s)\nabla_{xy}u(s).$$

We express this in ξ,η-coordinates on the interval $\overline{p_ip_j}$. As before, we set without loss of generality $x_1=x_2=0$, $y_1=0$, $y_2=h$, $\eta_1=0$, $\eta_2=h$ consequently $\mathbf{n}=(1,0)^T$. We find with $i,j\in\{1,2\}$

$$
\begin{aligned}
(5.3.11) &\int_{p_1}^{p_2} \mathrm{d}s \overline{\nabla v_i}(s)\mathbf{n}(s)\nabla_{xy}u(s) \\
&= h\int_0^1 \mathrm{d}\eta\, v_i(\eta)(1,0)J^{-T}\nabla_{\xi\eta}\left(v_j(\eta)u_j(\xi)\right) \\
&= h\int_0^1 \mathrm{d}\eta\, v_i(\eta)\begin{pmatrix} \left(J^{-1}\right)_{11} & \left(J^{-1}\right)_{12}\end{pmatrix}\begin{pmatrix} v_j\partial_\xi u_j \\ \partial_\eta v_j\, u_j\end{pmatrix} \\
&= \left(h\int_0^1 \mathrm{d}\eta\, v_i(\eta)\left(J^{-1}\right)_{11} v_j\right)\partial_\xi u_j + \left(h\int_0^1 \mathrm{d}\eta\, v_k(\eta)\left(J^{-1}\right)_{12}\partial_\eta v_j\right)u_j.
\end{aligned}
$$

The last equation is of central importance for the whole concept. It supplies the link between the discretized interior and exterior problems. It becomes apparent that we do not need to introduce the normal derivative. Instead, we trace back the problem in a formulation in ξ,η-coordinates.

We want to express the terms in parenthesis by local matrices. For continuous, piece-wise linear elements we obtain the elementary matrices, using the notation of the previous section,

$$
\mathrm{B}_1 \;:=\; \left(h\int_0^1 \mathrm{d}\eta\, v_i(\eta)\left(J^{-1}\right)_{11} v_j\right)
$$

which results in the local matrix for the kth segment associated to the (first) derivative in ξ-direction

$$
\mathrm{B}_1^{(k)} = \frac{1}{3}h_k\zeta_k\begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1\end{pmatrix}.
$$

In the same way we obtain

$$
\mathrm{B}_0 \;:=\; \left(h\int_0^1 \mathrm{d}\eta\, v_i(\eta)\left(J^{-1}\right)_{12}\partial_\eta v_j\right)
$$

with the result

$$
\mathrm{B}_0^{(k)} = \frac{1}{6}\begin{pmatrix} a-3a_1 & -a+3a_1 \\ 2a-3a_1 & -2a+3a_1\end{pmatrix}^{(k)}.
$$

Thus, we can extend the discretization scheme to the boundary terms, using the formal notation of (5.3.9),

$$
(5.3.12)\qquad \int_{\partial\Omega_{\mathrm{int}}} \mathrm{d}s\overline{\mathrm{v}}(s)\mathbf{n}(s)\nabla_{xy}u(s) \xrightarrow{\text{local approx}} \left(\mathrm{B}_0^{(k)}\mathbf{u} + \mathrm{B}_1^{(k)}\partial_\xi\mathbf{u}\right)_{\xi=0}.
$$

Algorithm.

Now we are able to assemble the whole discrete problem. Since most of all steps are completely parallel to the discussed 1D examples, we give only a brief overview and leave out the many obvious details.

Step 1: Discretization of the interior problem in standard manner, according to the left-hand side of (5.3.1).

Step 2: Discretization of the boundary terms of the interior problem, according to the right-hand side of (5.3.1), using the scheme (5.3.12). Dirichlet and Neumann data (with respect to ξ) appear as unknowns, cf. Section 5.2 and Section 5.2.2.

Step 3: Assembling of the Laplace transformed exterior problem. To this end, the discretized exterior problem has to be assembled according to (5.3.9), followed by a Laplace transformation of the system. In the spatial domain, exactly the terms $(h_k\zeta_k + \xi a_k)^i \, \partial_\xi^j u$, $i \in \{-1, 0, 1\}$, $j \in \{0, 1, 2\}$ appear, which already have appeared in the 1D situation.

Step 4: Formulate the linking condition exactly as in Section 5.2 as axis-integral condition, or as in Section 5.2.2 as cut-function formulation. Each segment supplies one pair of linking conditions.

Step 5: Discretize in the Laplace domain, either along the real axis (compare Section 5.2) or along the cuts (compare Section 5.2.2).

Step 6: Solve simultaneously the interior (spatial) and the exterior (spectral) problem.

Asymptotic Approximation.

Following the example of the asymptotic (high frequency approximation) of the DtN-number of Bessel's equation Section 2.4, pp. 28, we can apply the same factorization technique to the system of ODE's

$$L\mathrm{u}(\xi) := \left[\left(\sum_{j=0}^{2} \xi^j \mathrm{M}_{j2} \right) \partial_\xi^2 + \left(\sum_{j=0}^{2} \xi^j \mathrm{M}_{j1} \right) \partial_\xi^1 + \left(\sum_{j=0}^{2} \xi^j \mathrm{M}_{j0} \right) \right] \mathrm{u}(\xi) = 0,$$

which results from (5.3.10) by a multiplication with factors of the form $h\zeta + \xi a$ and an assembling of all local matrices. Now we can repeat the procedure of Section 2.4 to derive a family of boundary conditions at the position $\xi = 0$. The symbol of L is given by

$$\sigma(L) = \left(\sum_{j=0}^{2} \xi^j \mathrm{M}_{j2} \right) (-\rho^2) + \left(\sum_{j=0}^{2} \xi^j \mathrm{M}_{j1} \right) (i\rho) + \left(\sum_{j=0}^{2} \xi_{j0}^j \right).$$

Here we used ξ as distance variable as before, and ρ as corresponding dual variable. Note that there is no dual pair ϕ, ν as in the continuous case. Introducing the matrices

$$\mathrm{K}_2 = \left(\sum_{j=0}^{2} \xi^j \mathrm{M}_{j2} \right), \; \mathrm{K}_1 = \mathrm{K}_2^{-1} \left(\sum_{j=0}^{2} \xi^j \mathrm{M}_{j1} \right), \; \mathrm{K}_0 = \mathrm{K}_2^{-1} \left(\sum_{j=0}^{2} \xi^j \mathrm{M}_{j0} \right)$$

the symbol reads

$$\sigma(L) = \mathrm{K}_2 \left(-I\rho^2 + \mathrm{K}_1 i\rho + \mathrm{K}_0 \right).$$

Obviously, this is the discrete counterpart of the symbol of the continuous Helmholtz equation, cf. 2.4.13, p. 31. Accordingly, the factorization

$$\left(-\mathrm{I}\rho^2 + \mathrm{K}_0 \right) = \left(\mathrm{I}i\rho + i\mathrm{K}_0^{1/2} \right) \left(\mathrm{I}i\rho - i\mathrm{K}_0^{1/2} \right)$$

defines the highest order pseudodifferential operator $\Lambda_\pm^1 = \mp i\mathrm{K}_0^{1/2}$, and the same consideration as in Section 2.4 yields the two equations for the symbols $\Lambda_-^0(x,\rho)$ and $\Lambda_-^{-1}(x,\rho)$

$$\begin{array}{rcccl} 0 & = & \mathrm{K}_1 i\rho & + & \partial_x \Lambda_-^1 - \Lambda_-^0 \left(\Lambda_+^1 - \Lambda_-^1 \right) \\ 0 & = & & + & \partial_\xi \Lambda_-^0 - \Lambda_-^{-1} \left(\Lambda_+^1 - \Lambda_-^1 \right) + \Lambda_-^0 \Lambda_-^0. \end{array}$$

We wrote $\Lambda_\pm^j$ for the symbols of the pseudodifferential operators to indicate that these are parameter dependent matrices in our discrete case , where the parameters

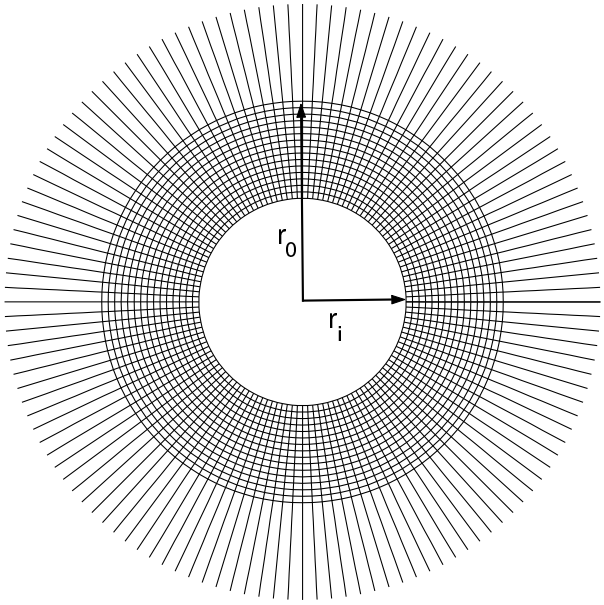


FIGURE 5.4.1. Regular mesh used for the radially symmetric scattering problem.

are ξ and ρ. If $K_0^{1/2}$ can be computed in a numerically cheap way, this gives a practicable way to compute asymptotic boundary conditions directly from the the given semi-discrete problem. Without following this way further here we remark that a cheap computation of a DtN-operator could supply an useful preconditioner for the algebraic system.

5.4. Numerical Examples

The following 2D examples are the very first numerical results and are presented here to give an impression of the technique. Future work is needed to both analyze and improve the numerical sub-problems.

The first numerical experiments are meant to repeat the numerical studies of Section 5.2.2, pp. 115, in a 2D setting. We compute the solution of the Helmholtz equation with $k = 1$ exterior to a disk with radius $r_\mathrm{i} = 2\pi$. On the boundary of the disk Dirichlet data $u|_\Gamma(\phi) = \cos(\nu\phi)$, $0 \le \phi \le 2\pi$ are prescribed. We carried out the experiments with $\nu = 0$ and $\nu = 10$ as characteristic angular frequencies. The interior computational domain is a circular domain between the radii r_i and r_0, where $r_0 > r_\mathrm{i}$ is the radius of the artificial circular boundary. In the following experiments we set $r_0 = 4\pi$, so that the radius of the artificial domain covers two wavelengths. The corresponding regular mesh is shown in Fig. 5.4.1.

Fig. 5.4.2 displays the convergence of the interior solution, if the initial mesh is subsequently uniformly refined, starting with a number of unknowns on the artificial boundary $N_\phi = 32$ and ending up with $N_\phi = 256$. During the refinement of the interior domain we kept fix the discretization of the spectral domain. Here we used 14 intervals and the implicit Runge-Kutta collocation scheme of consistence order 4, p. 118, which results in 28 additional unknowns per node on the artificial boundary.

A natural measure for the quality of the discretization in the spectral domain is the far-field error shown in Fig. 5.4.3. Roughly, the far-field errors behave similarly like the interior errors. The errors corresponding to the higher angular frequency

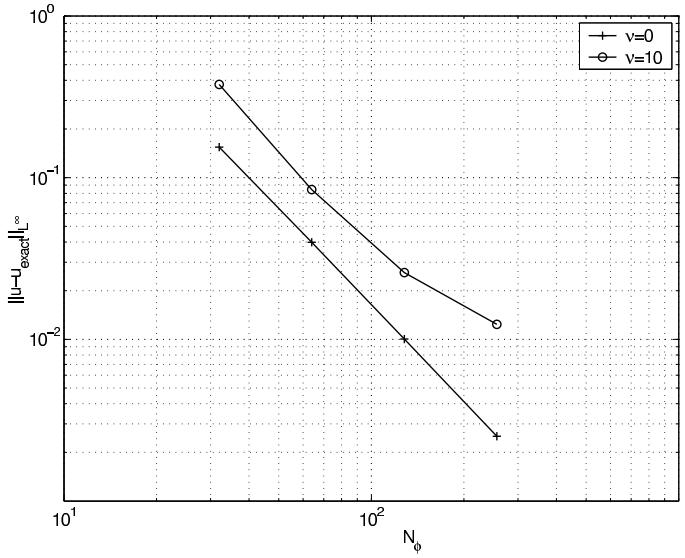


FIGURE 5.4.2. Convergence of the interior solution computed for a sequence of uniformly refined interior meshes. The discretization in the spectral domain is kept fix. N_ϕ denotes the number of nodes on the artificial boundary.

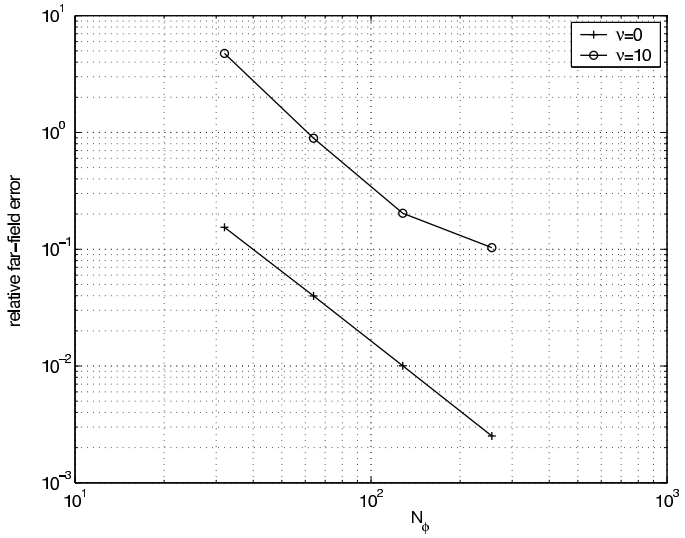


FIGURE 5.4.3. Far-field error corresponding to the radially symmetric scattering problem

$\nu = 10$ are larger than the errors corresponding to $\nu = 0$. This similar behavior shows that in the given situation the far-field error is dominated by the interior discretization error. The cut functions ψ corresponding to the frequencies $\nu = 0$ and $\nu = 10$ for the analyzed radially symmetric experiment are displayed in Fig. 5.4.4. The smooth behavior of these curves indicate that they can be efficiently treated by a number of numerical approximation methods so that future implementations of the cut function approach will result in fast methods.

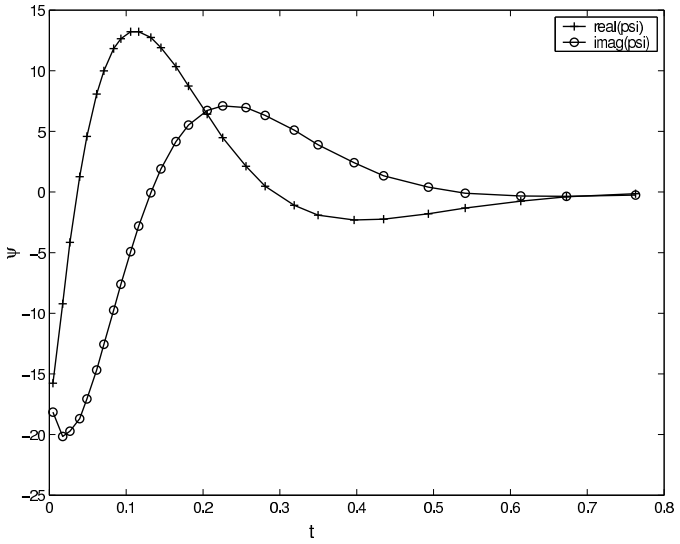


FIGURE 5.4.4. Cut function corresponding to $\nu = 10$. The +-signs indicate the collocation points.

Next we want to give briefly some other numerical experiments. Fig. 5.4.5 shows the triangulation of an elliptic computational domain and the matching ray-like discretization of the exterior domain. A plane wave propagates from left to right and causes the computed reflection from an interior Dirichlet boundary (circle).

Fig. 5.4.6 displays the the first convergence result obtained by our method. This result has a very preliminary character, since we have not implemented the very successful spectral approaches from Sections 5.2 and 5.2.2. Instead, we discretized the whole system as described above, but then we approximated the space of the Dirichlet functions on the boundary by a small number of Fourier modes. This approximation is the same as performed in the classical derivation of DtN-operators based on factorizations, see Section 2.1. In future implementations, this approximation will be replaced by the Laplace domain methods. Fig. 5.4.6 shows the convergence of the interior problem with respect to the dimension of the subspace, that is the number of Fourier modes used for the approximation of u on $\partial\Omega$, for three subsequently uniform refined meshes. The interesting result is that if the approximation quality of the subspace is sufficient, then the remaining error is the discretization error of the interior problem.

Application of radial rays. Fig. 5.4.7 shows an application to the same problem with an quadratic computational domain. The discretization of the exterior domain has been performed with *radial* rays. This allows in a natural way to consider corners, without to introduce any special technique.

Application of normal rays.

The example of Fig. 5.4.8 shows the reflection from a semi-transparent mirror. Here the exterior domain contains a waveguide, which extends from the left boundary to infinity. This waveguide inhomogeneity is easily taken into account by a *normal* ray discretization of the exterior domain.

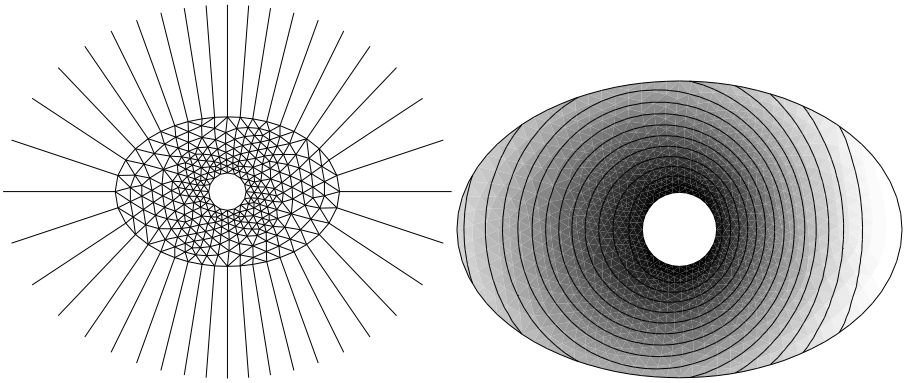


FIGURE 5.4.5. Inner and outer discretization of an elliptic computational domain and the corresponding numerical solution of the scattering problem.

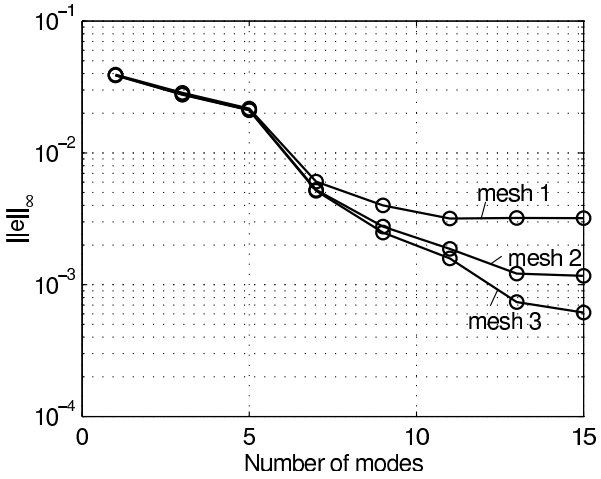


FIGURE 5.4.6. Convergence of the error $e := u_{\text{exact}} - u_{\text{discrete}}$ in maximum norm vs. the number of modes for an initial (mesh 1) and two subsequently refined triangulations (meshes 2 and 3). The error decreases until the discretization error is reached.

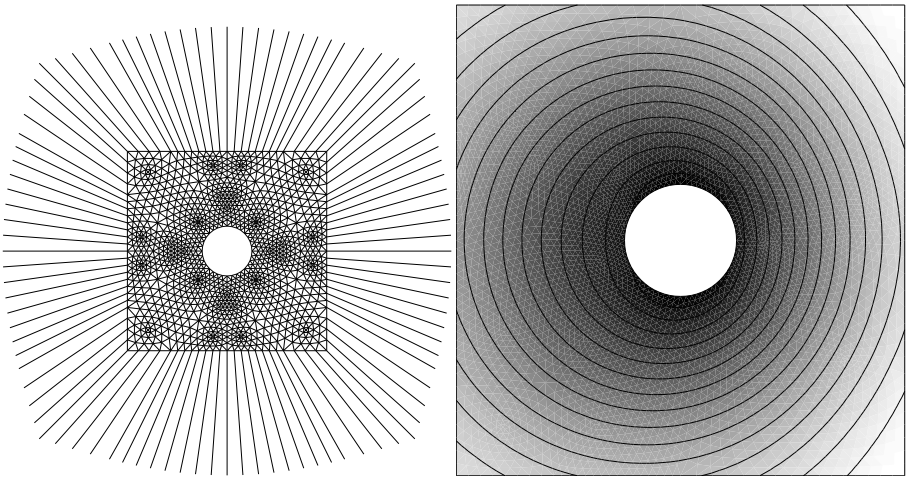


FIGURE 5.4.7. Inner and outer discretization of a quadratic computational domain and the corresponding numerical solution of the scattering problem.

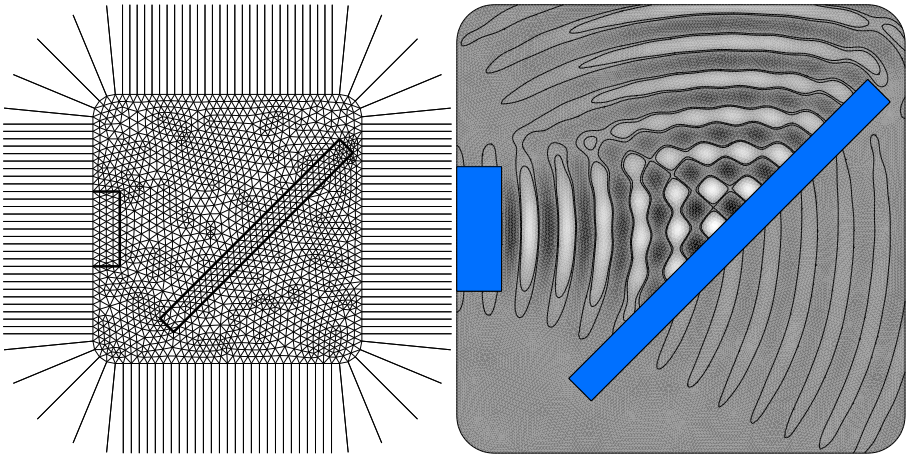


FIGURE 5.4.8. Inner and outer discretization of the reflection problem and the corresponding numerical solution. This is an example of an inhomogeneous scattering problem.

