



TECHNISCHE
UNIVERSITÄT
WIEN

DIPLOMARBEIT

Resonanzprobleme in Wellenleitern mit Hilfe von optimierten infiniten Elementen

ausgeführt am

Institut für
Analysis und Scientific Computing
TU Wien

unter der Anleitung von

Associate Prof. Dipl.-Math. Dr.rer.nat. Lothar Nannen

durch

Stefan Nuding

Matrikelnummer: 11705366

Wien, am 17. Mai 2023

Kurzfassung

Diese Arbeit beschäftigt sich mit der Frage, wie man Finite Elemente Methoden auf unbeschränkte Gebiete erweitern kann. Dazu wird hauptsächlich die Helmholtz Gleichung verwendet, mit deren Hilfe die theoretischen Aspekte präsentiert und vertieft werden. Ganz allgemein betrachtet, muss man im Falle eines unbeschränkten Gebietes eine zusätzliche Bedingung an Lösungen stellen, welche im Wesentlichen das Verhalten im „Unendlichen“ beschreibt. Dies geschieht in der vorliegenden Arbeit mit Hilfe des sogenannten Dirichlet-to-Neumann Operators (kurz DtN), welcher nicht nur in der Theorie eindeutige Lösungen garantiert, sondern ebenfalls die Betrachtung eines beschränkten Gebietes nach Setzen eines transparenten Randes ermöglicht. Aus Sicht der Numerik ist somit eine Diskretisierung jenes Operators notwendig. Dazu präsentieren wir die Learned Infinite Elemente Methode, die daraus besteht, die Finite Elemente Matrix über externe Freiheitsgrade so zu erweitern, dass deren Schur Komplement den exakten Operator „bestmöglich“ approximiert. Dabei werden die zusätzlichen Freiheitsgrade über ein gewisses Optimierungsproblem bestimmt. Essentiell für die Verwendung der oben genannten Methode ist die Voraussetzung, dass der exakte DtN Operator als unendliche Reihe mittels einer bestimmten Basis dargestellt werden kann. In den numerischen Experimenten wurde durch Lösen der Helmholtz Gleichung auf Wellenleitern mit Streukörpern nachgewiesen, dass die Learned Infinite Elemente äußerst effizient einsetzbar ist und insbesondere schon mit vergleichsweise wenigen zusätzlichen Freiheitsgraden hohe Konvergenzraten aufweist. In Kombination mit der Integralmethode wurde überdies die bislang nur für Streuprobleme verwendete Diskretisierung des DtN Operators ebenfalls zur Berechnung von Eigenwerten auf unbeschränkten Gebieten erweitert. In diversen Experimenten konnte demnach die Konvergenz anhand der Berechnung von Eigenwerten des negativen Laplace Operators auf unbeschränkten Wellenleitern bestätigt werden.

Abstract

This paper covers the question, how Finite Element Methods can be used when trying to solve partial differential equations on unbounded domains. Throughout this thesis, the Helmholtz equation will be used to present techniques and to illustrate theoretical aspects. In general, if one is interested in problems on unbounded domains, one needs to demand an additional condition to solutions, that vaguely speaking describes the behavior „nearby infinity“. This can be done with the so-called Dirichlet-to-Neumann (DtN) operator, which not only guarantees uniqueness of solutions theoretically, but also enables re-formulating of the problem to a bounded domain with the help of a transparent boundary. From the numerical point of view, one therefore needs to additionally discretize this operator to gather approximate solutions. One of the main topics will therefore be the presentation of the Learned Infinite Elements Method, which consists of extending the Finite Element matrix via so called external degrees of freedoms such that the respective Schur complement approximates the DtN operator in the „best possible“ way. These additional entries will be obtained through a certain optimization problem. An essential assumption to apply the above method is that the exact DtN operator can be described as a series with the help of a certain basis. In various numerical experiments it was shown by seeking solution of the Helmholtz equation in waveguides, that the Learned Infinite Elements Method resulted in high convergence rates even for relatively few exterior degrees of freedom. In combination with the integral method we could in addition use the previously described discretization of the DtN operator to calculate eigenvalues on unbounded domains. In several numerical experiments we could even verify convergence by calculating eigenvalues of the negative Laplace operator in unbounded waveguides.

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Diplomarbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe.

Wien, am 17. Mai 2023



Stefan Nuding

Inhaltsverzeichnis

1	Einleitung	1
2	Grundlegende Resultate	2
3	Helmholtz Probleme	7
3.1	Helmholtz Gleichung auf beschränkten Gebieten	8
3.2	Helmholtz Gleichung auf unbeschränkten Gebieten	10
3.2.1	Sommerfeld Bedingung	10
3.2.2	Dirichlet-to-Neumann Operator	13
3.3	Lösbarkeit und Eindeutigkeit auf unbeschränkten Gebieten	15
3.3.1	Eindeutige Lösbarkeit auf dem Ganzraum	18
3.3.2	Eindeutige Lösbarkeit in Wellenleitern	24
3.4	Konvergenz von diskreten Methoden	31
4	Diskretisierung des DtN Operators	37
4.1	Learned Infinite Elemente Verfahren	37
4.2	Numerische Resultate	43
4.2.1	Einführendes Beispiel	43
4.2.2	Inhomogenität in der Wellenzahl	47
4.2.3	Wellenleiter mit Streukörpern	55
5	Anwendung auf Eigenwertprobleme	59
5.1	Integralmethode	59
5.2	Einführendes numerisches Beispiel	63
5.3	Kombination mit Learned Infinite Elemente Verfahren	69
5.3.1	Aufsetzen der Problemstellung	73
5.3.2	Einzelne Eigenwerte innerhalb einer Kontur	75
5.3.3	Mehrere Eigenwerte innerhalb einer Kontur	82
5.3.4	Weitere Experimente	89
	Literatur	92