CHAPTER 6

Numerical Treatment of Schrödinger-Type Scattering Problems

This chapter applies the pole condition approach to time-dependent Schrödinger and to wide-angle one way equations. We want to discuss the problem: Given the (possibly variable coefficient) evolution problem and a (possibly non-uniform) time-discretization, how to derive corresponding transparent boundary conditions? We do not aim, in first place, to derive fast algorithms in the sense of e. g. fast convolution algorithms, but on the relation between the time-evolution and the pole condition.

Therefore we study evolution problems with only one spatial dimension. This, however, supplies the framework for the application of the methods developed in the previous sections, which, in turn, will lead to solvers for variable coefficient problems in higher dimensions. We give the principal structure of the corresponding algorithms; the analysis and realization of corresponding algorithms, however, is part of future research.

Our approach here is the same as in Section 5.1, namely a direct computation of the poles by a factorization technique and their elimination by a proper choice of the boundary conditions. In one space dimension this concept leads directly to the desired boundary conditions even in non-trivial cases like the one-way wide-angle equations. Based on these boundary conditions, we can prove the unconditional stability of the numerical propagation algorithms and of some discrete conservation properties.

6.1. Time-Dependent Schrödinger Equation

First we focus on the construction of transparent boundary conditions for discretized Schrödinger equations. The presentation follows the one given by the author in **[80]**, and the joint work with Friese and Yevick published in **[83]**.

Let the continuous evolution equation be given:

$$\begin{aligned} \partial_t u &= -\frac{i}{c}\left(\partial_x^2 u + V(x,t)u\right), \qquad x \in \mathbb{R}, \qquad t > 0 \\ u(x,0) &= u_0, \qquad \text{and} \qquad \lim_{x\to\pm\infty} u(x) = 0, \end{aligned} \tag{6.1.1}$$

cf. Section 1.1, p. 12. Here c is a negative real constant and $V(x,t)$ denotes a real and bounded potential. The operator $A = 1/c(\partial_x^2 + V(x,t))$ with $A : D(A) \subset L^2(\mathbb{R}) \to L^2(\mathbb{R})$ is symmetric. Hence, we have for all $u_0 \in D(A)$ the conservation property $\partial_t (u,u)_{L^2} = 0$. We consider the evolution equation (6.1.1) in the infinite domain $\Omega = \{x, t \in \mathbb{R} : t > 0\}$. Further, we assume that the initial function $u_0(x)$ has support only in a finite interval. We decompose the infinite domain Ω into an interior domain Ω_{int} containing the object of interest and its complement with respect to the whole domain Ω. In our 1D-case we define $\Omega_{\text{int}} = \{x, t \in \mathbb{R} : x_- < x < x_+,\ t > 0\}$, and the exterior domain consists of the two domains $\Omega_- = \{x, t \in \mathbb{R} : x < x_-,\ t > 0\}$ and $\Omega_+ = \{x, t \in \mathbb{R} : x > x_+,\ t > 0\}$, for the left and right semi-infinite sub-domains, respectively. Further, we assume for simplicity

that the potential function V depends in the exterior domains $\Omega_\pm$ only on t, i. e. , $V = V(t)$. A dependence of the potential on the transversal direction x can be taken into account, as well as higher space dimensions, by the same techniques as described in Sections 4 and 5. However, we want to simplify the problem to the extend possible such that we can concentrate on the treatment of the time evolution.

Besides the discrete pole condition based approach, which we present here, there is a large number of other methods tailored to evolution problems. In the following, we extend the discussion of some of the most important methods given in Section 1 to time-dependent methods.

Artificially absorbing layers and PML-method. From the physical point of view, it is a natural to modify the potential function in the exterior domain such that an artificial absorption is generated. The parameters describing the absorber-function have to be adjusted such that backward diffraction from the absorber is small over a prescribed spectral range (e. g. Kosloff and Kosloff **[65]** or Yevick **[92]**. The main advantage of such an approach is its simplicity for two and three-dimensional problems.

The PML-technique of Bérenger **[11]** has been constructed originally to solve time-dependent Maxwell's equations, hence this approach is a natural canditate for Schrödinger-type problems. The method turned out to be very efficient for many problems. Here we have the same simple formulation for 2D and 3D problems but also the same problem in finding the best choice of parameters for given task. A numerical study of the PML-technique with respect to Schrödinger-type equations in the context of optical beam propagation may be found in **[94]**.

Nonlocal boundary conditions. As for Helmholtz-type scattering problems, a solution representation based on Green' functions provides the basis for the construction of transparent boundary conditions. The main steps to derive such boundary conditions are the following. At first, the partial differential equation is converted to a second-order differential equation in space by Laplace transforming in *time*. The transformed differential equation is then solved allowing only for decaying modes in the exterior domains. The back transform to the time-domain yields convolution formulas which are used to construct transparent boundary conditions. The derivation of such formulas for a variety of constant coefficient problems is well understood, see the review paper of Hagstrom **[50]**. The back transform contains convolution integrals of the form

$$\int_0^T \mathrm{d}\tau\, f(t-\tau)g(\tau), \qquad 0 \le t \le T, \tag{6.1.2}$$

which can be discretized directly, as suggested by Baskakov and Popov **[8]**. However, such approaches may lead to numerical instabilities. An analysis of a discretization scheme based on the implicit mid-point rule for the time-evolution may be found in the thesis of Mayfield **[72]**. She showed that numerical stability, even in the case of uniform discretization both in time (with time-step Δt) and space (with space-step Δx), is given only in disjoined intervals for $\Delta t/\Delta x^2$. Alternatively, the problem of finding boundary conditions for the discretized equation, can be consistently formulated for discrete time. This repeats the continuous procedure in the time-discrete situation, cf. e. g. **[69]**. In this manner, Arnold **[4]** composes a boundary condition which incorporates both a uniform space and a uniform time discretization. Besides the uniformity of discretization, the time-independency of the potential in the exterior domain is an essential prerequisite. The key point in his approach is that the uniform space discretization of the interior domain is continued to the exterior domain as well. In contrast, the approach **[82]** supposes a given, possible

nonuniform, time-discretization and an arbitrary inner space-discretization. The space-variable in the exterior domain is not discretized. Additionally, the outer potential may be a function of time. We will label this approach the semi-discrete method. In **[85]** we have shown that our semi-discrete method covers Arnold's fully discrete method.

The non-local (in time) boundary conditions are often considered to be inefficient, if a long evolution history has to be taken into account. Therefore a number of fast solution techniques has been proposed. One strategy is to approximate the convolution integrals by analytical means such that they can be evaluated fast. In its core, the function f contained in the convolution integral (6.1.2) is approximated by a finite sum of exponential functions, say

$$f(t) = \sum_{j=1}^{m} c_j e^{\lambda_j t}$$

which is always possible if the Laplace transform of $f(t)$ is a rational function. For each exponential function $\exp(\lambda_j t)$ the convolution integral (6.1.2) may be expressed as a function $t \mapsto y$:

$$y(t, \lambda_j) := \int_0^t \mathrm{d}\tau \, e^{\lambda_j (t-\tau)} g(\tau).$$

Differentiation with respect to t shows that y satisfies the initial-value problem

$$y' = \lambda_j y + g, \qquad y(0) = 0.$$

The remarkable fact is, that the computation of the convolution integral, which is a nonlocal operation, can be replaced by a solution of an initial value problem, where the computation of a new time-step requires only local operations. The described situation occurs exactly for the wave equation in 3D, when the solution of the wave equation is represented by separated spherical harmonics. The idea and algorithms go back to Grote and Keller **[43, 42]**. Unfortunately, most of all practical problems do not have this simple structure, as became evident even from our analysis of the Helmholtz equation in Sections 1 and 4, where the characteristic singularities were cut singularities. However, it is a convincing idea, to approximate a given function by a finite number of exponentials such that the above strategy may be applied to more general problems. This concept is discussed in detail in the the review of Hagstrom **[50]**. Moreover, in **[1]** Alpert, Greengard and Hagstrom give explicit rigorous bounds on the degree m required to obtain a uniform approximation of f for the whole time-interval T.

A second strategy is to employ the fast convolution technique due to Hairer, Lubich and Schlichte **[51]** to obtain fast methods. Recently, such an approach has been combined with the initial-value formulation of the convolution problem. In **[70]** Lubich and Schädle have shown that nearly optimal algorithms can be constructed this way. As mentioned above, we will not consider the efficiency aspect of our algorithms here, since there is some reasonable hope that the pole condition approach together with the Laplace domain method will yield efficient algorithms, too.

Local approximations of non-local boundary conditions. In case of higher space-dimensions (≥ 2) one obtains local approximations in space using the theory of pseudo-differential operators, exactly as described in Section 2.4. After transforming back into the time-domain the resulting transparent boundary conditions in general become local in space but nonlocal in time, as discussed above. To avoid the additional numerical costs due to the non-locality in time, the dispersion relation between time- and space-variables may be rational approximated in the

dual domain. This construction scheme has been successfully applied by a number of authors. Following the pioneering work of Engquist and Majda **[28]** on hyperbolic equations, advanced approximation techniques have been proposed for mixed parabolic-hyperbolic systems (Halpern, **[52]**) or parabolic equations (Hagstrom, **[49]**). Again, the review of Hagstrom **[50]** gives an excellent overview.

Longitudinal Discretization.

Let us rewrite (6.1.1) as

$$\begin{aligned} \partial_t u &= f(u,t), \qquad (x,t) \in \Omega \\ \lim_{x \to \pm\infty} u(x,t) &= 0. \end{aligned}$$

To solve this equation numerically, we apply the implicit one-step discretization

$$\begin{aligned} u_{i+1} - u_i &= \tau f\left(\theta u_{i+1} + (1-\theta)u_i, t_i + \theta\tau\right) \\ \tau &= t_{i+1} - t_i, \qquad i = 0, 1, \ldots \\ 0 &< \theta \leq 1. \end{aligned}$$

Using the definition of $f(u,t)$ from (6.1.1), we find

$$u_{i+1} - u_i = -i\frac{\tau}{c}\left((\partial_x^2 + V)(\theta u_{i+1} + (1-\theta)u_i)\right). \tag{6.1.3}$$

By rearranging the terms we obtain a sequence of inhomogeneous Helmholtz-type equations

$$\begin{aligned} \partial_x^2 u_{i+1} - \lambda^2 u_{i+1} &= -\Theta\partial_x^2 u_i + \kappa^2 u_i \\ \Theta &= \frac{1-\theta}{\theta} \\ \lambda^2(x, t_i + \theta\tau) &= \frac{ic}{\tau\theta} - V(x, t_i + \theta\tau) \\ \kappa^2(x, t_i + \theta\tau) &= -\frac{ic}{\tau\theta} - \Theta V(x, t_i + \theta\tau). \end{aligned} \tag{6.1.4}$$

We now seek solutions u_i, $i \geq 1$ of (6.1.4) that vanish at infinity. We focus here on obtaining an exterior solution which enables the boundary conditions to be constructed, independently of the numerical method employed to solve the interior problem. For this purpose, we fix the right boundary at $x_+ = 0$, $t > 0$ and search for solutions $u_i(x)$, $i \geq 1$, $x \geq 0$ in the right exterior domain. These exterior solutions have to obey the boundary condition at infinity

$$\lim_{x \to \infty} u_i(x) = 0, \qquad i \geq 1. \tag{6.1.5}$$

Continuous treatment of the space coordinate. In the semi-discretized equation (6.1.3) the space-variable x appears as continuous variable. Therefore (6.1.3) forms a sequence of ordinary differential equations. In order to apply the pole condition, we perform a Laplace transformation with respect to the spatial coordinate x. In contrast to our previous notation, we denote the dual variable to x by p. Later we will introduce the quantity s as shift operator. With

$$\widehat{u}_i(p) = Lu_i(x) = \int_0^\infty \mathrm{e}^{-px} u_i(x)\, dx\,, \quad p \in \mathrm{C}, \quad \mathrm{Re}(p) > 0\,,$$

we find the equivalent transformed system to

$$\begin{aligned} &\left(p^2 - \lambda_{i+1}^2\right)\widehat{u}_{i+1}(p) + \left(\Theta p^2 - \kappa_{i+1}^2\right)\widehat{u}_i(p) \\ &\quad = p u_{i+1}(0) + \partial_x u_{i+1}|_{x=0} + \Theta\left(p u_i(0) + \partial_x u_i|_{x=0}\right). \end{aligned} \tag{6.1.6}$$

Discrete treatment of the space coordinate. The alternative idea is to consider a uniformly discretized space. That is, we associate the solution points at the ith propagation step with physical locations according to the formula

$$u_i^{(j)} = u_i(j \cdot h), \quad j \geq -1, \quad i \geq 0\,.$$

Here $u_i^{(-1)}$ is the rightmost inner value in Ω_i while u_i^0 is located on the boundary between the internal and the right external region. The equation corresponding to 6.1.6 is obtained by introducing the sequences

$$\begin{aligned} U_i &= \{u_i^{(0)}, u_i^{(1)}, u_i^{(2)}, \ldots\} \\ U_i^+ &= \{u_i^{(1)}, u_i^{(2)}, u_i^{(3)}, \ldots\} \\ U_i^- &= \{u_i^{(-1)}, u_i^{(0)}, u_i^{(1)}, \ldots\} \end{aligned} \tag{6.1.7}$$

with Z-transforms

$$\begin{aligned} ZU_i &= u_i^{(0)} + \frac{1}{z}u_i^{(1)} + \frac{1}{z^2}u_i^{(2)} + \ldots \\ ZU_i^+ &= u_i^{(1)} + \frac{1}{z}u_i^{(2)} + \frac{1}{z^2}u_i^{(3)} + \ldots \\ ZU_i^- &= u_i^{(-1)} + \frac{1}{z}u_i^{(0)} + \frac{1}{z^2}u_i^{(1)} + \ldots\,. \end{aligned}$$

By Z-transforming the finite-difference form of (6.1.4) and taking into account the relations between ZU_i, ZU_i^+ , and ZU_i^-, we obtain the discrete counterpart to (6.1.6)

$$\begin{aligned} &\left(z - (2 + h^2\lambda^2) + \frac{1}{z}\right) ZU_{i+1}(z) + \Theta\left(z - (2 + h^2\kappa^2/\Theta) + \frac{1}{z}\right) ZU_i(z) \\ &\quad = u_{i+1}^{(-1)} - zu_{i+1}^{(0)} + \Theta(u_i^{(-1)} - zu_i^{(0)}) \end{aligned} \tag{6.1.8}$$

Comparing (6.1.6) and (6.1.8) we do not find any qualitative difference in the structure of these equations with respect to the time-evolution. The only difference is that the factors before ZU_{i+1} and ZU_i are slightly more involved. Based on this similarity, we restrict all of our following considerations to (6.1.6). The corresponding results for (6.1.8) are obtained in complete analogy.

Construction of Discrete Transparent Boundary Conditions.

The main tool in deriving the transparent boundary conditions is the discrete analog of Mikusiński's operational calculus **[74]** . The difference is that Mikusiński considered an algebra based on continuous, complex valued functions over a semi-infinite time-interval, whereas our basic elements are semi-infinite sequences of complex numbers.

As we have seen, sequences like (6.1.7) play a central role. Let U and V such sequences. We need the operations addition (denoted by $U + V$), multiplication (denoted by UV, which is in fact a convolution) and the map $U \mapsto V = \mathrm{A}U$, where A is an infinite dimensional matrix. The sequence $s = \{0, 1, 0, \ldots\}$ is called the shift operator. In the Appendix A.2 the properties needed for our purposes are collected. In order to reformulate our problem based on the algebraic calculus, we rewrite the sequence (6.1.6) into a matrix equation with infinite dimensional matrices:

$$
(6.1.9)\qquad
\left[p^2 \begin{pmatrix} 1 & & & & \\ \Theta & \ddots & & & \\ & \ddots & 1 & & \\ & & \Theta & 1 & \\ & & & \ddots & \ddots \end{pmatrix} - \begin{pmatrix} \lambda_1^2 & & & & \\ \kappa_1^2 & \ddots & & & \\ & \ddots & \lambda_i^2 & & \\ & & \kappa_{i+1}^2 & \lambda_{i+1}^2 & \\ & & & \ddots & \ddots \end{pmatrix} \right] \begin{pmatrix} \widehat{u}_1 \\ \vdots \\ \widehat{u}_i \\ \widehat{u}_{i+1} \\ \vdots \end{pmatrix}
$$

$$
= \begin{pmatrix} 1 & & & & \\ \Theta & \ddots & & & \\ & \ddots & 1 & & \\ & & \Theta & 1 & \\ & & & \ddots & \ddots \end{pmatrix} \left[p \begin{pmatrix} u_1(0) \\ \vdots \\ u_i(0) \\ u_{i+1}(0) \\ \vdots \end{pmatrix} + \begin{pmatrix} u_1'(0) \\ \vdots \\ u_i'(0) \\ u_{i+1}'(0) \\ \vdots \end{pmatrix} \right] .
$$

Denoting the matrix containing the Θ's by Q and the matrix containing the λ_j^2's and κ_j^2's by T we write, instead of (6.1.9), equivalently

$$
(6.1.10)\qquad (p^2\mathrm{Q} - \mathrm{T})U = \mathrm{Q}\,(p\,\{u(0)\} + \{u'(0)\})\,.
$$

Uniform discretization. Starting with uniform discretization and time-independent exterior potentials, i. e. , $\lambda_1^2 = \lambda_2^2 = \ldots = \lambda^2$ and $\kappa_1^2 = \kappa_2^2 = \ldots = \kappa^2$, we obtain from (6.1.9), using the shift-operator definition,

$$
(6.1.11)\qquad \left(p^2 - \frac{\lambda^2 + s\kappa^2}{1 + s\Theta} \right) U = p\{u(0)\} + \{u'(0)\}\,.
$$

This way, we have expressed the operators T and Q completely in terms of the shift operator s. Both T and Q are in the space of sequences $\mathcal{C}$, therefore $\mathrm{Q}^{-1}\mathrm{T}$ is in the ring of fractions $\mathcal{C}_q$ (see the Appendix A.2), hence all rational operations with respect to $\mathrm{Q}^{-1}\mathrm{T}$ are well defined. The term $\mathrm{Q}^{-1}\mathrm{T}$ has the algebraic structure of a lower triangular matrix, since T and Q are lower triangular matrices. Because $\mathrm{Im}(\lambda^2) \neq 0$, one proves easily that there is always a uniquely defined lower triangular matrix L such that

$$
\mathrm{L} \cdot \mathrm{L} = \mathrm{L}^2 = \mathrm{Q}^{-1}\mathrm{T} \qquad \text{with} \quad \mathrm{Re}(\mathrm{diag}(\mathrm{L})) > 0\,.
$$

Formally, we write

$$
(6.1.12)\qquad \mathrm{L} = \sqrt{\frac{\lambda^2 + s\kappa^2}{1 + s\Theta}}.
$$

It is clear that for the expression (6.1.12) a Taylor series expansion in s at $s = 0$ exists. To any given order, we can take a sufficient number of terms of the series expansion such that the square of the series expansion equals to $\mathrm{Q}^{-1}\mathrm{T}$ up to the given order. Therefore we have equivalently to (6.1.12)

$$
\mathrm{L} = \lambda + a_1 s + a_2 s^2 + \ldots, \qquad \text{with} \quad \lambda = \sqrt{\lambda^2},\ \mathrm{Re}(\lambda) > 0,
$$

see Example A.2.6 of the Appendix A.2 for a more detailed discussion. It holds $\mathrm{Im}\,\lambda^2 < 0$, by its definition in (6.1.4). Hence λ as defined above has a negative imaginary part. We obtain the exterior solutions $\{U\}$, as function of the boundary data,

$$
U = \frac{p\,\{u(0)\} + \{u'(0)\}}{(p - \mathrm{L})(p + \mathrm{L})}.
$$

This expression has poles at $\pm\mathrm{L}$. The operator $+\mathrm{L}$ contains $+\lambda$, the operator $-\mathrm{L}$ contains $-\lambda$. The special choice

Pole condition for the uniformly discretized Schrödinger equation:

$$\{u'(0)\} = -\mathrm{L}\,\{u(0)\} \tag{6.1.13}$$

supplies the desired transparent boundary conditions. This becomes clear, if we consider the related exterior solutions U

$$U = \frac{\{u(0)\}}{p + \mathrm{L}}.$$

Since L is a lower triangular matrix with real parts of the diagonal entries all greater than zero, it follows that U contains only decaying modes.

REMARK. We have motivated the pole condition by the decay behavior of the exterior solution. Therefore we excluded poles with positive real part. Since $\mathrm{Im}\,\lambda < 0$ for $\mathrm{Re}\,\lambda > 0$, we could require as well that all poles with negative imaginary part must be excluded. This, however, is exactly our pole condition defined in 3.4.8, p. 56.

Non-uniform discretization (algebraic approach). The difficulty, which appears in generalizing the uniform approach is that the operator T cannot be expressed in terms of the shift-operator s, because all λ_i^2's and κ_i^2's may be different from each other. Nevertheless, the whole approach remains the same, only the simplification of the operator representation is lost. Equation (6.1.10) now reads

$$\left(p^2 + \frac{1}{1 + s\Theta}\mathrm{T}\right)\{U\} = p\,\{u(0)\} + \{u'(0)\}\,. \tag{6.1.14}$$

As the argumentation concerning the factorization of operators $\mathrm{Q}^{-1}\mathrm{T}$ is equally valid here, we solve the problem of the transparent boundary conditions in the nonuniform case, if we find a factorization

$$\mathrm{L}\cdot\mathrm{L} = \mathrm{L}^2 = \frac{1}{1 + s\Theta}\mathrm{T} \qquad \text{with} \quad \mathrm{Re}(\mathrm{diag}(\mathrm{L})) > 0.$$

This, however, is a standard task in numerical linear algebra: find the square root of a lower triangular matrix with non-vanishing diagonal elements such that the lower triangular matrix contains only diagonal elements with positive real part. Among these methods are the direct Cholesky-like factorization, which amounts to $\mathcal{O}(n^2)$ operations in the n-th propagation step, Krylov-subspace methods, and suitable basis-transformations, which amount to $\mathcal{O}(n)$ operations. Once the matrix L is obtained, we finally arrive again at the boundary condition (6.1.13).

Laplace domain method.

We refer to the algebraic representation of the time-discrete evolution problem in its general form (6.1.10). Instead to compute the poles explicitly, as we have done so far, we may add to each equation of the system the pole condition as an additional constraint. Let us take, for illustrative purposes, the pole condition in its direct (axis integral) form, as given in Lemma 4.4.3.

Following the representation given in Section 5.2, the full algebraic system consists of three parts:

Part I: Interior problem in discrete form given by (6.1.4).

Part II: Pole condition in discrete form. It consists of a discretization of U along the real axis

$$0 = \int_{\infty}^{-\infty} \mathrm{d}\tau \frac{U(\tau)}{\tau - p_+}$$

for some fixed p_+ with $\mathrm{Im}\,(p_+) > 0$, and $\tau \in \mathbb{R}$.

Part III: Laplace transformed exterior problem in discrete form. It consists of a discretization of (6.1.10) along the real axis

$$(\tau^2 \mathrm{Q} - \mathrm{T}) U(\tau) = \mathrm{Q}\left(\tau \{u(0)\} + \{u'(0)\}\right), \qquad \tau \in \mathbb{R}.$$

We require the simultaneous solution of all three parts. The discretization of U in the spectral domain, and the linking to the interior problem can be realized exactly as discussed in detail in Sections 5.2 and 5.2.2.

Let n_p the number of propagation steps and n_U the number of discretization points of U on the real axis. Then the proposed approach has the obvious properties:

(1) A fixed discretization of U along the real axis with n_U points requires $O(n_p n_U)$ additional operations to compute the boundary condition, which is optimal.
(2) Since we can solve the system equation by equation, the additional memory requirement is $O(n_U)$, i. e., it is independent of the number of steps.

The analysis and implementation of this algorithm is part of the current research.

6.2. Wide Angle One-Way Equations

The following is based on the representation given by the author, Friese and Yevick in **[83]** and extends the foregoing to a more general case. We consider a scalar wave propagation in 2D modeled by the Helmholtz equation

$$\partial_z^2 u + \partial_x^2 u + k^2 u = 0 \tag{6.2.1}$$

in the entire $\mathbb{R}^2$, where the wavenumber k is a position dependent, bounded function $(x, z) \mapsto k$. In many physical situations, we can further distinguish a principal propagation direction, here taken to be the z-direction and a transverse, x direction. For the particular case of a position-independent wavenumber, the operator $\partial_z^2 + \partial_x^2 + k^2$ can be explicitly factorized, leading to the exact one-way Helmholtz equation

$$\partial_z u = ik\sqrt{1 + \frac{1}{k^2}\partial_x^2}\, u\,. \tag{6.2.2}$$

In the above expression, the formal square root operator is a pseudo-differential operator, which can be given a precise meaning by the calculus of pseudo-differential operators, see Section 2.4. In the following, we will replace square root operator $\sqrt{1+X}$, with $X = 1/k^2\partial_x^2$, by an rational expression. Such approximation can be formally written as

$$\partial_z u = ik \frac{\left(1 - b_0'\partial_x^2\right)\left(1 - b_1'\partial_x^2\right)\cdots\left(1 - b_m'\partial_x^2\right)}{\left(1 - b_0\partial_x^2\right)\left(1 - b_1\partial_x^2\right)\cdots\left(1 - b_n\partial_x^2\right)} u \tag{6.2.3}$$

with complex coefficients $b_0', \ldots, b_m'$ and $b_0, \ldots, b_n$. The approximation quality and the well-posedness of this formal equation has been examined in e. g. **[89]** and **[6]**. Additionally, the non-commutativity of the factors in (6.2.3) in case of non-constant coefficients $b_0', \ldots, b_m'$ and $b_0, \ldots, b_n$ on the computational domain is the source of theoretical investigations which are not discussed here. Rather, we adopt the usual technique of *frozen coefficients*, in which the functions are considered to be constant in deriving the rational approximation, cf. e. g, **[6]**.

From a physical perspective, while originally developed in computational underwater acoustics propagation methods based on 6.2.3 have proved critically important in many other contexts, as evident from some representative references, e. g. **[20, 19, 48, 93, 59]**. A high-order Padé approximation to the Helmholtz propagator ensures that widely-divergent beams are described to great accuracy. Thus the Padé approximation to the Helmholtz equation supplies a model which ranges between the Schrödinger equation and the Helmholtz equation. It combines the simplicity of an evolution algorithm with the generality of the Helmholtz equation.

In the following sections, we will accordingly investigate the discrete solution to the following problem:

(1) Let the coefficients $b_0', \dots, b_m'$ and $b_0, \dots, b_n$ be piecewise continuous real functions on the bounded computational domain $\Omega := (x_-, x_+) \times (0, z_{\max})$ and real constants outside the computational domain.
(2) Let u be the solution of (6.2.3) on the unbounded domain $\mathbb{R} \times [0, z_{\max}]$ with initial conditions $u(x, 0) = u_0(x)$ compactly supported on the interval $[x_-, x_+] \subset \mathbb{R}$ together with the asymptotic boundary condition $\lim_{|x| \to \infty} u(x, z) = 0$ holds for all $0 \le z \le z_{\max}$.
(3) Determine the solution $u(x)$ on the bounded computational domain Ω that agrees exactly with the unbounded result, restricted to Ω.

The goal of this section is to construct transparent boundary conditions for arbitrary higher order evolution equations of the type (6.2.3) that insure the realization of the third point above. Similar to the Schrödinger equation, various techniques have been developed for the solution of such boundary problems. Most of all the techniques discussed in Section 6.1 has been applied to the wide-angle one-way equation, too.

Recently, a specialization of the problem defined in this introduction based on a Padé (2,2) approximation to the Helmholtz operator has been successfully analyzed by Arnold and Erhardt **[5]**. In this reference, the authors follow the common way with forth- and back Laplace transforming in *time*. As outlined in Section 6.1, this implies to compute the *inverse* Laplace transform, which restricts the applicability of the method.

We extend the properties of our method described above as follows:

- We allow for an arbitrary high order Padé-approximation of the square root operator on the entire transversal axes.
- We do not need to introduce a Green's function representation for the chosen Padé approximation.
- We do not need to compute an explicite inverse Laplace transform.
- We prove unconditional stability of the algorithm.

As a result, we are able to develop numerical techniques even for high Padé orders. It should be noted that this presentation is restricted to uniform step-sizes in the propagation direction, so that we can apply the shift-operator technique already used in the previous sections. Similar formulas can however be derived for non-equidistant step sizes using the algebraic approach of Section 6.1, but the analysis as well as the resulting formulas would be far more complicated. So we drop this here.

Standard Wide-Angle Approximations.

We first summarize the standard wide-angle approximation to the Helmholtz equation (6.2.1) that will form the basis for our subsequent considerations. The first step in this analysis is to define a new field variable $\widetilde{u}(x, z) := u(x, z) \exp(-ik_0 z)$. The reference-wavenumber k_0 is chosen to effectively minimize the mean phase

TABLE 1. Padé coefficients

	$(2,0)$		$(2,2)$		$(4,2)$			$(4,4)$		
j	0	2	0	2	0	2	4	0	2	4
c'_j	1	1/2	1	3/4	1	1	1/8	1	5/4	5/16
c_j	1		1	1/4	1	1/2		1	3/4	1/16

velocity of the wavevector components of $\widetilde{u}(x,z)$ in the z-direction. The resulting spectrally shifted Helmholtz equation, where we drop the tilde on $u(x,z)$ for notational simplicity, is

$$\left(\partial_z^2 + 2ik_0\partial_z + \partial_x^2 + k^2 - k_0^2\right) u(x,z) = 0\,. \tag{6.2.4}$$

Formally factorizing the above equation in analogy with (6.2.2) yields the following one-way equation after the scaling $z := z \cdot k_0$ and $x := x \cdot k_0$

$$\partial_z u = -i\left(1 - \sqrt{1+X^2}\right) u, \qquad \text{with } X^2 := \frac{k^2 - k_0^2}{k_0^2} + \partial_x^2. \tag{6.2.5}$$

Since $k = \text{const}$ in the exterior domain, rational approximations of the form

$$\sqrt{1+X^2} \simeq \frac{c'_0 + c'_2 X^2 + \ldots + c'_{2m} X^{2m}}{c_0 + c_2 X^2 + \ldots + c_{2n} X^{2n}}\,. \tag{6.2.6}$$

can be applied to the square-root operator. The most accurate of these are generally obtained for $m = n$ or $m = n+1$, see **[89]**. We will derive a general procedure for deriving transparent boundary conditions that can be applied to all such approximations in succeeding sections. However, to simplify our numerical algorithm, we will consider only Padé-type approximations of order $(2m, 2n)$. This yields an interpolation error $O\left(X^{2m+2n+2}\right)$ for $X \to 0$.

For the simple Padé approximation of (6.2.6) the coefficients c'_{2i}, c_{2j}, $i = 0,\ldots,m$ and $j = 0,\ldots,n$ can be obtained through either an explicit factorization (**[7]**) or by the Newman-procedure **[53]**. The latter technique, upon which our computer codes are based, can be extended to wide-angle equations other than those based on Padé-approximations. Some of the resulting coefficients can be found in Table 1.

Longitudinal Discretization.

The implicit midpoint discretization of (6.2.5) results in

$$\frac{u_i(x) - u_{i-1}(x)}{\Delta z} = -i\left(1 - \sqrt{1+X^2}\right)\frac{u_i(x) + u_{i-1(x)}}{2}. \tag{6.2.7}$$

Here $u_i(x)$, $0 < i \leq n$ denotes $u(x, z_0 + i\Delta z)$ with z_0 the initial value of the longitudinal distance and Δz the propagation step length. Note here that the subscripts i denote the longitudinal stepnumber. Superscript i values instead denote the ith integer power of the corresponding quantity while algebraic factors of i refer to the imaginary unit. From (6.2.7), we immediately obtain the discrete evolution equation

$$\left(1 + \frac{i\Delta z}{2}\left(1 - \sqrt{1+X^2}\right)\right) u_i(x) = \left(1 - \frac{i\Delta z}{2}\left(1 - \sqrt{1+X^2}\right)\right) u_{i-1}(x)\,.$$

After replacing the square-root operator by its rational approximant we obtain an equation of the form

$$u_i(x) = \frac{P'\left(X^2\right)}{P\left(X^2\right)} u_{i-1}(x). \tag{6.2.8}$$

Here $P\left(X^2\right)$ and $P'\left(X^2\right)$ are the following polynomials of degree m in the variable X^2

$$\begin{aligned} P'\left(X^2\right) &= \left(1-i\frac{\Delta z}{2}\right)C\left(X^2\right)+i\frac{\Delta z}{2}C'\left(X^2\right) \qquad \text{and} \\ P\left(X^2\right) &= \left(1+i\frac{\Delta z}{2}\right)C\left(X^2\right)-i\frac{\Delta z}{2}C'\left(X^2\right). \end{aligned}$$

The quantities $C'\left(X^2\right)$ and $C\left(X^2\right)$ are themselves polynomials defined by $C'\left(X^2\right) = c'_0+c'_2X^2+\ldots+c'_{2m}X^{2m}$ and $C\left(X^2\right)=c_0+c_2X^2+\ldots+c_{2n}X^{2n}$, see (6.2.6) and Tab. (1). Applying a complex root finder yields

$$P'\left(X^2\right)=c'\prod_{j=1}^{k}\left(1-a'_jX^2\right) \text{ and } P\left(X^2\right)=c\prod_{j=1}^{k}\left(1-a_jX^2\right), \tag{6.2.9}$$

with c and $c'$are constants. Choosing $X^2=0$ it follows $P'(0)/P(0)=c'/c$ and comparing with (6.2.7) shows $c'/c=1$. Finally, inserting $X^2:=\left(k^2-k_0^2\right)/k_0^2+\partial_x^2$ leads to the desired factorization.

Discrete Evolution Equation.

From the rational approximation (6.2.8) and the factorization (6.2.9) we obtain the longitudinally discretized form of the evolution equation

$$u_i(x)=\left(\frac{1-a'_k\partial_x^2}{1-a_k\partial_x^2}\right)\cdots\left(\frac{1-a'_2\partial_x^2}{1-a_2\partial_x^2}\right)\left(\frac{1-a'_1\partial_x^2}{1-a_1\partial_x^2}\right)u_{i-1}(x), \tag{6.2.10}$$

which is the exact counterpart of the continuous evolution equation (6.2.3). In terms of the intermediate functions $g_i^{(1)}(x),\ldots\, g_i^{(k-1)}(x)$ given by

$$\begin{aligned} g_i^{(1)}(x) &= \left(\frac{1-a'_1\partial_x^2}{1-a_1\partial_x^2}\right)u_{i-1}(x) \\ g_i^{(2)}(x) &= \left(\frac{1-a'_2\partial_x^2}{1-a_2\partial_x^2}\right)g_i^{(1)}(x) \\ &\vdots \\ g_i^{(k-1)}(x) &= \left(\frac{1-a'_{k-1}\partial_x^2}{1-a_{k-1}\partial_x^2}\right)g_i^{(k-2)}(x) \\ u_i(x) &= \left(\frac{1-a'_k\partial_x^2}{1-a_k\partial_x^2}\right)g_i^{(k-1)}(x), \end{aligned}$$

the factorized, order $2k$ discrete evolution problem (6.2.10) can be recast into the following system of k second-order differential equations

$$\begin{aligned} \left(1-a_1\partial_x^2\right)g_i^{(1)}(x)-\left(1-a'_1\partial_x^2\right)u_{i-1}(x) &= 0 \\ \left(1-a_2\partial_x^2\right)g_i^{(2)}(x)-\left(1-a'_2\partial_x^2\right)g_i^{(1)}(x) &= 0 \\ &\vdots \\ \left(1-a_{k-1}\partial_x^2\right)g_i^{(k-1)}(x)-\left(1-a'_{k-1}\partial_x^2\right)g_i^{(k-2)}(x) &= 0 \\ \left(1-a_k\partial_x^2\right)u_i(x)-\left(1-a'_k\partial_x^2\right)g_i^{(k-1)}(x) &= 0. \end{aligned} \tag{6.2.11}$$

This sequence of equations may be rewritten conveniently into matrix form

(6.2.12)
$$\left[\begin{pmatrix} 1 & & & & \\ \ddots & \ddots & & & \\ & -1 & 1 & & \\ & & -1 & 1 & \\ & & & \ddots & \ddots \end{pmatrix} + \begin{pmatrix} -a_1\partial_x^2 & & & & \\ \ddots & \ddots & & & \\ & a_k'\partial_x^2 & -a_k\partial_x^2 & & \\ & & a_1'\partial_x^2 & -a_1\partial_x^2 & \\ & & & \ddots & \ddots \end{pmatrix}\right] \mathrm{g}(x) = 0,$$

where the unknown functions are assembled into the vector

$$\mathrm{g}(x) = \left(g_1^{(1)}, \ldots, g_1^{(k-1)}, u_1, \ldots, \underbrace{g_i^{(1)}, \ldots, g_i^{(k-1)}, u_i}_{i\text{th macro step}}, \ldots \right)^T .$$

Here quantities distinguished by a subscript i are evaluated during the ith physical propagation step, while the $k-1$ intermediate partial steps that comprise each physical step and that are associated with different terms in the Padé approximation are distinguished through their superscripts.

Given an initial condition with respect to x, at some fixed point, say x_+, we can solve the system (6.2.12) equation by equation. Thus at step i the function $u_{i-1}(x)$ maps over $k-1$ intermediate functions $g_i^{(j)}(x)$ to the solution $u_i(x)$. Hence we consider a mapping

$$u_{i-1}(x) \longmapsto \underbrace{g_i^{(1)}(x) \mapsto g_i^{(2)}(x) \mapsto \ldots \mapsto g_i^{(k-1)}(x) \longmapsto u_i(x)}_{\mathrm{g}_i(x) := \left(g_i^{(1)}(x), g_i^{(2)}(x), \ldots g_i^{(k-1)}(x), u_i(x)\right)} \tag{6.2.13}$$

$$u_{i-1}(x) \longmapsto \mathrm{g}_i(x)\,, \tag{6.2.14}$$

which depends on the initial conditions (with respect to x). In the lowest order case $k = 1$ the vector $\mathrm{g}_i(x)$ consists of the single function $u_i(x)$.

To write (6.2.11) or (6.2.12), respectively, as a simple matrix equation with finite dimension, we introduce the discrete shift-operator s_k with the property that $s_k u_i(x) = u_{i-1}(x)$. This operator shifts our sequence by one physical step, corresponding to k terms in g. That is, if s is the shift operator which shifts the sequence in g by one place, we have $s_k = s^k$. Since in the following the index shift occurs only in powers of s, we redefine $s := s_k$ to yield the more intuitive formulation denoted by $su_i(x) = u_{i-1}(x)$.

Introducing the notation g' for $\partial_x \mathrm{g}$ and eliminating $u_{i-1}(x)$ in (6.2.12) in terms of $u_i(x)$ then yields

$$\left(\mathrm{E}(s) + \mathrm{A}(s)\partial_x^2\right) \mathrm{g}_i(x) = 0 \text{ for all macro steps } i \geq 1, \tag{6.2.15}$$

which must be solved subject to boundary conditions

$$\mathrm{g}'_{i,+} = \mathrm{B}_+(s)\mathrm{g}'_{i,+} \text{ and } \mathrm{g}'_{i,-} = \mathrm{B}_-(s)\mathrm{g}_{i,-}. \tag{6.2.16}$$

A comparison with the infinite dimensional system (6.2.12) demonstrates that the $k \times k$-matrices $\mathrm{E}(s)$ and $\mathrm{A}(s)$ and the k-element vectors $\mathrm{g}_i(x)$, $\mathrm{g}_{i,\pm}$ and $\mathrm{g}'_{i,\pm}$ are given by

$$\mathrm{E}(s) = \begin{pmatrix} 1 & & & & -s \\ -1 & 1 & & & \\ & & \ddots & & \\ & & -1 & 1 & \\ & & & -1 & 1 \end{pmatrix}, \mathrm{A}(s) = \begin{pmatrix} -a_1 & & & & sa_1' \\ a_2' & -a_2 & & & \\ & & \ddots & & \\ & & a_{k-1}' & -a_{k-1} & \\ & & & a_k' & -a_k \end{pmatrix},$$

$$\mathrm{g}_i(x) = \begin{pmatrix} g_i^{(1)}(x) \\ g_i^{(2)}(x) \\ \vdots \\ g_i^{(k-1)}(x) \\ u_i(x) \end{pmatrix}, \qquad \mathrm{g}_{i,\pm} = \begin{pmatrix} g_i^{(1)} \\ g_i^{(2)} \\ \vdots \\ g_i^{(k-1)} \\ u_i \end{pmatrix}_{x=x_\pm}, \qquad \mathrm{g}'_{i,\pm} = \begin{pmatrix} \dot{g}_i^{(1)} \\ \dot{g}_i^{(2)} \\ \vdots \\ \dot{g}_i^{(k-1)} \\ \dot{u}_i \end{pmatrix}_{x=x_\pm}.$$

We refer to $\mathrm{E}(s)$ and $\mathrm{A}(s)$ as operators, as they depend explicitly on the shift-operator. Further, they may be represented as $\mathrm{D}(s) = \sum_{j=0}^{i} \mathrm{D}_j s^j$, where the $k \times k$ matrices D_j, $j = 0, \ldots, i$ are complex. We will label operators of this general type convolution operators or convolution matrices; the properties of such matrices that are required to derive our subsequent algorithms are collected in the Appendix. The Dirichlet-to-Neumann operators $\mathrm{B}_\pm(s)$ that implement the desired boundary conditions must be constructed such that the asymptotic boundary condition $\lim_{|x|\to\infty} u_i(x) = 0$ is fulfilled for all propagation steps. We will show later by construction that these boundary operators are also of convolution type.

Discrete Transparent Boundary Conditions.

We now present our derivation of transparent boundary conditions for general wide-angle methods. To simplify the discussion, we consider $\mathrm{g}_i(x)$ only in the right exterior domain, $x \geq x_+$, and further shift the position of the right boundary x_+ to the origin according to $x \mapsto x + x_0$ and omit the $\pm$-subscript. As well, we designate the boundary value $\mathrm{g}_{i,+}$ by $\mathrm{g}_{i,0}$. All our results of course apply equally to the left exterior domain.

Consider (6.2.15) as an initial value problem in the right exterior domain with data given on the shifted boundary $x_+ = 0$. Our objective is to construct an operator with the property that the corresponding exterior solution decays asymptotically for any given Dirichlet data $\mathrm{g}_{i,0}$. To do this, we construct the Laplace transform $\widehat{\mathrm{g}}_i$ of the exterior solution vector $\mathrm{g}_i(x)$. In the exterior domain, the Laplace transform of the equation system (6.2.15 is

$$\left(\mathrm{E}(s) + p^2 \mathrm{A}(s)\right) \widehat{\mathrm{g}}_i(p) = \mathrm{A}(s) \left(p\mathrm{g}_{i,0} + \mathrm{g}'_{i,0}\right).$$

The operator $\mathrm{A}(s)$ is invertible, since $a_j \neq 0$, $j = 1, \ldots, k$, cf. Section 6.2 and Lemma A.2.4. Therefore we can write equivalently

$$\left(p^2 \mathrm{I} - \mathrm{C}^2(s)\right) \widehat{\mathrm{g}}_i(p) = p\mathrm{g}_{i,0} + \mathrm{g}'_{i,0},$$

where the $k \times k$-matrix $\mathrm{C}^2(s) := -\mathrm{A}(s)^{-1}\mathrm{E}(s)$ is of convolution type and I is the $k \times k$-identity-matrix. If we assume for the moment that we can obtain the square roots, $\pm\mathrm{C}(s)$, of the matrix $\mathrm{C}^2(s)$ we have $\left(p^2\mathrm{I} - \mathrm{C}^2(s)\right) = \left(p\mathrm{I} + \mathrm{C}(s)\right)\left(p\mathrm{I} - \mathrm{C}(s)\right)$, since $\mathrm{C}(s)$ and I commute. Thus the ansatz $\mathrm{g}'_{i,0} = \mathrm{B}\mathrm{g}_{i,0}$, see (6.2.15), with $\mathrm{B}(s) = -\mathrm{C}(s)$ eliminates all poles which belong to asymptotically increasing solutions, hence we derived the

Pole condition for wide angle one-way equations
$\mathrm{g}'_{i,0} = \mathrm{B}\mathrm{g}_{i,0}, \qquad$ with $\mathrm{B}(s) = -\mathrm{C}(s)$

Further, we obtain the *solution representation*,

(6.2.17)

$$\widehat{\mathrm{g}}_i(p) = \left(p\mathrm{I} + \mathrm{C}(s)\right)^{-1} \mathrm{g}_{i,0} \quad \text{subject to boundary conditions } \mathrm{g}'_{i,0} = \mathrm{B}(s)\mathrm{g}_{i,0}.$$

To derive the nonlocal boundary conditions, which is equivalent to finding the boundary operator $\mathrm{B}(s)$, we apply the same procedure as in **[82]** or **[85]**. That is, we construct the matrix $\mathrm{C}(s)$ such that all poles p_j, $j = 1, \ldots, k$ of $(p\mathrm{I} + \mathrm{C}(s))^{-1}$ are located in the right half of the complex plane, i.e. $\mathrm{Re}\, p_j > 0$, $j = 1, \ldots, k$. This ensures that the exterior solution decays appropriately as $x \to \pm\infty$. The operator $(p\mathrm{I} + \mathrm{C}(s))^{-1}$ possesses a pole in p if the matrix $(p\mathrm{I} + \mathrm{C}(s)|_{s=0})^{-1}$ possesses a pole in p, cf. the properties of operators of convolution type given in the Appendix.

The square root of $\mathrm{C}^2(s) := -\mathrm{A}^{-1}(s)\mathrm{E}(s)$ is obtained by decomposing the matrices $\mathrm{A}(s)$ and $\mathrm{E}(s)$ into its components with respect to s.. The first of these is independent of the shift operator s , namely $\mathrm{A}_0 := \mathrm{A}|_{s=0}$ and $\mathrm{E}_0 := \mathrm{E}|_{s=0}$ while the second is s-dependent. We thus have $\mathrm{A} = \mathrm{A}_0 + s\mathrm{A}_1$ and $\mathrm{E} = \mathrm{E}_0 + s\mathrm{E}_1$ with

$$\mathrm{A}_1 := \begin{pmatrix} 0 & \cdots & 0 & a_1' \\ & & & 0 \\ & \mathbf{0} & & \vdots \\ & & & 0 \end{pmatrix} \quad \mathrm{E}_1 := \begin{pmatrix} 0 & \cdots & 0 & -1 \\ & & & 0 \\ & \mathbf{0} & & \vdots \\ & & & 0 \end{pmatrix}.$$

For convenience, we define new matrices

$$\tilde{\mathrm{A}}_1 := \mathrm{A}_0^{-1}\mathrm{A}_1 \quad \tilde{\mathrm{E}}_0 := \mathrm{A}_0^{-1}\mathrm{E}_0 \quad \tilde{\mathrm{E}}_1 := \mathrm{A}_0^{-1}\mathrm{E}_1\,.$$

From the previous definitions together with the ansatz $\mathrm{C}(s) = \sum_{j=0}^{i-1} \mathrm{C}_j s^j$, where $\mathrm{C}_j \in \mathbb{C}^{k\times k}$, $0 \le j \le i$, we observe that we must find matrices C_j such that

(6.2.18)
$$\left(\mathrm{I} + \tilde{\mathrm{A}}_1 s\right)\left(\mathrm{C}_0 + \mathrm{C}_1 s + \mathrm{C}_2 s^2 + \ldots\right)\left(\mathrm{C}_0 + \mathrm{C}_1 s + \mathrm{C}_2 s^2 + \ldots\right) = -\tilde{\mathrm{E}}_0 - \tilde{\mathrm{E}}_1 s.$$

It is a property of matrices of convolution type that the expression (6.2.18) can be computed element-wise and further ordered with respect to equal powers of s such that the resulting operator is of convolution type. That is we are allowed to reorder, for example, $\mathrm{C}_1 s \mathrm{C}_0$ to $\mathrm{C}_1\mathrm{C}_0 s$. (This property, of course, is not generally true.) Now, as is evident by comparing coefficients, to solve (6.2.18) we must first find C_0 such that $\mathrm{C}_0^2 = -\tilde{\mathrm{E}}_0$. Because the matrix $\tilde{\mathrm{E}}_0$ is a quotient of two lower triangular matrices, the matrix C_0 is also lower triangular. Further, the diagonal entries C_0 can be chosen such that $\mathrm{Re}(\mathrm{C})_{jj} > 0$, $j = 1, \ldots, k$, through the Algorithm (3) below. Hence all poles corresponding to C_0**,** with C_0 constructed according to Algorithm 3, are located in the right half of the complex plane.

Algorithm 3 Calculate $\mathrm{C}_0 = \sqrt{-\tilde{\mathrm{E}}_0}$

```
for i = 1 to k do
  c_ii = sqrt(-Ẽ_{0,ii})
  if i > 1 then
    for j = i − 1 to1 do
      c_ij = ( −Ẽ_{0,ij} − Σ_{m=j+1}^{i−1} c_im c_mj ) / (c_ii + c_jj)
    end for
  end if
end for
```

Subsequently the sequence $\mathrm{C}_1, \mathrm{C}_2, \ldots, \mathrm{C}_{n-1}$ for the following n propagation steps is obtained by comparing coefficients of equal powers of s in (6.2.18). The corresponding pseudo-code is given in Algorithm 4, and consists mainly of solutions of Sylvester equations. From the structure of the algorithm we observe that if C_0 is computed, the entire sequence is uniquely determined.

Algorithm 4 Recursive calculation of C_j, $j = 1, \ldots, n-1$

$\mathrm{Z} := -\left(\mathrm{E}_1 + \mathrm{A}_1\mathrm{C}_0^2\right)$
Compute C_1 from $\mathrm{C}_1\mathrm{C}_0 + \mathrm{C}_0\mathrm{C}_1 = \mathrm{Z}$
for $k = 2$ to $n-1$ **do**
  $\mathrm{Z} \Leftarrow -\mathrm{A}_1\mathrm{Z}$
  Compute C_k from $\mathrm{C}_k\mathrm{C}_0 + \mathrm{C}_0\mathrm{C}_k = \mathrm{Z} - \sum_{j=1}^{k-1} \mathrm{C}_j\mathrm{C}_{k-j}$
end for

Finally, the boundary condition at every step $0 < i \leq n$ is obtained from C_j, $j = 0, \ldots n-1$, by employing the definitions $\mathrm{B} = -\mathrm{C}$ and

$$\begin{aligned} \mathrm{g}'_{i,0} &= \mathrm{B}(s)\mathrm{g}_{i,0} \\ &= \left(\mathrm{B}_0 + s\mathrm{B}_1 + \ldots + s^{i-1}\mathrm{B}_{i-1}\right)\mathrm{g}_{i,0} \\ &= \mathrm{B}_0\mathrm{g}_{i,0} + \mathrm{B}_1\mathrm{g}_{i-1,0} + \ldots + \mathrm{B}_{i-1}\mathrm{g}_{1,0}\,. \end{aligned} \tag{6.2.19}$$

Equation (6.2.19) provides the algorithmic basis for constructing nonlocal boundary conditions for any wide-angle approximation and discrete propagation method, as different propagation methods can be distinguished simply through the values of the defining coefficients $a'_1, \ldots, a'_k$ and $a_1, \ldots, a_k$.. As the operator $\mathrm{B}(s)$ possesses a Taylor representation in s, its action can be represented by matrix-vector multiplications of the Taylor coefficients B_j with boundary values describing the history of the evolution process. In our numerical implementation we order the boundary vectors $\mathrm{g}_{i,0}$ from lower to larger step numbers and introduce a composite boundary vector $\mathrm{g}_0 = \left(\mathrm{g}^T_{1,0}, \ldots\, \mathrm{g}^T_{i-1,0}, \mathrm{g}^T_{i,0}\right)^T$. Similarly, we generate the composite boundary matrices

$$\mathrm{B} = \left(\mathrm{B}_{i-1}, \mathrm{B}_{i-2}, \ldots\, \mathrm{B}_0\right), \quad 1 \leq i \leq n \tag{6.2.20}$$

$$\mathrm{C} = -\mathrm{B}. \tag{6.2.21}$$

in place of the boundary operator $\mathrm{B}(s)$, after which the normal derivative $\mathrm{g}'_{i,0}$ from system (6.2.19) is computed through a matrix-vector multiplication. This procedure for implementing the discrete boundary condition in terms of a composite matrix C is summarized in Algorithm 7.

Finite-Element Discretization.

From the representation (6.2.11), we will now generate a finite-element discretization on the interior domain. For illustrative purposes, consider the first equation of the system (6.2.11) at step i, $0 < i \leq n$. Multiplying this equation by a trial function $v \in H^1(\Omega)$, $\Omega = (x_-, x_+)$, integrating over Ω, and integrating by parts yields

$$\begin{aligned} \left(v, g_i^{(1)}\right) + \left(\partial_x v, a_1 \partial_x g_i^{(1)}\right) - \left(\bar{v} a_1 \partial_x g_i^{(1)}\right)\Big|_{x_-}^{x_+} = \\ (v, u_{i-1}) + (\partial_x v, a'_1 \partial_x u_{i-1}) - (\bar{v} a'_1 \partial_x u_{i-1})|_{x_-}^{x_+}\,. \end{aligned} \tag{6.2.22}$$

The variational problem corresponding to this equation is therefore to find a function $g_i^{(1)} \in H^1(\Omega)$ such that (6.2.22) holds for any $v \in H^1(\Omega)$. The other equations of the system (6.2.11) can be reformulated similarly. The resulting system is then discretized by replacing the infinite-dimensional function space $H^1(\Omega)$ by a finite-dimensional space $V_h = \mathrm{span}\{v_1, v_2, \ldots, v_N\}$ with $V_h \subset H^1(\Omega)$. Hence the corresponding discrete problem is to determine a discrete approximation $g_{h,i}^{(1)}$ of $g_i^{(1)}$

with $g_{h,i}^{(1)} \in V_h$ such that for all $v_h \in V_h$

$$\begin{aligned}(6.2.23)\qquad &\left(v_h, g_{h,i}^{(1)}\right) + \left(\partial_x v_h, a_1 \partial_x g_{h,i}^{(1)}\right) - \left(\overline{v}_h a_1 \partial_x g_{h,i}^{(1)}\right)\Big|_{x_-}^{x_+} \\ &= \quad (v_h, u_{h,i-1}) + (\partial_x v_h, a_1' \partial_x u_{h,i-1}) - (\overline{v}_h a_1' \partial_x u_{h,i-1})|_{x_-}^{x_+}\,.\end{aligned}$$

To solve the above system, we employ standard linear C^0-elements. Let v_1 and v_N denote the leftmost and the rightmost finite elements. A compact notation for (6.2.23) results if we set $v_1|_{x_-} = 1,\ v_N|_{x_+} = 1$, and define the vectors $\mathrm{b}_i^{(1)}, \mathrm{b}_i'^{(1)} \in \mathcal{C}^N$ with $N = \dim V_h$ by

$$\mathrm{b}_i^{(1)} = \begin{pmatrix} \left(-a_1 \partial_x g_{h,i}^{(1)}\right)\Big|_{x=x_-} \\ 0 \\ \vdots \\ 0 \\ \left(a_1 \partial_x g_{h,i}^{(1)}\right)\Big|_{x=x_+} \end{pmatrix} \text{ and } \mathrm{b}_i'^{(1)} = \begin{pmatrix} (-a_1' \partial_x u_{h,i-1})|_{x=x_-} \\ 0 \\ \vdots \\ 0 \\ (a_1' \partial_x u_{h,i-1})|_{x=x_+} \end{pmatrix},$$

and introduce the mass and stiffness matrices $\mathrm{M} \in \mathbb{R}^{N\times N}$, $\mathrm{A}_1, \mathrm{A}_1' \in \mathbb{C}^{N\times N}$ in standard fashion as $(\mathrm{M})_{i,j} = (v_{h,i}, v_{h,j})$ and $(A)_{i,j} = (\partial_x v_{h,i}, a_1 \partial_x v_{h,j})$. Defining as well the vectors $\mathrm{g}_i^{(1)} = \left(g_{h,i,1}^{(1)}, \ldots, g_{h,i,N}^{(1)}\right)^T \in \mathbb{C}^N$ and $\mathrm{u}_{i-1} = (u_{h,i-1,1}, \ldots, u_{h,i-1,N})^T \in \mathbb{C}^N$, that are the discrete counterparts of the continuous functions $g_i^{(1)}(x)$, $u_i(x)$ we have

$$(6.2.24)\qquad (\mathrm{M} + \mathrm{A}_1)\,\mathrm{g}_i^{(1)} - \mathrm{b}_i^{(1)} = (\mathrm{M} + \mathrm{A}_1')\,\mathrm{u}_{i-1} - \mathrm{b}_i'^{(1)}.$$

If we know the solution u_{i-1} in the interior domain together with its normal derivative on the boundary and the normal derivative of $\mathrm{g}_i^{(1)}$ we can obtain the unknown intermediate vector $\mathrm{g}_i^{(1)}$. Repeating this procedure for each of the equations of (6.2.11), we generate the following block matrix equation in terms of the matrices and vectors introduced in the preceding paragraph,

$$\begin{aligned}(6.2.25)\quad &\begin{pmatrix} \mathrm{M}+\mathrm{A}_1 & & & \\ & \mathrm{M}+\mathrm{A}_2 & & \\ & & \ddots & \\ & & & \mathrm{M}+\mathrm{A}_k \end{pmatrix} \begin{pmatrix} \mathrm{g}_i^{(1)} \\ \mathrm{g}_i^{(2)} \\ \vdots \\ \mathrm{u}_i \end{pmatrix} - \begin{pmatrix} \mathrm{b}_i^{(1)} \\ \mathrm{b}_i^{(2)} \\ \vdots \\ \mathrm{b}_i^{(k)} \end{pmatrix} \\ &= \begin{pmatrix} \mathrm{M}+\mathrm{A}_1' & & & \\ & \mathrm{M}+\mathrm{A}_2' & & \\ & & \ddots & \\ & & & \mathrm{M}+\mathrm{A}_k' \end{pmatrix} \begin{pmatrix} \mathrm{u}_{i-1} \\ \mathrm{g}_i^{(1)} \\ \vdots \\ \mathrm{g}_i^{(k-1)} \end{pmatrix} - \begin{pmatrix} \mathrm{b}_i'^{(1)} \\ \mathrm{b}_i'^{(2)} \\ \vdots \\ \mathrm{b}_i'^{(k)} \end{pmatrix}.\end{aligned}$$

To solve the system (6.2.25), the vectors $\mathrm{b}_i^{(j)}$, $\mathrm{b}_i'^{(j)}$, $j = 1, \ldots k$ must be constructed in accordance with the boundary conditions. The relationship between the discretized evolution equation (6.2.25) and the boundary condition (6.2.19) is determined by first decomposing the boundary condition at each boundary according to

$\mathrm{b}_i^{(j)} = \mathrm{b}_{i,-}^{(j)} + \mathrm{b}_{i,+}^{(j)}$ with

$$\mathrm{b}_{i,-}^{(j)} = \begin{pmatrix} \left(-a_j \partial_x g_{i,h}^{(j)}\right)\Big|_{x=x_-} \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \text{ and } \mathrm{b}_{i,+}^{(j)} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \left(a_j \partial_x g_{i,h}^{(j)}\right)\Big|_{x=x_+} \end{pmatrix}, \ \mathrm{b}_{i,\pm}^{(j)} \in \mathbb{C}^N \ .$$

Performing the same decomposition for each vector $\mathrm{b}_i'^{(j)}$ and assembling all nonzero entries of the vectors $\mathrm{b}_{i,\pm}^{(j)}$, $\mathrm{b}_{i,\pm}'^{(j)}$, $j = 1, \dots k$ into the four vectors $\mathrm{b}_{i,\pm}$, $\mathrm{b}'_{i,\pm} \in \mathbb{C}^k$ we arrive at

(6.2.26)

$$\mathrm{b}_{i,\pm} = \begin{pmatrix} \left(\pm a_1 \partial_x g_{i,h}^{(1)}\right)\Big|_{x=x_\pm} \\ \left(\pm a_2 \partial_x g_{i,h}^{(2)}\right)\Big|_{x=x_\pm} \\ \vdots \\ \left(\pm a_{k-1} \partial_x g_{i,h}^{(k-1)}\right)\Big|_{x=x_\pm} \\ (\pm a_k \partial_x u_i)|_{x=x_\pm} \end{pmatrix} \text{ and } \mathrm{b}'_{i,\pm} = \begin{pmatrix} \pm (a_1' \partial_x u_{i-1})|_{x=x_\pm} \\ \pm \left(a_2' \partial_x g_{i,h}^{(1)}\right)\Big|_{x=x_\pm} \\ \vdots \\ \pm \left(a_{k-1}' \partial_x g_{i,h}^{(k-2)}\right)\Big|_{x=x_\pm} \\ \pm \left(a_k' \partial_x g_{i,h}^{(k-1)}\right)\Big|_{x=x_\pm} \end{pmatrix}.$$

We now derive an equation relating the vectors $\mathrm{b}_{i,\pm}$ and $\mathrm{b}'_{i,\pm}$ to the boundary condition (6.2.19). Regarding first $\mathrm{b}_{i-1,+}$, we have from (6.2.26) and (6.2.19)

$$\begin{aligned} \mathrm{b}_{i,+} &= \mathrm{diag}(a_1, \dots, a_k) \mathrm{g}'_{0,i} \\ &= \mathrm{diag}(a_1, \dots, a_k) \mathrm{B}(s) \mathrm{g}_{0,i} \\ &= \mathrm{diag}(a_1, \dots, a_k) \left(\mathrm{B}_0 + \mathrm{B}_1 s + \mathrm{B}_2 s^2 + \dots + \mathrm{B}_{i-1} s^{i-1}\right) \mathrm{g}_{0,i}. \end{aligned}$$

All of the expressions $\mathrm{B}_j s^j \mathrm{g}_n(0) = \mathrm{B}_j \mathrm{g}_{n-j}(0)$ above with $j = 1, \dots, n-1$ can be immediately evaluated based on the observation that the shift operator s decreases the index of B_j by unity. Further, as the matrix B_0 is a lower triangular matrix we can arrange the algorithm such that the boundary condition at the current step only depends on boundary values at previous steps. To this end we decompose $\mathrm{b}_\pm$ as

$$\begin{aligned} \mathrm{b}_\pm &= \mathrm{B}_{\mathrm{d},\pm} \mathrm{g}_{0,i} + \mathrm{B}_{\mathrm{r},\pm} \mathrm{g}_i \\ \text{(6.2.27)} \quad \text{with } \mathrm{B}_{\mathrm{d},\pm} &:= \pm \,\mathrm{diag}(a_1, \dots, a_k)|_{x=x_\pm} \,\mathrm{diag}\left(\mathrm{B}_{0,\pm}\right) \\ \text{and } \mathrm{B}_{\mathrm{r},\pm}(s) &:= \pm \,\mathrm{diag}(a_1, \dots, a_k)|_{x=x_\pm} \\ \text{(6.2.28)} \qquad &\quad \cdot \left(\mathrm{B}_0 - \mathrm{diag}\left(\mathrm{B}_0\right) + s\mathrm{B}_1 + \dots + s^{i-1}\mathrm{B}_{i-1}\right)\big|_{x=x_\pm} , \end{aligned}$$

where the matrices $\mathrm{diag}\left(\mathrm{B}_{0,\pm}\right)$ contain only the main diagonals of $\mathrm{B}_{0,\pm}$. The diagonal matrices $\mathrm{B}_{\mathrm{d},\pm}$ are then inserted into the matrix of the final system which yields updated matrices A_j satisfying

$$\text{(6.2.29)} \qquad \mathrm{A}_j = \mathrm{A}_j - \begin{pmatrix} \mathrm{B}_{\mathrm{d},-}(j,j) & & \\ & 0 & \\ & & \mathrm{B}_{\mathrm{d},+}(j,j) \end{pmatrix}, \quad j = 1, \dots k\,.$$

The reduced matrix $\mathrm{B}_{\mathrm{r},\pm}$, which is a lower triangular matrix with a zero diagonal, only couples previously determined boundary values and is therefore placed on the right-hand side of the evolution equation, cf. Algorithm 7.

A corresponding expression for b'_+ results from the observation that the vector g_i is ordered as $\left(g_i^{(1)}, g_i^{(2)}, \dots\, g_i^{(k-1)}, u_i\right)^T$, reflecting the algebraic structure of the

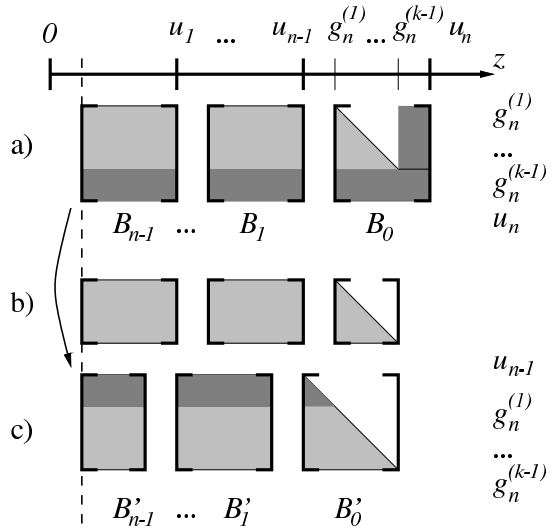


FIGURE 6.2.1. Construction of the boundary operator B′ from the boundary operator B

boundary condition (6.2.19). The discrete evolution system (6.2.25), however, requires the alternate ordering (6.2.26) $\left(u_{i-1}, g_i^{(1)}, g_i^{(2)}, \ldots\ g_i^{(k-1)}\right)^T$. Thus we obtain a condition of the form

$$\begin{pmatrix} \partial u_{i-1} \\ \vdots \\ \partial g_i^{(k-1)} \end{pmatrix} = \left(\mathrm{B}'_0 + \mathrm{B}'_1 s + \ldots + \mathrm{B}1_{i-1} s^{i-1}\right) \begin{pmatrix} u_{i-1} \\ \vdots \\ g_i^{(k-1)} \end{pmatrix}$$

with an operator $\mathrm{B}'(s) := \mathrm{B}'_0 + \mathrm{B}'_1 s + \ldots + \mathrm{B}'_{i-1} s^{i-1}$, if we rearrange the columns and rows of the operator B appropriately.

In the first step of this transformation, which is illustrated in Fig. 6.2.1 we remove the last row and the last column of the composite boundary matrix $\mathrm{B} = -\mathrm{C}$ to obtain the reduced matrix shown in part b) of Fig. 6.2.1. We then place the former last row of B at the top of the reduced matrix, after shifting the row to the left and adjusting it to the dimension of the reduced matrix, as illustrated in part c) of Fig. 6.2.1. Finally, the resulting matrix is multiplied with the diagonal matrix $\mathrm{diag}\,(a'_1, \ldots\ a'_k)$ according to (6.2.26). The pseudo-code for these operations is given in Algorithm 5.

Algorithm 5 Computation of the operator B′ (see Fig. 6.2.1)

1: Compute $\mathrm{B}' := -\mathrm{C}(1 : k-1, 1 : nk-1)$

2: Compute $\mathrm{b}' := -\left[\mathrm{C}(k, k+1 : nk), \underbrace{0, 0, \ldots\ 0}_{k-1}\right]$

3: Compute $\mathrm{B}' \Leftarrow \mathrm{diag}\,(a'_1, \ldots\ a'_k) \begin{bmatrix} 0 & \mathrm{b}' \\ 0 & \mathrm{B}' \end{bmatrix}$

To evolve the field over n propagation steps, we must first initialize the two boundary operators B_+ and B_- with dimension $k \times (kn)$, acting on the right and left boundary, respectively. At the end of the simulation we will then possess two vectors g_+ and g_- of dimension $nk+1$ that contain the required boundary values.

The initialization of the propagation algorithm, which includes the computation of the standard finite element matrices, the updating of these matrices (which corresponds to incorporating the boundary conditions) and the computation of the boundary matrices $\mathrm{B}_{\pm}$ are summarized in Algorithm 6.

Algorithm 6 Computation of the finite element matrices A_j, M and the boundary matrices $\mathrm{B}_{\pm}$

Compute $\mathrm{C}_{0,\pm}$	{acc. to Algorithm 3}
Compute $\mathrm{C}_{j,\pm}$, $j = 1 : n-1$	{acc. to Algorithm 4}
Compute $\mathrm{C}_{\pm} := [\mathrm{C}_{n-1}, \mathrm{C}_{n-2}, \dots \mathrm{C}_0]_{\pm}$	{acc. to (6.2.21)}
Compute A_j, $j = 1, \dots, k$ and M	{acc. to (6.2.23), (6.2.24)}
Compute $\mathrm{B}'_{\pm}$	{acc. to Algorithm 5}
Compute B_{d}	{acc. to (6.2.27)}
Update A_j, $j = 1 : k$ using B_{d}	{acc. to (6.2.29) }
Compute $\mathrm{B}_{\pm} := -\mathrm{diag}(a_1, \dots, a_k)_{\pm} [\mathrm{C}_{n-1}, \mathrm{C}_{n-2}, \dots \mathrm{C}_0 - \mathrm{diag}(\mathrm{C}_0)]_{\pm}$	{acc. to (6.2.28)}
$\mathrm{B}_{\pm} \Leftarrow [0, \mathrm{B}_{\pm}(:, 1 : nk-1)]$	
$\mathrm{B}_{\pm} := [\mathrm{B}_{n-1}, \mathrm{B}_{n-2}, \dots \mathrm{B}_0] \Leftarrow \mathrm{B}_{\pm} - \mathrm{B}'_{\pm}$	

The structure and the numerical details of the resulting propagation algorithm is finally described in Algorithm 7.

Algorithm 7 Propagation algorithm

$\mathrm{g} := \mathrm{u}_0$	{set the initial values}
$\mathrm{g}_- := (\mathrm{g}(1))$, $\mathrm{g}_+ := (\mathrm{g}(N))$	{save the boundary values}
for $i = 1$ to n **do** {propagate n steps}	
for $j = 1$ to k **do** {solve k intermediate problems}	
$\mathrm{b} = (\mathrm{M} + \mathrm{A}'_j)\,\mathrm{g}$	
$c_- = \mathrm{B}_-(j, (n-i)k+1 : (n-1)k+j)\mathrm{g}_-$	
$c_+ = \mathrm{B}_+(j, (n-i)k+1 : (n-1)k+j)\mathrm{g}_+$	
$\mathrm{b} \Leftarrow \mathrm{b} + \begin{pmatrix} c_- \\ 0 \\ c_+ \end{pmatrix}$	
Compute g from $(\mathrm{M} + \mathrm{A}_j)\,\mathrm{g} = \mathrm{b}$	
$\mathrm{g}_- \Leftarrow (\mathrm{g}_-, \mathrm{g}(1))^T$, $\mathrm{g}_+ \Leftarrow (\mathrm{g}_+, \mathrm{g}(N))^T$	{save the boundary values}
end for	
$\mathrm{u}_i := \mathrm{g}$	{solution of the i-th step}
end for	

6.3. Stability

The stability of the wide-angle propagation Algorithm 7 can be verified through a natural extension of our earlier analysis for the Schrödinger-type Padé-(2,0) approximant **[85]**. We summarize the result in

THEOREM 6.3.1. *Let the finite element space V_h used for the discretization (6.2.23) be independent of the step number. Let the coefficients a' and a of the discrete evolution equation (6.2.10) are complex conjugates with* $\mathrm{Im}(a) \neq 0$. *Then the wide angle propagation Algorithm 7 together with the factorization Algorithms 3 and 4 is stable for each stepsize Δz. In particular, both $L^2(\mathbb{R})$ and $H^1(\mathbb{R})$ norms of the solution are conserved along the propagation.*

PROOF. We start from the representation of the exterior solution given by (6.2.17)

$$\widehat{\mathrm{g}}_i(p) = (p\mathrm{I} + \mathrm{C}(s))^{-1}\,\mathrm{g}_{i,0}, \quad \text{with } \mathrm{C}(s) = \mathrm{C}_0 + \mathrm{C}_1 s + \ldots + \mathrm{C}_{i-1}s^{i-1},$$

in which by construction (Algorithm 3) $\mathrm{Re}\,(\mathrm{diag}(\mathrm{C}_0)) > 0$. The operator $p\mathrm{I} + \mathrm{C}(s)$ is of convolution type. Its inverse exists and is again of convolution type $(p\mathrm{I} + \mathrm{C}(s))^{-1} = \mathrm{D}_0 + \mathrm{D}_1 s + \ldots$, as long $\det(p\mathrm{I} + \mathrm{C}_0) \neq 0$. Further it holds $\mathrm{C}_0^{-1} = \mathrm{D}_0$. Hence the zeros of $\det(p\mathrm{I} + \mathrm{C}_0)$ are exactly the poles of the solution vector $\widehat{\mathrm{g}}_i(p)$. Consequently, $\mathrm{g}_i(x)$ consists only of exponentially decaying functions, in particular $\lim_{x\to\infty} \mathrm{g}_i(x) = 0$.

Let $\Omega_- = (-\infty, x_-)$, $\Omega_+ = (x_+, \infty)$, and $\Omega = (x_-, x_+)$. The unconditional stability of the propagation algorithm follows directly from this conservation law.

We consider first the discrete variational equation, (6.2.23), and abbreviate $g_{h,i}^{(1)}$ by g and $u_{h,i-1}$ by u and a_1, a_1' by a, a'. Accordingly, (6.2.23) reads

$$(v_h, g) + (\partial_x v_h, a\partial_x g) - (\bar{v}_h a \partial_x g)|_{x_-}^{x_+} = (v_h, u) + (\partial_x v_h, a'\partial_x u) - (\bar{v}_h a' \partial_x u)|_{x_-}^{x_+}.$$

We must compute the function g from its predecessor u. If we regard the special choice $v_h = g$ and $v_h = u$, we obtain the equation system

$$\begin{aligned}(g,g) + (\partial_x g, a\partial_x g) - (\bar{g} a \partial_x g)|_{x_-}^{x_+} &= (g,u) + (\partial_x g, a'\partial_x u) - (\bar{g} a' \partial_x u)|_{x_-}^{x_+}\\ (u,g) + (\partial_x u, a\partial_x g) - (\bar{u} a \partial_x g)|_{x_-}^{x_+} &= (u,u) + (\partial_x u, a'\partial_x u) - (\bar{u} a' \partial_x u)|_{x_-}^{x_+}.\end{aligned}$$

We compute the conjugate complex of the second equation

$$(g,u) + (a\partial_x g, \partial_x u) - \left(u\overline{a\partial_x g}\right)\Big|_{x_-}^{x_+} = (u,u) + (a'\partial_x u, \partial_x u) - \left(u\overline{a'\partial_x u}\right)\Big|_{x_-}^{x_+}.$$

The sum of the last equation and the first of the above system yields

$$\begin{aligned}&(g,g) + (\partial_x g, a\partial_x g) - \left(\bar{g} a \partial_x g + u\overline{a\partial_x g}\right)\Big|_{x_-}^{x_+}\\ &\quad = (u,u) + (a'\partial_x u, \partial_x u) - \left(\bar{g} a' \partial_x u + u\overline{a'\partial_x u}\right)\Big|_{x_-}^{x_+}.\end{aligned}$$

The same procedure applies to the two exterior domains. For the right exterior domain

$$\begin{aligned}&(g,g)_+ + (\partial_x g, a\partial_x g)_+ - \left(\bar{g} a \partial_x g + u\overline{a\partial_x g}\right)\Big|_{x_+}^{\infty}\\ &\quad = (u,u)_+ + (a'\partial_x u, \partial_x u)_+ - \left(\bar{g} a' \partial_x u + u\overline{a'\partial_x u}\right)\Big|_{x_+}^{\infty}\end{aligned}$$

with an analogous result for the left exterior domain. Summing all three contributions and noting that the terms at infinity vanish while the normal derivatives at $x_\pm$ cancel we obtain

$$\begin{aligned}&(g,g)_- + (g,g) + (g,g)_+ + (\partial_x g, a\partial_x g)_- + (\partial_x g, a\partial_x g) + (\partial_x g, a\partial_x g)_+\\ &\quad = (u,u)_- + (u,u) + (u,u)_+ + (a'\partial_x u, \partial_x u)_- + (a'\partial_x u, \partial_x u) + (a'\partial_x u, \partial_x u)_+\end{aligned}$$

The imaginary part of the above equation yields

$$\begin{aligned}&\mathrm{Im}\,(a)\left[(\partial_x g, \partial_x g)_- + (\partial_x g, \partial_x g) + (\partial_x g, \partial_x g)_+\right]\\ &\quad = -\mathrm{Im}\,(a')\left[(\partial_x u, \partial_x u)_- + (\partial_x u, \partial_x u) + (\partial_x u, \partial_x u)_+\right]\end{aligned}$$

Since a and a' are complex conjugates with $\Im(a) \neq 0$, we obtain the conservation law

$$\begin{aligned}&(\partial_x g, \partial_x g)_- + (\partial_x g, \partial_x g) + (\partial_x g, \partial_x g)_+\\ &\quad = (\partial_x u, \partial_x u)_- + (\partial_x u, \partial_x u) + (\partial_x u, \partial_x u)_+.\end{aligned}$$

This means that the $H^1(\mathbb{R})$ semi-norm of two subsequent steps is conserved and implies the conservation of the $L^2(\mathbb{R})$ norm

$$(g,g)_- + (g,g) + (g,g)_+ = (u,u)_- + (u,u) + (u,u)_+.$$

Since the $L^2(\mathbb{R})$ norm is conserved for every intermediate step, it is conserved over the entire discrete evolution. Since the initial data u_0 is required to be compactly supported on the interior domain, we obtain the bound

$$||g_i||_{L^2(\Omega)}, i \geq 0 \leq ||u_0||_{L^2(\Omega)}$$

. □

6.4. Numerical Experiments

We now verify our theoretical considerations by computing the reflection from the computational window boundary of a Gaussian input beam described by

$$u(x,0) = u_0(x) = \text{const } \exp\left(-(x/10)^2\right) \exp\left(ik_0 x \sin\phi\right)$$

propagating in air. We set $k(x) = k_0 = 2\pi/\lambda$ where the free space wavelength $\lambda = 1.55\mu m$, the propagation step size $\Delta z = 0.4\mu m$ and $\phi = \pi/4$. In our calculations which are meant to duplicate the corresponding numerical experiments in a predecessor to this work **[32]**, the computational domains are either $\Omega = (-50, 50) \times (0, 400)\mu m^2$ or $\Omega = (-50, 50) \times (0, 100)\mu m^2$ while the transverse step sizes Δx vary between $0.01\mu m \leq \Delta x \leq 0.2\mu m$. We consider first the intrinsic error associated with applying the implicit mid-point rule to plane wave solutions of the Helmholtz equation. Recall that for the exact one-way Helmholtz propagator

$$u(\Delta z) = u(0) \exp\left(-i\Delta z k_0 \sqrt{1 - \sin^2\phi}\right) ,$$

for which the implicit mid-point rule yields

$$u_{\text{ IMR}}(\Delta z) = u_{\text{ IMR}}(0) \frac{1 - i\Delta z k_0/2 \left(1 - \sqrt{1 - \sin^2\phi}\right)}{1 + i\Delta z k_0/2 \left(1 - \sqrt{1 - \sin^2\phi}\right)} .$$

In this expression ϕ, $-\pi/2 < \phi < \pi/2$ is the angle between the propagation direction and the z-axis. This yields a resulting phase error $\log(u(\Delta z)) - \log(u_{\text{ IMR}}(\Delta z))$ which we compare in Fig. 6.4.1 to that obtained by instead applying the Padé (2,0) (Schrödinger-type) approximation to the exact propagator. While at Padé order 2 the phase error of the Padé approximation is far greater, the opposite is true for a Padé (8,8) approximant as evident from Fig. 6.4.2.

Next we display in Fig. 6.4.3 the spectral norms of the matrices B_i for the boundary conditions associated with both the (2,0) and (8,8) Padé approximants as a function of the number of propagation steps. Here the matrices B_i are defined and computed as in Algorithm 6. Both approximations decay asymptotically as $||\mathrm{B}_i||_2 = \text{const } \mathrm{i}^{-3/2}$, independent of the order of the approximation. Note that every second coefficient of the Padé (2,0) approximation vanishes.

We now propagate the field from $z = 0$ to $z = 400\mu m$ with the (8,8) Padé procedure. The transverse grid spacing is here $\Delta x = 0.2\mu m$ while the computational domain is $\Omega = (-50, 50) \times (0, 400)\mu m^2$. The contour lines for the logarithmic amplitude over the first $100\mu m$ of propagation are shown in Fig. 6.4.4. While the incident field propagates as expected along the $\theta = \pi/4$-direction, residual reflections are generated by the finite transverse discretization error.

To underline the wide-angle property of the Padé $(8, 8)$ approximant, we first note that employing the $(2, 0)$ in place of the $(8, 8)$ Padé approximation for the square root operator leads to considerable phase errors, as evident from Fig. 6.4.5.

Performing the numerical experiment in the presence of second Gaussian beam that describes an an angle of $+\pi/4$with respect to the z-axis. Fig. 6.4.6 demonstrates the wide-angle nature of the underlying propagation method.

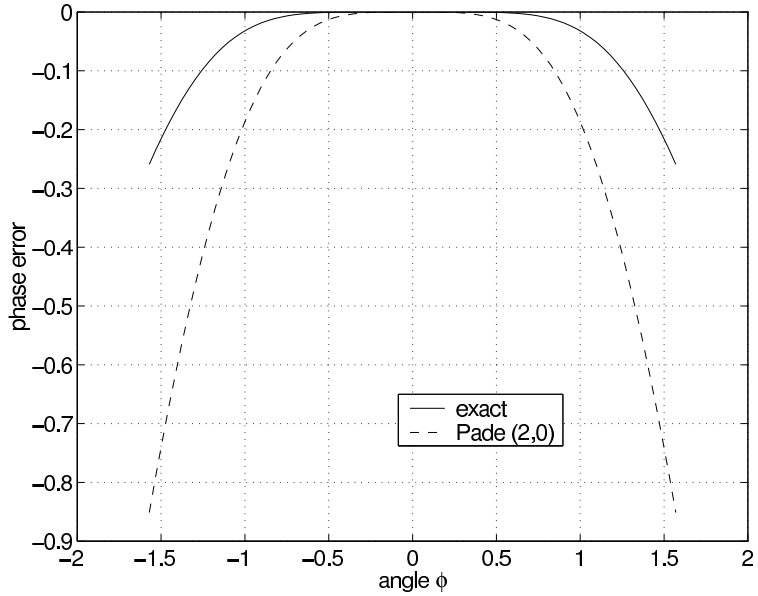


FIGURE 6.4.1. The phase error associated with the exact implicit mid-point discretization (solid line) compared to that of the corresponding Padé (2,0) approximant (dashed line)

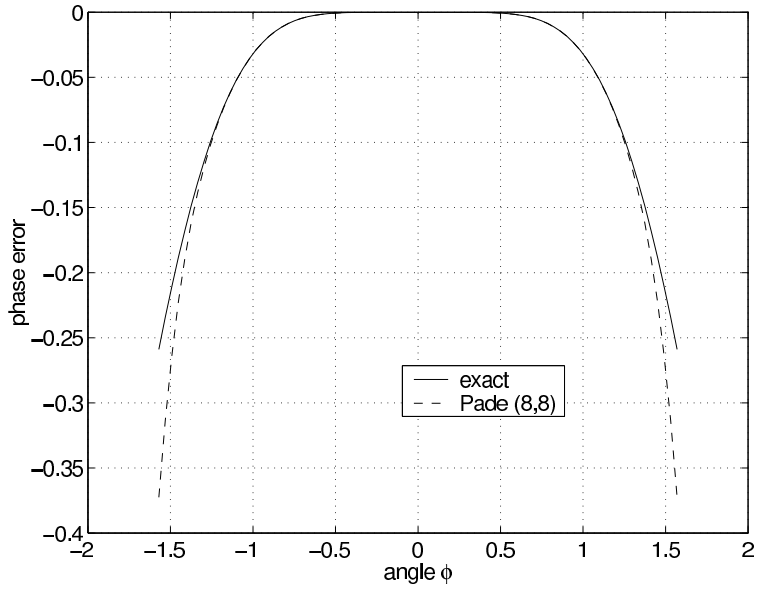


FIGURE 6.4.2. As in Fig. 6.4.1, but for a (8,8) Padé approximant.

The influence of the discretization error with respect to the transverse step length Δx, is evident if we repeat our previous $(8,8)$ Padé simulation, Fig. 6.4.4 with $\Delta x = 0.01\mu m$, cf. Fig. 6.4.7. The boundary reflection, which vanishes in the $\Delta x \to 0$ limit is indeed significantly reduced. The actual dependence of the reflection from the discretization error can be deduced from Fig. 6.4.8 which graphs the discrete $\mathrm{L}^2(-50,50)$ norm as a function of propagation distance for transverse step sizes of $0.2\mu m, 0.1\mu m, 0..05\mu m, 0.025\mu m$, and $0.01\mu m$ The figure clearly shows

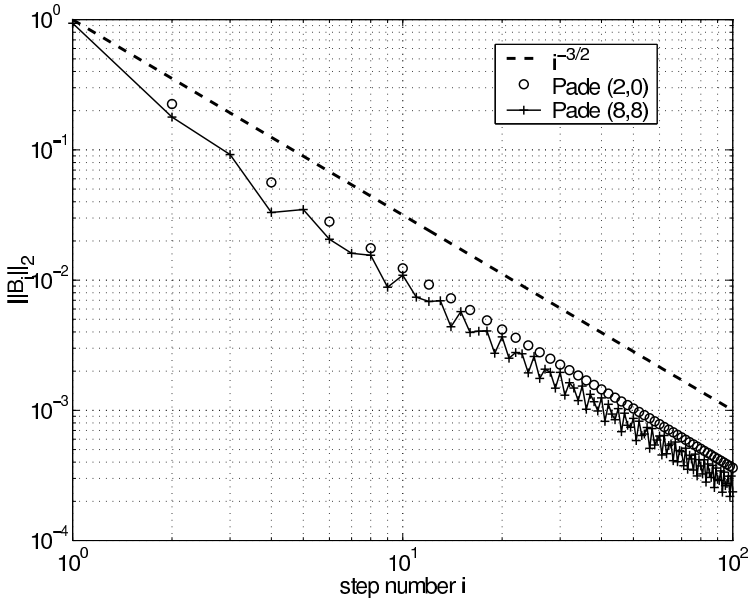


FIGURE 6.4.3. Spectral norm of the boundary matrices B_i as a function of the number of propagation steps.

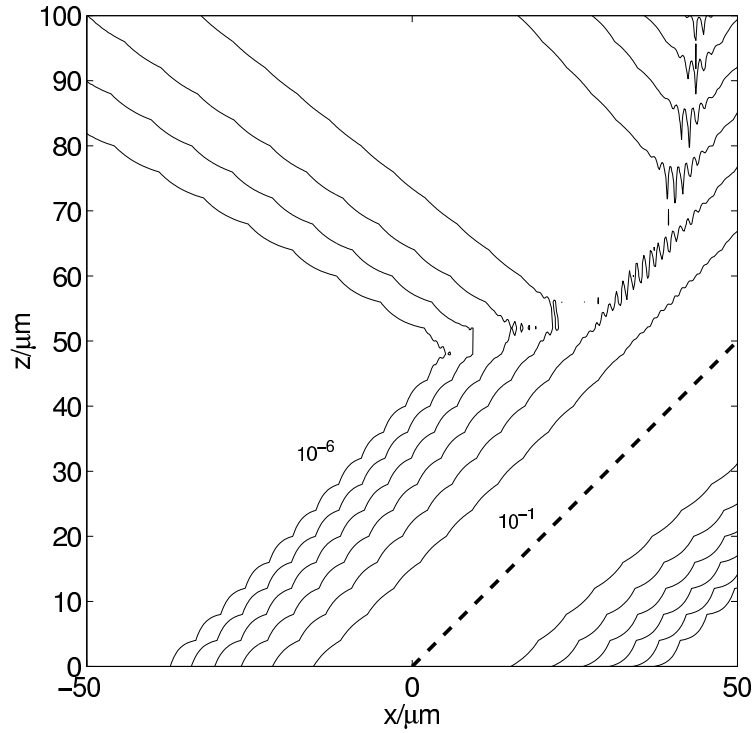


FIGURE 6.4.4. Gaussian beam propagation calculated with a (8,8) Padé propagator and $\Delta x = 0.2\mu m$. The dashed line represents the exact propagation angle $\theta = \pi/4$. The reflected field vanishes as $\Delta x \to 0$.

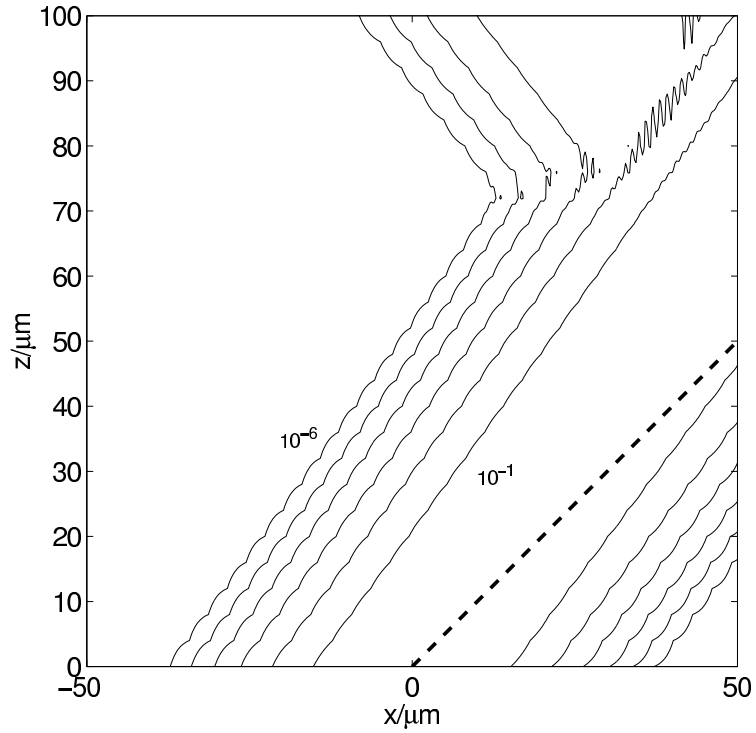


FIGURE 6.4.5. As in Fig. 6.4.4 except for a Padé (2,0) propagator. Note the error in the propagation angle.

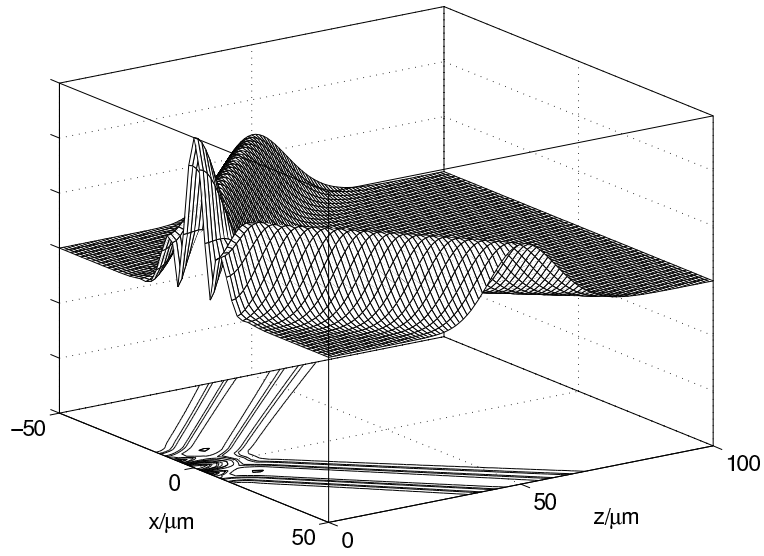


FIGURE 6.4.6. Propagation of two Gaussian beams at a relative angle of $\pi/2$.

that halving the transverse grid point spacing reduces the norm of the reflected field by a factor of 4. This behavior is entirely consistent with the $O\left(\Delta x^2\right)$ discretization error of the underlying linear finite elements. Finally, to demonstrate the stability of our algorithm (subject to arithmetic error), we display in Fig. 6.4.9 $||u||$ computed

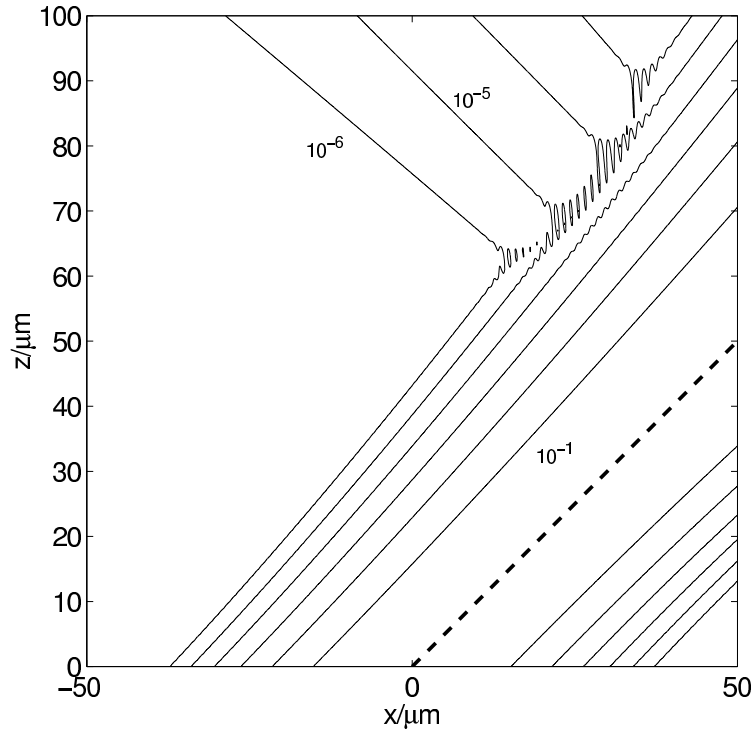


FIGURE 6.4.7. As in Fig. 6.4.4, but with $\Delta x = 0.01\mu m$.

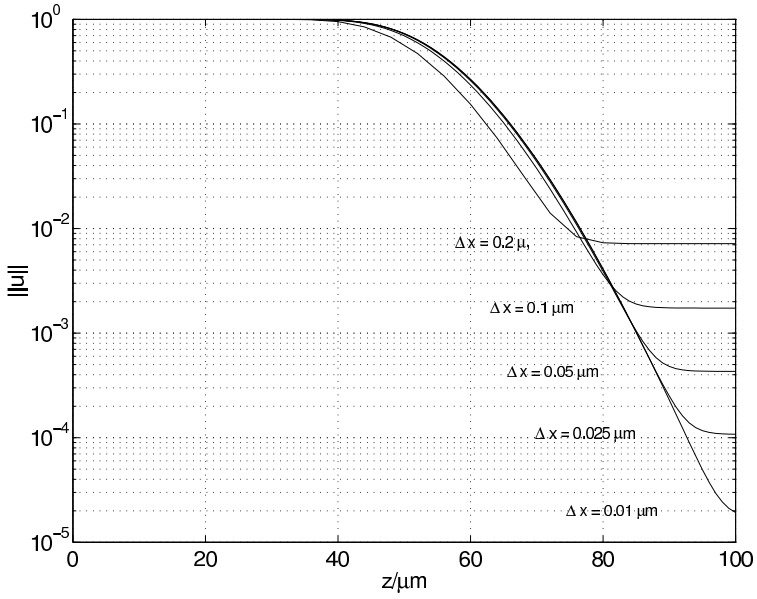


FIGURE 6.4.8. The $L^2(-50, 50)$ norm $||u||$ as a function of the propagation distance for varying step sizes Δx.

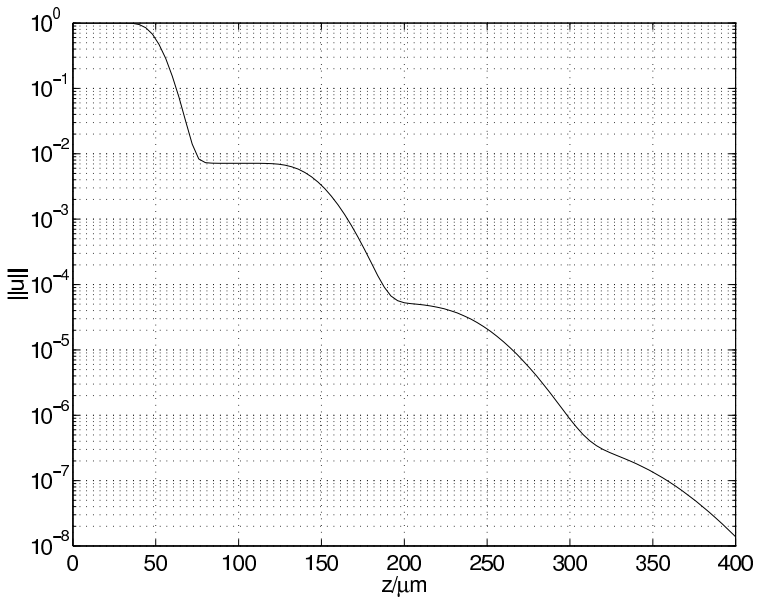


FIGURE 6.4.9. The discrete L_2-norm of the field within the computational window as a function of longitudinal distance for $0 \leq z \leq 400\mu m$ for the parameters of Fig. 6.4.4.

over a longer longitudinal interval $0 \leq z \leq 400\mu m$ for an (8,8) Padé propagator. The parameters are the same as in Fig. 6.4.4. Clearly the resulting curve, which displays successive plateaus corresponding to integer number of reflections of the Gaussian beam from the computational window boundaries indicates the stability of the algorithm.

Conclusions

Main Results. The central result of this work is a new characterization of wave propagation on unbounded media. The most important notion is the concept of *pole condition.* All asymptotically outgoing waves have to satisfy the pole condition. The pole condition generalizes the famous Sommerfeld radiation condition to problems with inhomogeneous exterior domains and is equivalent to Sommerfeld's condition in the simpler case of homogeneous exterior domains. The pole condition is constructive. It allows both a theoretical analysis of wave propagation and a construction of numerical alorithms. As a side-effect of the approach, a new representation formula for the solution in the exterior domain has been established. There are different possibilities for numerical realizations of the theoretical concept. For problems in space dimensions $d \geq 2$ we developed a new class of methods which we call *Laplace domain methods.* The Laplace domain methods solve the interior problem (given in the spatial domain) and the exterior problem (given in the *spectral* domain) simultaneously such that the pole condition is automatically satisfied.

The main focus of the numerical experiments is on the 2D Helmholtz equation. Nevertheless we applied the technique also to the time-dependent 1D Schrödinger equation and, in a generalization of this approach, to higher order approximations of the 1D one-way wide angle equations.

Open Problems and Current Research. Since our whole concept is new, many questions in theory, algorithms and applications are still open. We want to mention the most important ones.
Theory:

- In Section 4.2 we have shown the equivalence of the Sommerfeld radiation condition and the pole condition for the finite dimensional case. In a joint paper with Hohage and Zschiedrich **[58]**, we extended this to the infinite dimensional case. The technique used there is a perturbation analysis of Riccati's equation, which is far-away from the concept of our algorithms. Is there a proof closer related to the numerical algorithms?
- In our analysis we excluded potentials of the form $\mathcal{O}(1/|\mathbf{x}|)$ for $|\mathbf{x}| \to \infty$. The Laplace domain method in its axis integral form, cf. Section 5.2, however, is able to deal even with this type of potentials. How to extend the theory to include such potentials?
- Our theory is based throughout on separable problems which is a consequence of the applied modal analysis. In contrast, the numerical algorithms are able to deal with non-separable problems as well. Can we reformulate the theory to cover non-separable problems?

Algorithms:

- The Laplace domain methods analyzed in Sections 5.2 and 5.2.2 for separable problems in 2D have to be implemented and analyzed for general non-separable problems in 2D and 3D.
- The Laplace domain method for time-dependent problems, cf. p. 150, has to be implemented and analyzed. The central question here is whether or not it is possible to obtain stability results as in Theorem 6.3.1 for the Laplace domain method. This would offer the possibility to develop a new class of propagation algorithms.
- A detailed analysis of the operation account and memory requirements has to be done.

Applications:

- The discretization of the exterior domain is based on *linear* rays. To extend the applicability, curved rays might be considered also. Our consideration of hyperbolic-elliptic coordinates in Section 4, p. 59, can be seen as a first step in this direction.
- Further, it seems to be possible to extend the construction scheme to vectorial elements of Nédélec-type. This would allow to apply the pole condition idea to Maxwell problems. Such results will be contained in the PhD thesis of Zschiedrich **[97]**.
- The concept can be applied in a natural way to eigenproblems on unbounded domains. A first successful study of 1D eigenproblems based on the pole condition approach will be reported in the diploma thesis of Geng **[33]**. The numerical results clearly demonstrate the superiority of the pole condition approach over the PML approach.

APPENDIX A

Basic Properties of the Continuous and Discrete Transforms

A.1. Basic Properties of the Laplace Transform

For convenience, we assemble in the following some basic properties of Laplace transform, as far as we need them for our further elaboration. Our representation follows mainly the one given by Dettman in **[24]**. We assume the following:

(H1) The real-valued function $u(\xi)$ is twice continuously differentiable in any closed finite interval $[0,a]$, $u(\xi) \in C^2[0,a]$, $0 < a < \infty$.

(H2) The second derivative of u is exponentially bounded, $|u''(\xi)| \leq C\,\mathrm{e}^{d\xi}$, $0 \leq \xi < \infty$, for some real d.

An immediate consequence is that besides u'' also $u'$and u are also exponentially bounded. It follows directly $u, u' \leq C\mathrm{e}^{b\xi}$ with $b = \max(0,d)$.

LEMMA A.1.1. *The Laplace transform* $\mathrm{L}(u)$

$$\mathrm{L}(u) := \int_0^\infty \mathrm{e}^{-s\xi} u(\xi)\, d\xi,\ s \in \mathbb{C}$$

is analytic in the half-plane $\mathrm{Re}(s) > b$.

PROOF. The integral converges uniformly for $\mathrm{Re}(s) > b$ since $|u(\xi)| \leq C\mathrm{e}^{b\xi}$. Hence the derivative with respect to s can be computed under the integral sign, which shows the existence of the derivative. □

As usually, we denote the Laplace transform of $u(\xi)$ by $\widehat{u}(s) := \mathrm{L}\left(u(\xi)\right)(s)$ and its derivative by $\widehat{u}'(s)$ or by $\partial_s \widehat{u}$.

LEMMA A.1.2. *The Laplace transform of the expression* $\xi^n u(\xi)$ *is analytic for* $\mathrm{Re}(s) > b$. *It holds*

$$\mathrm{L}(\xi^n u) = (-1)^n \partial_s^n \widehat{u}(s)\,.$$

PROOF. Taking the first derivative in the half-plane $\mathrm{Re}(s) > b$ supplies $\widehat{u}'(s) = \int_0^\infty \partial_s \mathrm{e}^{-s\xi} u(\xi)\, d\xi$. This is equivalent to $\mathrm{L}\left(-\xi u(\xi)\right) = \widehat{u}'(s)$. Repeating this n-times gives the result. □

LEMMA A.1.3. *Let* u *satisfy the hypothesis (H1) and (H2). Then it holds in the half-plane* $\mathrm{Re}(s) > b$

$$\begin{aligned} \mathrm{L}(u'(\xi)) &= s\widehat{u}(s) - u(0) \\ \mathrm{L}(u''(\xi)) &= s^2\widehat{u}(s) - su(0) - u'(0) \end{aligned}$$

PROOF. The function $u(\xi)$, $u'(\xi)$, and $u''(\xi)$ are all exponentially bounded. Hence there Laplace transform exists in the half-plane $\mathrm{Re}(s) > b$. Integration by parts yields $\int_0^\infty \mathrm{d}\xi\, \mathrm{e}^{-s\xi} u' = -u(0) + s\int_0^\infty \mathrm{d}\xi\, \mathrm{e}^{-s\xi} u = -u(0) + s\widehat{u}(s)$. Setting $v'(\xi) = u''(\xi)$ and repeating the procedure for v'gives the second correspondence. □

COROLLARY A.1.4. *Let u satisfy the hypothesis (H1) and (H2). Then there is a constant M and a half-plane $\mathrm{Re}(s) \geq \rho$ such that*

$$|\widehat{u}(s)| \leq \frac{M}{|s|}\,.$$

PROOF. The exponential bound of $\widehat{u}'$ gives the estimate

$$|\widehat{u}'| \leq \int_0^\infty \mathrm{d}\xi\, C\mathrm{e}^{b\xi}\mathrm{e}^{-\mathrm{Re}(s)\xi} = \frac{C}{\mathrm{Re}(s)-b} \leq \frac{C}{\rho - b}$$

On the other hand, Lemma A.1.3 shows

$$\begin{aligned} |\widehat{u}(s)| &\leq \frac{|\widehat{u}'| + |u(0)|}{|s|} \\ &\leq \frac{\frac{C}{\rho-b} + |u(0)|}{|s|} \\ &= \frac{M}{|s|} \end{aligned}$$

where $M = \max\left(C\,/(\rho - b\,,\, |u(0)|\right)$. □

COROLLARY A.1.5. *Let u satisfy the hypothesis (H1) and (H2). It holds*

$$\lim_{\mathrm{Re(s)}\to\infty} s\widehat{u}(s) = u(0)\,.$$

This follows immediately from the fact that (H1) and (H2) it follows $\lim_{\mathrm{Re(s)}\to\infty} |\widehat{u}'(s)| = 0$.

LEMMA A.1.6. *Let u satisfy the hypothesis (H1) and (H2). Let s_∞ be defined as a point in the complex plane with some finite imaginary part and an infinite, positive real part. It holds in a half-plane $\mathrm{Re}(s) > b$*

$$\mathrm{L}\left(\frac{1}{\xi_0 + \xi}u(\xi)\right) = \mathrm{e}^{\mathrm{s}\xi_0}\int_s^{s_\infty} \mathrm{d}s_1\, \mathrm{e}^{-s_1\xi_0}\widehat{u}(s_1),$$

where the Laplace transformed expression remains analytic in this half-plane.

PROOF. Let

$$\widehat{g}(s) = \mathrm{e}^{\mathrm{s}\xi_0}\int_s^{s_\infty} \mathrm{d}s_1\, \mathrm{e}^{-s_1\xi_0}\widehat{u}(s_1),\ \text{in}\,\mathrm{Re}(s) > b$$

It follows

$$\begin{aligned} \widehat{g}'(s) &= \xi_0\mathrm{e}^{\mathrm{s}\xi_0}\int_s^{s_\infty} \mathrm{d}s_1\, \mathrm{e}^{-s_1\xi_0}\widehat{u}(s_1) - \widehat{u}(s) \\ &= \xi_0\widehat{g}(s) - \widehat{u}(s). \end{aligned}$$

By Lemma A.1.3,

$$u(\xi) = (\xi + \xi_0)\, g(x).$$

□

COROLLARY A.1.7. *It holds in a half-plane $\mathrm{Re}(s) > b$*

$$\mathrm{L}\left(\frac{1}{(\xi_0 + \xi)^n}u(\xi)\right) = \mathrm{e}^{\mathrm{s}\xi_0}\int_s^{s_\infty} \mathrm{d}s_n \ldots \int_{s_2}^{s_\infty} \mathrm{d}s_1\, \mathrm{e}^{-s_1\xi_0}\widehat{u}(s_1),$$

where the Laplace transformed expression remains analytic in this half-plane.

The iterated integrals of Corollary A.1.7 can be generalized to the following (obviously not widely known) lemma, which supplies the Laplace transform of the product of two functions $p\left(\frac{1}{\xi+\xi_0}\right)$ and $u(\xi)$ via a convolution along the real axis.

LEMMA A.1.8. *Let p be a power series of the form $p(t) = \sum_{m=1}^{\infty} p_m t^m$ with convergence radius $\rho > \frac{1}{\xi_0}$, further $v(\xi) = p\left(\frac{1}{\xi}\right)$, and let u be continuous and bounded on the positive real axis. Then*

$$\mathrm{L}\left(v\left(\xi+\xi_0\right)u(\xi)\right)(s) = \int_s^\infty \mathrm{d}s_1\, \mathrm{e}^{(s-s_1)\xi_0}\left(L^{-1}v\right)(s_1-s)\,\widehat{u}(s_1) \tag{A.1.1}$$

in the half-plane $\mathrm{Re}(s) > 0$. If the convergence radius of $p(t)$ even $\rho > \frac{1}{\xi_0-\widetilde{\xi_0}} > \frac{1}{\xi_0}$ for some $0 < \widetilde{\xi_0} < \xi_0$, then there exists a constant $C > 0$ such that

$$\left|\mathrm{e}^{-s\xi_0}\left(L^{-1}v\right)(s)\right| \leq C\mathrm{e}^{-\xi_0 \mathrm{Re}\, s + \left(\xi_0-\widetilde{\xi}_0\right)|s|} \tag{A.1.2}$$

$$\left|\frac{\mathrm{d}}{\mathrm{d}s}\left(\mathrm{e}^{-s\xi_0}\left(L^{-1}v\right)(s)\right)\right| \leq C\mathrm{e}^{-\xi_0 \mathrm{Re}\, s + \left(\xi_0-\widetilde{\xi}_0\right)|s|} \tag{A.1.3}$$

for all $s \in \mathbb{C}$.

PROOF. It holds

$$\left(\mathrm{L}^{-1}v\right)(s) = \sum_{m=1}^{\infty} \frac{p_m}{(m-1)!} s^{m-1}.$$

For $m = 1$ Eq. (A.1.1) recovers Lemma A.1.6. We want to show the result for $m > 1$ by induction. Assume that (A.1.1) holds true for $p = t^m$. We show that it holds also for $p = t^{m+1}$. By Lemma A.1.6

$$\begin{aligned}
&\mathrm{L}\left(\left(\frac{1}{\xi_0+\xi}\right)^{m+1} u(\xi)\right)(s) \\
&= \mathrm{L}\left(\frac{1}{\xi_0+\xi}\frac{u(\xi)}{(\xi_0+\xi)^m}\right)(s) \\
&= \int_s^\infty \mathrm{d}s_2\, \mathrm{e}^{(s-s_2)\xi_0}\left(\int_{s_2}^\infty \mathrm{d}s_1\, \mathrm{e}^{(s_2-s_1)\xi_0}\mathrm{L}^{-1}\left(\left(\frac{1}{\xi}\right)^m\right)(s_1-s_2)\,\widehat{u}(s_1)\right) \\
&= \int_s^\infty \mathrm{d}s_2 \int_{s_2}^\infty \mathrm{d}s_1\, \mathrm{e}^{(s-s_1)\xi_0}\frac{1}{(m-1)!}(s_1-s_2)^{m-1}\,\widehat{u}(s_1) \\
&= \int_s^\infty \mathrm{d}s_1 \int_s^{s_1} \mathrm{d}s_2\, \mathrm{e}^{(s-s_1)\xi_0}\frac{1}{(m-1)!}(s_1-s_2)^{m-1}\,\widehat{u}(s_1) \\
&= \int_s^\infty \mathrm{d}s_1\, \mathrm{e}^{(s-s_1)\xi_0}\frac{1}{m!}(s_1-s)^m\,\widehat{u}(s_1).
\end{aligned}$$

So far the statement holds if p is a polynomial. We extend this to the case that p is an infinite series. From the definition of the convergence radius follows the exponential bound

$$\begin{aligned}\left|\left(\mathrm{L}^{-1}v\right)(s)\right| &\leq \sum_{m=0}^{\infty}\frac{|p_{m+1}|}{m!}|s^m| \\ &= \sum_{m=0}^{\infty}\frac{|p_{m+1}|}{m!}\frac{\left(\xi_0-\widetilde{\xi}_0\right)^m}{\left(\xi_0-\widetilde{\xi}_0\right)^m}|s^m| \\ &\leq C\sum_{m=0}^{\infty}\frac{\left(\xi_0-\widetilde{\xi}_0\right)^m}{m!}|s^m| \\ &\leq C\,\mathrm{e}^{\left(\xi_0-\widetilde{\xi}_0\right)|s|}\end{aligned}$$

with $C=\left|\xi_0-\widetilde{\xi}_0\right|\sup_{m\geq 0}|p_{m+1}|\left(\xi_0-\widetilde{\xi}_0\right)^{-(m+1)}<\infty$, because all partial sums above are bounded by the right-hand side given by the exponential factor $\exp\left(\xi_0-\widetilde{\xi}_0\right)|s|$ and because the power series converges uniformely for $|t|<1/\xi_0<1/(\xi_0-\widetilde{\xi}_0)$. The latter fact makes it neccessary that $\lim_{m\to\infty}|p_{m+1}|\left(\xi_0-\widetilde{\xi}_0\right)^{-(m+1)}=0$, hence $C<\infty$. Obviouysly, we can repeat the same conclusion for

$$\left|\left(\mathrm{L}^{-1}v\right)'(s)\right|=\left|\sum_{m=0}^{\infty}\frac{|p_{m+1}|}{(m-1)!}|s^{m-1}|\right|$$

resulting in

$$\left|\left(\mathrm{L}^{-1}v\right)'(s)\right|\leq C\,\mathrm{e}^{\left(\xi_0-\widetilde{\xi}_0\right)|s|}$$

with $C=\left|\xi_0-\widetilde{\xi}_0\right|^2\sup_{m\geq 0}|p_{m+1}|\left(\xi_0-\widetilde{\xi}_0\right)^{-(m+1)}<\infty$

It can be shown by Lebegue's Dominated Convergence Theorem that

$$\begin{aligned}\int_0^{\infty}\mathrm{d}\xi\,\mathrm{e}^{-s\xi}v(\xi+\xi_0)u(\xi) &= \lim_{M\to\infty}\int_0^{\infty}\mathrm{d}\xi\,\mathrm{e}^{-s\xi}\sum_{m=1}^{M}\frac{p_m}{(\xi+\xi_0)^m}u(\xi) \\ &= \lim_{M\to\infty}\int_s^{\infty}\mathrm{d}s_1\,\mathrm{e}^{(s-s_1)\xi_0}\sum_{m=1}^{M}\frac{p_m}{(m-1)!}(s-s_1)^{m-1}\,\widehat{u}(s_1) \\ &= \int_s^{\infty}\mathrm{d}s_1\,\mathrm{e}^{(s-s_1)\xi_0}\left(L^{-1}v\right)(s_1-s)\,\widehat{u}(s_1)\end{aligned}$$

□

LEMMA A.1.9. *Let $u:\mathbb{R}_+\to\mathbb{C}$ be bounded. Further let u have a bounded analytic continuation such that $u(r)$ is holomorphic for all complex r with $\operatorname{Re} r\geq 0$ and $0\leq\arg r<\pi/2$. Let $\alpha\neq 0$ be a complex number with $0\leq\arg\alpha<\pi/2$. Then it holds*

$$L\left(u\left(\alpha\xi\right)\right)(s)=\frac{1}{\alpha}\left(Lu\left(\xi\right)\right)\left(\frac{s}{\alpha}\right)\qquad \textit{for } \operatorname{Re}s>0,\quad \xi\in\mathbb{R}_+.$$

PROOF. By definition, it holds

$$L\left(u\left(\alpha\xi\right)\right)(s)=\int_0^{\infty}\mathrm{e}^{-s\xi}u(\alpha\xi)\,\mathrm{d}\xi,\qquad \operatorname{Re}s>0.$$

This is identical to

$$L\left(u\left(\alpha\xi\right)\right)(s)=\int_0^{\infty}\frac{1}{\alpha}\mathrm{e}^{-\frac{s}{\alpha}\alpha\xi}u(\alpha\xi)\,\mathrm{d}\alpha\xi,\qquad \operatorname{Re}s>0.$$

Setting $\zeta := \alpha\xi$ and applying Cauchy's Theorem, we obtain, with $P_\infty := \lim_{\xi\to\infty} \alpha\xi$,

$$\begin{aligned}\int_0^{P_\infty} \frac{1}{\alpha} \mathrm{e}^{-\frac{s}{\alpha}\zeta} u(\zeta)\, d\zeta &= \int_0^{\infty} \frac{1}{\alpha} \mathrm{e}^{-\frac{s}{\alpha}\xi} u(\xi)\, \mathrm{d}\xi \\ &= \frac{1}{\alpha} L\left(u\left(\xi\right)\right)\left(\frac{s}{\alpha}\right).\end{aligned}$$

□

A.2. Discrete Operational Calculus

This is the discrete analog to Mikusiński's operational calculus **[74]**. Our goal is to define an algebra with infinite sequences as elements such that operations like quotients and square roots of sequences and are well defined. Let U denote the infinite sequence of complex numbers $\{u_0, u_1, \ldots, u_j, \ldots\}$, $|u_j| < \infty$ for all $j \geq 0$. Let $\mathcal{C}$ be the set of all such sequences. In particular, if the u_j are functions $u_j : \mathbb{C} \to \mathbb{C}$ of the complex argument p, we write $U\left(p\right) := \{u_0\left(p\right), u_1\left(p\right), \ldots, u_j\left(p\right), \ldots\}$.

Definition A.2.1. We define the Null element and the unity element of $\mathcal{C}$ by

$$[0] = \{0, 0, \ldots,\} \quad \text{and} \quad [1] = \{1, 0, 0, \ldots,\}.$$

For simplicity we write often 0 instead of [0], and 1 instead of [1], respectively, if no ambiguity occurs. We define the operations addition and multiplication between two elements $U, V \in \mathcal{C}$ by

$$U + V = \{u_0 + v_0,\, u_1 + v_1,\, \ldots\} \quad \text{and} \quad UV = \{g_0, g_1, \ldots\} \text{ with } g_j = \sum_{i=0}^{j} u_{j-i} v_i\,.$$

From these definitions we find for any $U, V \in \mathcal{C}$ that $U + V \in \mathcal{C}$ and $UV \in \mathcal{C}$, and further $UV = VU$, $U(VW) = (UV)W$, and $U(V + W) = UV + UW$. Hence, $\mathcal{C}$ is a commutative ring. It follows by induction that for $U, V \in \mathcal{C}$ the equation $UV = 0$ implies either $U = 0$ or $V = 0$ or both. Hence we can extend the ring $\mathcal{C}$ to the ring of fractions $\mathcal{C}_q = \{U/V \,|\, U, V \in \mathcal{C}, V \neq 0\}$. To fractions U/V and U'/V' are defined to be equal if and only if $UV' = VU'$. The addition and multiplication of two fractions of sequences are defined by

$$\frac{U}{V} + \frac{U'}{V'} = \frac{UV' + VU'}{VV'} \quad \text{and} \quad \frac{U}{V}\frac{U'}{V'} = \frac{UU'}{VV'}\,,$$

respectively. Some special elements of $\mathcal{C}_q$ are labeld with own symbols. We define the multiplication operator by $[a] := \{a,\, 0,\, 0, \ldots\}$, $a \in \mathbb{C}$. The operation $[a]\,U$ realizes a multiplication of each element of U by a, for simplicity we write aU.

Definition A.2.2. The shift-operator s is defined by $s := \{0,\, 1,\, 0,\, 0, \ldots\}$.

By definition, it holds $sU = \{0,\, u_0,\, u_1, \ldots,\}$ and $ss = s^2 = \{0,\, 0,\, 1,\, 0,\, \ldots,\}$, etc. Two sequences $U, V \in \mathcal{C}$ are equal, if their elements are equal, $u_i = v_i$ for all $i \geq 0$. In particular, if we compare sequences elementwise, we use the following notation: $(sU)_{i+1} = (U)_i = u_i$. Even shorter we write $su_{i+1} = u_i$ which has to be understand in the foregoing sense. The following property holds: $V := s^k U$ implies $v_j = 0$ for $j = 0, \ldots, k-1$. Further, we say that $U(p)$ has a complex pole λ if $\lim_{p\to\lambda} |u_j(p)| = \infty$ for some $j \geq 0$.

Definition A.2.3. A linear mapping $\mathrm{A} : \mathcal{C} \to \mathcal{C}$ is defined through infinite matrices $\mathrm{A} = (a_{ij})_{i,j=0,\ldots,\infty}$, $a_{ij} \in \mathbb{C}$ which map U into V, $U \mapsto V = \mathrm{A}U$ by a matrix-vector multiplication $V_i = \sum_{j=0}^{i} a_{ij} U_j$. We denote the set of all of such matrices by $\mathcal{L}$.

In our application equations between sequences appear of the form

$$\left[\begin{pmatrix} \mathrm{A} & & & \\ & \mathrm{A} & & \\ & & \mathrm{A} & \\ & & & \ddots \end{pmatrix} + \begin{pmatrix} \mathrm{B} & & & \\ & \mathrm{B} & & \\ & & \mathrm{B} & \\ & & & \ddots \end{pmatrix} s^k \right] U = V\,, \tag{A.2.1}$$

where A and B are lower triangular, complex $k \times k$ matrices. Each row of this equation corresponds to one propagation step (we call this a macro-step), which itself consists of k intermediate steps, according to the order of the Padé approximation. In view of this separation of matrices, we separate the sequences U and V into subsequences of length k, and motivated by the matrix form of the basic equation we denote this subsequences column vectors, that is we compose $U = \{\mathrm{u}_0^T, \mathrm{u}_1^T, \dots,\}$ and $V = \{\mathrm{v}_0^T, \mathrm{v}_1^T, \dots,\}$ with $\mathrm{u}_j, \mathrm{v}_j \in \mathbb{C}^k$. Thus we can replace (A.2.1) by the shorter form

$$\left(\mathrm{A} + \mathrm{B}s^k\right) \mathrm{u}_j = \mathrm{v}_j,\ j \geq 0\,. \tag{A.2.2}$$

The second important type of matrices which occur are lower triangular block matrices of *convolution* type

$$\begin{pmatrix} \mathrm{B}_0 & & & \\ \mathrm{B}_1 & \mathrm{B}_0 & & \\ \mathrm{B}_2 & \mathrm{B}_1 & \mathrm{B}_0 & \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{A.2.3}$$

where the B_j, $j \geq 0$, are $k \times k$ complex matrices. These matrices are special forms of infinite Töplitz matrices, cf. e. g. **[14]**. In the special case $k = 1$ we can assign to each matrix B of convolution type a sequence $B \in \mathcal{C}$ defined by $B = \{\mathrm{B}_0, \mathrm{B}_1, \mathrm{B}_2, \dots\}$ such that $\mathrm{B}U = BU$ for all $U \in \mathcal{C}$. A typical mapping $U \mapsto V$ within our factorization procedure possesses the form

$$\left[\begin{pmatrix} \mathrm{A} & & & \\ & \mathrm{A} & & \\ & & \mathrm{A} & \\ & & & \ddots \end{pmatrix} + \begin{pmatrix} \mathrm{B}_0 & & & \\ \mathrm{B}_1 & \mathrm{B}_0 & & \\ \mathrm{B}_2 & \mathrm{B}_1 & \mathrm{B}_0 & \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \right] U = V, \tag{A.2.4}$$

where A and B_0 are lower triangular, complex $k \times k$ matrices and B_j, $j \geq 1$ are arbitrary, complex $k \times k$ matrices. By our notation introduced above, the jth row ($j \geq 0$) reads

$$\left(\mathrm{A} + \sum_{i=0}^{j} \mathrm{B}_i s^{ik}\right) \mathrm{u}_j = \mathrm{v}_j,\ j \geq 0\,, \tag{A.2.5}$$

which we further abbreviate by $(\mathrm{A} + \mathrm{B}(s))\,\mathrm{u}_j = \mathrm{v}_j$ with an operator $\mathrm{B}(s) := \sum_{i=0}^{j} \mathrm{B}_i s^{ik}$. We denote equations of type A.2.5 as operator equations, because they are formulated using the shift operator. We say that the inverse of the operator appearing in brackets on the left-hand side of A.2.4 possesses a pole, if $\det(\mathrm{A} - \mathrm{B}_0) = 0$. Due to the equivalence of A.2.4 and the operator equation A.2.5 we say accordingly, that the operator equation A.2.2 possesses a pole if $\det(\mathrm{A} - \mathrm{B}_0) = 0$.

PROPOSITION A.2.4. *The product of two infinite matrices of convolution type is again of convolution type.*

PROOF. We consider the mapping $U \mapsto V$

$$\left[\begin{pmatrix} \mathrm{A}_0 & & & \\ \mathrm{A}_1 & \mathrm{A}_0 & & \\ \mathrm{A}_2 & \mathrm{A}_1 & \mathrm{A}_0 & \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \mathrm{B}_0 & & & \\ \mathrm{B}_1 & \mathrm{B}_0 & & \\ \mathrm{B}_2 & \mathrm{B}_1 & \mathrm{B}_0 & \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}\right] U = V, \tag{A.2.6}$$

with matrices B_j as above and A_j of the same type as B_j. The jth row of (A.2.6) reads

$$\left[\sum_{i=0}^{j} \underbrace{\left(\sum_{l=0}^{i} \mathrm{A}_l \mathrm{B}_{i-l}\right)}_{\mathrm{C}_i} s^{ik}\right] \mathrm{u}_j = \mathrm{v}_j,\ j \geq 0\,, \tag{A.2.7}$$

and the matrices C_i represent the resulting matrix. □

The expression in brackets on the left-hand side of (A.2.7) can be identified with the formal product

$$\left(\sum_{i=0}^{j} \mathrm{A}_i s^{ik}\right)\left(\sum_{i=0}^{j} \mathrm{B}_i s^{ik}\right) = \mathrm{A}_0\mathrm{B}_0 + \sum_{l+m=1} \mathrm{A}_l \mathrm{B}_m s^k + \ldots + \sum_{l+m=j} \mathrm{A}_l \mathrm{B}_m s^{jk},$$
$$j, l, m \geq 0.$$

Thus given an operator expression in polynomial form $\sum_{i=0}^{j} \mathrm{C}_i s^{ki}$ and matrices B_i, $i = 0, \ldots, j$, we can determine the coefficients A_i, $i = 0, \ldots, j$, of a factorization $\left(\sum_{i=0}^{j} \mathrm{A}_i s^{ik}\right)\left(\sum_{i=0}^{j} \mathrm{B}_i s^{ik}\right) \mathrm{u}_j = \sum_{i=0}^{j} \mathrm{C}_i s^{ki} \mathrm{u}_j$ by a comparison of coefficients. In the following we consider matrices of convolution type $\mathrm{B}(s) = \sum_{i=0}^{j} \mathrm{B}_i s^{ik}$, $\mathrm{B}_i \in \mathbb{C}^{k\times k}$ with invertible matrices B_0. In particular, we consider the special choice $\mathrm{B}_0 = \mathrm{I}$. The general case is easily traced back to this case.

LEMMA A.2.5. *Given an operator* $\mathrm{B}(s)$ *of convolution type,*

$$\mathrm{B}(s) = \mathrm{I} + \sum_{i=1}^{j} \mathrm{B}_i s^{ik},\ \mathrm{B}_i \in \mathbb{C}^{k\times k}\,.$$

It holds:

(i) the inverse $\mathrm{B}^{-1}(s)$ *of the operator* $\mathrm{B}(s)$ *exists and* $\mathrm{B}^{-1}(s) \in \mathcal{L}$.

(ii) $\mathrm{B}^{-1}(s)$ *possesses the convolution form (A.2.3). The operator* $\mathrm{C}(s) := \mathrm{B}^{-1}(s)$ *has a representation*

$$\mathrm{C}(s) = \mathrm{I} + \sum_{i=1}^{j} \mathrm{C}_i s^{ik},\ \mathrm{C}_i \in \mathbb{C}^{k\times k}\,, \tag{A.2.8}$$

where the C_i *are recursively determined by* $\mathrm{C}_i = -\sum_{l=0}^{i-1} \mathrm{B}_{i-l}\mathrm{C}_l$, $\mathrm{C}_0 = \mathrm{I}$.

PROOF. (i) Gaussian elimination.

(ii) Proof by induction. Given $U \in \mathcal{C}$, we set $\mathrm{v}_j := \left(\mathrm{I} + \sum_{i=1}^{j} \mathrm{B}_i s^{ik}\right) \mathrm{u}_j$. We want to show that there exists an operator $\mathrm{C}(s)$ with $\mathrm{u}_j = \left(\mathrm{I} + \sum_{i=1}^{j} \mathrm{C}_i s^{ik}\right) \mathrm{v}_j$. Let $j = 0$ and $j = 1$. It holds, by definition of $\mathrm{B}(s)$, $\mathrm{u}_0 = \mathrm{v}_0$ and $\mathrm{B}_1\mathrm{u}_0 + \mathrm{u}_1 = \mathrm{v}_1$, hence

$$\begin{aligned} \mathrm{u}_0 &= \mathrm{v}_0 \\ \mathrm{u}_1 &= -\mathrm{B}_1\mathrm{v}_0 + \mathrm{v}_1\,. \end{aligned}$$

It follows $\mathrm{C}_1 = -\mathrm{B}_1$. Assume that the representation formula (A.2.8) is true for all indices l up to some $j \geq 1$. Hence it holds

$$\begin{aligned} \mathrm{u}_l &= \sum_{i=0}^{l} \mathrm{C}_i \mathrm{v}_{l-i}\,, \quad l = 0, \ldots, j \text{ and } \mathrm{C}_0 := \mathrm{I} \\ \mathrm{u}_j &= \mathrm{v}_j - \sum_{i=1}^{j} \mathrm{B}_i \mathrm{u}_{j-i}\,, \end{aligned} \tag{A.2.9}$$

where the last equation is just the definition of $\mathrm{B}(s)$. It follows, expressing u_l by the appropriate representation (A.2.9),

$$\begin{aligned} \mathrm{u}_j &= \mathrm{v}_j - \sum_{i=1}^{j} \mathrm{B}_i \mathrm{u}_{j-i} \\ &= \mathrm{v}_j - \sum_{i=1}^{j} \sum_{m=0}^{j-i} \mathrm{B}_i \mathrm{C}_m \mathrm{v}_{j-i-m} \qquad \text{(A.2.10)} \\ &= \mathrm{v}_j - \sum_{k=j-1}^{0} \sum_{m=0}^{k} \mathrm{B}_{j-k} \mathrm{C}_m \mathrm{v}_{k-m} \\ &= \mathrm{v}_j - \sum_{k=0}^{j-1} \sum_{m=0}^{k} \mathrm{B}_{j-k} \mathrm{C}_m \mathrm{v}_{k-m} \qquad \text{(A.2.11)} \\ &= \mathrm{v}_j - \sum_{k=0}^{j-1} \sum_{l=k}^{0} \mathrm{B}_{j-k} \mathrm{C}_{k-l} \mathrm{v}_l \\ &= \mathrm{v}_j - \sum_{k=0}^{j-1} \sum_{l=0}^{k} \mathrm{B}_{j-k} \mathrm{C}_{k-l} \mathrm{v}_l \\ &= \mathrm{v}_j - \sum_{l=0}^{j-1} \underbrace{\left(\sum_{k=l}^{j-1} \mathrm{B}_{j-k} \mathrm{C}_{k-l} \right)}_{-\mathrm{C}_{j-l}} \mathrm{v}_l \qquad \text{(A.2.12)} \end{aligned}$$

In (A.2.10) we replaced $j - i$ by k and eliminated i, in (A.2.11) we replaced $k - m$ by l and eliminated m. In (A.2.12) we changed the order of summation. It follows $\mathrm{C}_i = -\sum_{l=0}^{i-1} \mathrm{B}_{i-l} \mathrm{C}_l$ for $i \leq j$. This gives in recursive manner $\mathrm{C}_1 = -\mathrm{B}_1$, $\mathrm{C}_2 = -\mathrm{B}_2 - \mathrm{B}_1 \mathrm{C}_1$, etc. Note, that we used the induction assumption (A.2.9) only up to the index $l = j - 1$. By assumption, this representation holds true for $l = j$. In what follows, we show that (A.2.9) remains true for $l = j + 1$. By definition

$$\mathrm{u}_{j+1} = \mathrm{v}_{j+1} - \sum_{i=1}^{j+1} \mathrm{B}_i \mathrm{u}_{j+1-i}\,.$$

Hence we can repeat the procedure setting $j := j + 1$ which yields

$$\mathrm{u}_{j+1} = \mathrm{v}_{j+1} - \sum_{l=0}^{j} \underbrace{\left(\sum_{k=l}^{j} \mathrm{B}_{j+1-k} \mathrm{C}_{k-l} \right)}_{-\mathrm{C}_{j+1-l}} \mathrm{v}_l$$

It follows $\mathrm{C}_i = -\sum_{l=0}^{i-1} \mathrm{B}_{i-l} \mathrm{C}_l$ at step $j + 1$ for $i \leq j + 1$. That is exact the same recursion as for the jth step, hence the the matrices C_i, $0 \leq i \leq j$ corresponding to steps j and $j + 1$, respectively, are the same. □

EXAMPLE A.2.6. Let the operator equation

$$\left(\sum_{i=0}^{j} a_i s^i\right)\left(\sum_{i=0}^{j} a_i s^i\right) u_j = (1+s)u_j$$

be given. We to compute coefficients a_i, $i \geq 0$, such that this equation holds for all given u_j, $j \geq 0$. The operator on the right-hand side is of type (A.2.2) with $k = 1$, A $= 1$, B $= 1$. A comparison of coefficients supplies $a_0^2 = 1$, $2a_0a_1 = 1$, $2a_0a_2 + 2a_1a_1 = 0, \ldots$. Once a_0 is determined, the whole sequence follows uniquely. The operator equation should hold for any $U \in \mathcal{C}$. Hence it is convenient to write

$$\left(\sum_{i=0}^{j} a_i s^i\right)^2 = 1 + s$$

and to define the square root of $1 + s$ by

$$\sqrt{1+s} = \left(\sum_{i=0}^{j} a_i s^i\right),$$

where the coefficients a_i are determined by the above procedure. Observe, that by construction the coefficients a_i are in fact the Taylor coefficients of $\pm\sqrt{1+s}$, s considered as complex variable and the square root defined with the usual branch cut. Observe further, that the formal expression $\sqrt{1+s}u_j$ has to be read as $\sum_{i=0}^{j} a_i s^i u_j = \sum_{i=0}^{j} a_i u_{j-i}$. Hence the Taylor coefficients in the formal expression have set to zero for $i > j$.

Bibliography

[1] B. Alpert, L. Greengard, and T Hagstrom. Rapid evolution of non-reflecting boundary kernels for time-domain wave paropagation. *SIAM J. Mumer. Anal.*, 37:1138–1164, 2000.

[2] S. Amini, P. J. Harris, and D. T. Wilton. Coupled Boundary and Finite Element Methods for the Solution of the Dynamic Fluid-Structure Interaction Problem. In C. A. Brebbia and S. A. Orszag, editors, *Lecture Notes in Engineering*, volume 77. Springer Verlag, New York, Berlin, 1992.

[3] S. Amini and S. M. Kirkup. Solution of Helmholtz equation in the exterior domain by elementary boundary integral methods. *J. Comput. Phys.*, 118:208–221, 1995.

[4] A. Arnold. Numerically absorbing boundary conditions for quantum evolution equations. In *Proceedings of the International Workshop on Computational Electronics, Tempe, USA,1995.* VLSI-Design 63-5, 1995.

[5] Anton Arnold and Matthias Erhardt. Discrete transparent boundary conditions for wide angle parabolic equations in underwater acoustics. *J. Comput. Phys.*, 145(2):611–638, 1998.

[6] A. Bamberger, B. Engquist, L. Halpern, and P. Joly. Higher order paraxial wave equation approximations in heterogeneuos media. *SIAM J. Appl. Math.*, 48(1):129–154, 1988.

[7] A. Bamberger, B. Engquist, L. Halpern, and P. Joly. Parabolic wave equation approximations in heterogeneuos media. *SIAM J. Appl. Math.*, 48(1):99–128, 1988.

[8] V. A. Baskakov and A. V. Popov. Implementation of transparent boundaries for numerical solution of the schrödinger equation. *Wave Motion*, 14:123–128, 1991.

[9] A. Bayliss, M. Gunzburger, and E. Turkel. Boundary conditions for the numerical solution of elliptic equations in exterior domains. *SIAM J. Appl. Math.*, 42:430–451, 1982.

[10] A. Bayliss and E. Turkel. Radiation boundary conditions for wave-like equations. *Communic. on Pure and Applied Math.*, 33:707–725, 1980.

[11] J.-P. Bérenger. A Perfectly Matched Layer for the Absorbtion of Electromagnetic Waves. *J. Comput. Phys.*, 114:185–200, 1994.

[12] J.-P. Bérenger. Three Dimensional Perfectly Matched Layer for the Absorbtion of Electromagnetic Waves. *J. Comput. Phys*, 127:363–379, 1996.

[13] Max Born and Emil Wolf. *Principle of Optics*. Pergamon Press Oxford, New York, Seoul, Tokyo, 6 edition, 1980.

[14] A. Bötcher and S. M Grudsky. *Töplitz Matrices, Asymptotic Linear Algebra, and Functional Analysis*. Birkhäuser, Basel, Boston, Berlin, 2000.

[15] P. G. Carrier, L. Greengard, and V. Rokhlin. A fast adaptive multipole algorithm for particle simulations. *SIAM J. Sci. Stat. Comput.*, 9(4):159–184, 1988.

[16] L. Collatz. *Differentialgleichungen*. Teubner, 7 edition, 1990.

[17] Francis Collino and Peter Monk. The perfectly matched layer in curvilinear coordinates. *SIAM J. Sci. Comput.*, 19(6):2061–2090, 1998.

[18] Franis Collino. Perfectly Matched Absorbing Layers for the Paraxial Equations. *J. Comp. Phys.*, 131(1):164–180, 1997.

[19] M.D. Collins. Higher-order Padé approximations for accurate and stable elastic parabolic equations with application to interface wave propagation. *J. Acoust. Soc. Am.*, 89:1050–1057, 1991.

[20] M.D. Collins. A split-step Padé solution for the parabolic equation method. *J. Acoust. Soc. Am.*, 93:1736–1742, 1993.

[21] D. Colton and R. Kress. *Inverse Acoustic and Electromagnetic Scattering Theory*. Springer-Verlag Berlin, Heidelberg, New York, 1992.

[22] L. Demkovicz and K. Gerdes. Convergence of the infinite element methods for the Helmholtz equation. *Numerische Mathematik*, 79(1):11–42, 1998.

[23] John W. Dettman. *Mathematical methods in physics and engineering*. McGraw-Hill, 2 edition, 1969.

[24] John W. Dettman. *Applied complex variables*. Dover Publications, Inc., New York, 1984.

[25] Peter Deuflhard and Folkmar Bornemann. *Numerische Mathematik II. Integration gewöhnlicher Differentialgleichungen*. Walter de Gruyter, Berlin, New York, 1994.

[26] Doetsch. *Handbuch der Laplace-Transformation*. Birkhäuser Verlag, Basel, Stuttgart, 1955.
[27] H. W. Engl. *Integralgleichungen*. Springer, Vienna/New York, 1997.
[28] B. Engquist and A. Majda. Absorbing Boundary Conditions for the Numerical Simulation of Waves. *Mathematics of Computation*, 31(139):629–651, July 1977.
[29] Bjorn Engquist and Andrew Majda. Radiation Boundary Condition for Acoustic and Elastic Wave Propagation. *Communications on Pure and Applied Mathematics*, XXXII:313–357, 1979.
[30] Oliver G. Ernst. *Fast Numerical Solution of Exterior Helmholtz Problems with Radiation Condition By Imbedding*. PhD thesis, Department of Scientific Computing and Computational Mathematics of Stanford University, 1994.
[31] Richard P. Feynmann, Robert B. Leighton, and Matthew Sands. *The Feynmann lectures on physics*. Addisson Wesley Publishing Company, Reading Massachusetts, 1963.
[32] Tilmann Friese, Frank Schmidt, and David Yevick. Transparent boundary conditions for a wide-angle approximation of the one-way Helmholtz equation. *J. Comput. Phys.*, 165(2):645–659, 2000.
[33] Roland Geng. Berechnung der Eigenlösung des Schrödinger-Operators auf unbeschränkten Gebieten in 1d. Master's thesis, Freie Universität Berlin, Fachbereich Mathematik und Informatik, submission: 2001. in preparation.
[34] Klaus Giebermann. *Schnelle Summationsverfahren zur numerischen Lösung von Integralgeichungen für Streuprobleme im R^3*. PhD thesis, Universität Karlsruhe, Fakultät für Mathematik, 1997.
[35] D. Givoli. Nonreflecting boundary conditions. *J. Comput. Phys.*, 94:1–29, 1991.
[36] D. Givoli. A spatially exact non-reflecting boundary condition for time dependent problems. *Comput. Methods Appl. Mech. Engrg.*, 95:97–113, 1992.
[37] D. Givoli and J. B. Keller. Nonreflecting boundary conditions for elastic waves. *Wave Motion*, 12:261–279, 1990.
[38] Dan Givoli and Joseph B. Keller. A finite element method for large domains. *Computer Methods in Applied Mechanics and Engineering*, 76(6):41–66, 1989.
[39] Gene H. Golub and Charles F. Loan. *Matrix computations*. The John Hopkins University Press, third edition 1996 edition, 1983.
[40] L. Greengard. *The rapid evaluation of potential fields in particle systems*. MIT Press, Cambridge, Massachusetts, 1987.
[41] L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. *J. Comp. Phys.*, 73(2):325–348, 1987.
[42] M. J. Grote and J. B. Keller. Exact nonreflecting boundary conditions for time-dependent wave equation. *SIAM J. of Applied Mathematics*, 55(2):280–297, April 1995.
[43] M. J. Grote and J. B. Keller. On nonreflecting boundary conditions. *J. Comput. Phys*, 122:231–243, 1995.
[44] M. J. Grote and J. B. Keller. Nonreflecting boundary conditions for time-dependent scattering. *J. Comput. Phys.*, 127:52–65, 1996.
[45] M. J. Grote and J. B. Keller. Nonreflecting boundary conditions for Maxwell's equations. *J. Comput. Phys.*, 1997. submitted.
[46] W. Hackbusch, C. Lage, and S.A. Sauter. On the Efficient Realization of Sparse Matrix Techniques for Integral Equations with Focus on Panel Clustering, Cubature and Software Design Aspects. In Wolfgang E. Wendland, editor, *Boundary Element Topics: proceedings of the Final Conference of the Priority Research Programme Boundary Element Methods 1989-1995 of the German Research Foundation*, pages 51–75. Springer Verlag, 1997.
[47] W. Hackbusch and Z. P. Nowak. On the fast matrix multiplication in the boundary element method by panel clustering. *Numerische Mathematik*, 54:463–491, 1989.
[48] G.R. Hadley. Multistep method for wide-angle beam propagation. *Optics Letters*, 17:1743–1745, 1992.
[49] T. M. Hagstrom. Asymptotic expansions and boundary conditions for time-dependent problems. *SIAM J. Num. Anal.*, 23:948–958, 1986.
[50] Thomas Hagstrom. Radiation boundary condition for the numerical simulation of waves. *Acta Numerica*, 8:47–106, 1999.
[51] E. Hairer, C. Lubich, and M. Schlichte. Fast numerical solution of weakly singular Volterra integral equations. *J. Comput. Appl. Math.*, 23(1):87–98, 1988.
[52] L. Halpern. Artificial boundary conditions for incompletely parabolic perturbations of hyperbolic systems. *SIAM J. Math. Anal.*, 22:1256–1283, 1991.
[53] L. Halpern and L. N. Trefethen. Wide-angle one-way wave equations. *J. Acoust. Soc. Amer.*, 84(4):1397–1404, 1988.
[54] Laurence Halpern and Jeffrey Rauch. Error analysis for absorbing boundary conditions. *Numerische Mathematik*, 51:459–467, 1987.

[55] Isaac Harari and Thomas. J. R. Hughes. Analysis of continuous formulations underlying the computation of time-harmonic acoustics in exterior domains. *Comput. Methods Appl. Mech. Engrg.*, 97:103–124, 1997.

[56] Erikki Heikkola, Yuri A. Kuznetsov, Pekka Neittaanmäki, and Jari Toivanen. Fictitious Domain Methods for the Numerical Solution of Two-Dimensional Scattering Problems. *Journal of Comput. Phys*, 145(1):89–109, 1998.

[57] Thorsten Hohage, Frank Schmidt, and Lin Zschiedrich. Solving time-harmonic scattering problems based on the pole condition:Convergence of the PML method. Technical report, Konrad-Zuse-Zentrum (ZIB), 2001. ZIB-Report 23-01, in preparation.

[58] Thorsten Hohage, Frank Schmidt, and Lin Zschiedrich. Solving time-harmonic scattering problems based on the pole condition:Theory. Preprint 01-01, Konrad-Zuse-Zentrum (ZIB), 2001.

[59] W.P. Huang and C.L. Xu. A wide-angle vector beam propagation method. *Photon. Technol. Lett.*, 4:1118–1119, 1992.

[60] F. Ihlenburg. *Finite Element Analysis of Acoustic Scattering*, volume 132 of *Applied Mathematical Sciences*. Springer-Verlag New York Berlin Heidelberg, 1998.

[61] Frank Ihlenburg. On fundamental aspects of exterior approximations with infinite elements. *Journal of Computational Acoustics*, 1999. manuscript, submitted 1999.

[62] Claes Johnson. *Numerical solution of partial differential equations by the finite element method*. Cambridge University Press Cambridge, New York, New Rochelle Melbourne Sydney, 1987.

[63] S. N. Karp. A convergent "far-field" expansion for two dimensional radiation functions. *Comm. Pure Appl. MAth.*, 14:427–434, 1961.

[64] J. B. Keller and D. Givoli. Exact non-reflecting boundary conditions. *J. Comput. Phys.*, 82:172–192, 1989.

[65] R. Kosloff and D. Kosloff. Absorbing boundaries for wave propagation problems. *J. Comput. Phys.*, 63:363–376, 1986.

[66] R. Kress. *Linear Integral Equations*. Springer-Verlag Berlin, Heidelberg, New York, 1989.

[67] M. Lassas and E. Somersalo. On the existence and convergence of the solution of pml equations. *Computing*, 60:229–241, 1998.

[68] R. Leis. *Initial Boundary Value Problems in Mathematical Physics*. Teubner, 1986.

[69] Ch. Lubich. Convolution quadrature and discretized operational calculus i, ii. *Numer. Math.*, 52:129–145 and 413–425, 1988.

[70] Ch. Lubich and A. Schädle. Fast convolution for non-reflecting boundary conditions. Preprint, Mathematisches Institut, Universität Tübingen, April 2001.

[71] Reinhard März. *Integrated optics: design and modeling*. Artec House, Inc., 1995.

[72] B. Mayfield. *Nonlocal Boundary Conditions for the Schrödinger Equation*. PhD thesis, University of Rhode Island, Providence, RI, 1989.

[73] Erhard Meister. Modern Wiener-Hopf Methods in Diffraction Theory. In D. Bainov and V. Covachev, editors, *5th Int. Coll. on Differential Equations*, pages 245–256, 1995.

[74] J. Mikusiński. *Operational Calculus*. Pergamon Press, London, New York, 1959.

[75] P. Petropoulus. Reflectionless sponge layers as absorbing boundary conditions for the numerical solution of Maxwell equations in rectangular, cylindrical, and spherical coordinates. *SIAM J. Appl. Math.*, 60:1037–1058, 2000.

[76] V. Rokhlin. Rapid solution of integral equations of classical potential theory. *J. Comput. Phys.*, 60(2):187–207, 1985.

[77] V. Rokhlin. Diagonal forms of translation operators for the Helmholtz equation in three dimensions. *Appl. Comput. Harmon. Anal.*, 1(1):82–93, 1993.

[78] A. Ruhe. An algorithm for numerical determination of the structure of a general matrix. *BIT*, 10:196–216, 1970.

[79] S. Sauter. *Über die effiziente Verwendung des Galerkinverfahrens zur Lösung Fredholmscher Integralgleichungen*. PhD thesis, Universität Kiel, 1992.

[80] F. Schmidt. Construction of discrete transparent boundary conditions for Schrödinger-type equations. *Surv. Math. Ind.*, 9:87–100, 1999.

[81] F. Schmidt. Discrete nonreflecting boundary conditions for the Helmholtz equation. In A. Bermúdez, D. Gómez, C. Hazard, P. Joly, and J. E. Roberts, editors, *Fifth International Conference on Mathematical and Numerical Aspects of Wave Propagation*, pages 921–925, Waves 2000, Santiago de Compostella, Spain, 2000.

[82] F. Schmidt and P. Deuflhard. Discrete transparent boundary conditions for the numerical solution of Fresnel's equation. *Computers Math. Applic.*, 29:53–76, 1995.

[83] F. Schmidt, T. Friese, and D. Yevick. Transparent boundary conditions for split-step Padé approximations of the one-way Helmholtz equation. Submitted to J. Comp. Phys., Preprint SC 99-46, ZIB, 1999.

[84] F. Schmidt and R. März. On the reference wave vector of paraxial Helmholtz equations. *IEEE Journal of Lightwave Technology*, 14:2395–2400, 1996.

[85] F. Schmidt and D. Yevick. Discrete transparent boundary conditions for Schrödinger-type equations. *J. Comput. Phys.*, 134:96–107, 1997.

[86] Frank Schmidt. An Alternate Derivation of the Exact DtN-Map on a Circle. Preprint SC 98-32, Konrad-Zuse-Zentrum Berlin, December 1998.

[87] Michal E. Taylor. *Pseudodifferential Operators*. Princeton University Press, Princeton, New Jersey, 1981.

[88] Michal E. Taylor. *Partial Differential Equations*. Springer Verlag, New York, 1996.

[89] L. N. Trefethen and L. Halpern. Well-posedness of one-way wave equations and absorbing boundary conditions. *J. Acoust. Soc. Amer.*, 47:421–435, 1986.

[90] Calvin H. Wilcox. A generalization of theorems of Rellich and Atkinson. *Proc. Amer. Math. Soc.*, pages 271–276, 1956.

[91] Baolin Yang and Peter G. Petropoulus. Plane-wave analysis and comparison of split-field, bi-axial, and uniaxial PML methods as abcs for pseudospectral electromagnetic wave simulations in curvilinear coordinates. *J. Comp. Phys.*, 146(2):747–774, 1998.

[92] D. Yevick. Optimal absorbing boundary conditions. *J. Opt. Soc. Am. A*, 12:107–110, 1995.

[93] D. Yevick, J. Yu, W. Bardyszewski, and M. Glasner. Stability issues in vector electric field propagation. *Photon. Technol. Lett*, 7:656–658, 1995.

[94] D. Yevick, J. Yu, and F. Schmidt. Analytic studies of absorbing and impedance-matched boundary layers. *IEEE Photonics Technology Letters*, 9:73–75, 1997.

[95] E. Zeidler. *Applied Functional Analysis: Applications to Mathematical Physics*, volume 108 of *Applied Mathmatical Sciences*. Springer-Verlag New York Berlin Heidelberg, 1997.

[96] E. Zeidler. *Applied Functional Analysis: Main Principles and Th eir Applications*, volume 109 of *Applied Mathmatical Sciences*. Springer-Verlag New York Berlin Heidelberg, 1997.

[97] Lin Zschiedrich. *Transparent boundary conditions for time-harmonic scattering problems and time-dependent Schrödinger equations*. PhD thesis, Fachbereich Mathematik und Informatik, FU Berlin, submission: 2002. in preparation.

Index