1 Einleitung

In der Numerik ist die Finite Elemente Methode ganz allgemein eine der wichtigsten Methoden, um partielle Differentialgleichungen zu lösen. Dabei ist vor allem die Beschränktheit der Geometrie und die Vorgabe von Randdaten essentieller Bestandteil, um Lösungen zu erhalten. Betrachtet man hingegen Streu- oder Resonanzprobleme, die aus physikalischer Sicht Wellenphänomene beschreiben, so sind diese Voraussetzungen oft nicht erfüllt. Man denke dazu zum Beispiel an Schallwellen, die sich in einem möglicherweise sehr großen Gebiet ausbreiten. Somit steht man gezwungenermaßen vor der Frage, wie man in einem solchen Setting trotzdem die bekannte Finite Elemente Methode anwenden kann.

In dieser Arbeit werden wir versuchen diese Frage durch Lösen der *Helmholtz Gleichung* zu beantworten. Dazu werden wir in Kapitel drei zunächst deren physikalischen Ursprung herleiten und daran anschließend klären, welche Eigenschaften man im Falle eines unbeschränkten Gebietes zusätzlich fordern muss, um überhaupt Eindeutigkeit von Lösungen erwarten zu können. Um schlussendlich eine schwache Formulierung auf einem beschränkten Gebiet zu erhalten, führen wir den sogenannten *Dirichlet-to-Neumann* Operator ein, welcher zunächst nur implizit definiert ist. Anhand zweier konkreter Beispiele werden wir auch eine explizite Darstellung herleiten, mit der es uns gelingen wird, Eindeutigkeit von Lösungen bei bestimmten Problemen zu zeigen.

Aus theoretischer Sicht mag der DtN Operator zwar nützlich erscheinen, doch für eine praktische Implementierung wirft dieser nur noch mehr Fragen auf. Daher werden wir in Kapitel vier die in [HLP21] vorgestellte *Learned Infinite Elemente Methode* präsentieren, die unter gewissen Voraussetzungen eine äußerst effiziente Diskretisierung jenes Operators darstellt. Diese Methode werden wir anschließend auf diverse Problemstellungen anwenden und analysieren.

Im letzten Abschnitt werden wir zeigen, dass sich die Learned Infinite Elemente Methode ebenfalls zur Berechnung von Eigenwerten auf unbeschränkten Gebieten anwenden lässt. Dazu stellen wir zunächst die aus [Bey12] stammende Integralmethode vor, die dann in Kombination mit der oben genannten Diskretisierung des DtN Operators zur Berechnung von Eigenwerten des negativen Laplace Operators in Wellenleitern verwendet wird.

2 Grundlegende Resultate

Wir starten in diesem Abschnitt damit, die wichtigsten grundlegenden Resultate aus der (Funktional-) Analysis und der Numerischen Mathematik anzuführen, die in dieser Arbeit verwendet werden. Die meisten davon führen wir ohne Beweis an und verweisen stattdessen lediglich auf weiterführende Literatur. Als Erstes definieren wir die sogenannten *Sobolevräume*, welche wohl als Fundament der Theorie schwacher Lösungen gelten.

Definition 2.1. Sei $k \in \mathbb{N}$. Unter dem *Sobolevraum* $H^k(\Omega)$ bezeichnen wir die Menge aller Funktionen $u \in L^2(\Omega)$, so dass

$$D^\alpha u \in L^2(\Omega) \quad \text{für alle } |\alpha| \leq k,$$

wobei wie gewohnt $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ ein Multiindex ist und

$$D^\alpha u := \frac{\partial^{|\alpha|} u}{\partial \alpha_1 \dots \partial \alpha_n}$$

die dazugehörige schwache Ableitung bezeichnet. Versehen mit der folgenden Norm und Skalarprodukt

$$\|u\|_{H^k}^2 := \sum_{|\alpha| \leq k} \|D^\alpha u\|_{L^2(\Omega)}^2, \quad (u, v)_{H^k} := \sum_{|\alpha| \leq k} (D^\alpha u, D^\alpha v)_{L^2(\Omega)},$$

ist der oben definierte Sobolevraum $H^k(\Omega)$ ebenfalls ein Hilbertraum.

Die obige Definition macht zunächst nur für $k \in \mathbb{N}$ Sinn. Tatsächlich kann man Sobolevräume auch für beliebige $k \in \mathbb{R}$ definieren, siehe zum Beispiel [Hac92, Kapitel 6.2.5]. In dieser Arbeit werden wir vor allem die Wahl $k = 1/2$ benötigen. Dafür geben wir die nachfolgende Definition an:

Definition 2.2. Sei $\Omega \subset \mathbb{R}^n$ mit $\partial\Omega \in C^{0,1}$ (Lipschitz stetig) und bezeichne $\gamma : H^1(\Omega) \rightarrow L^2(\partial\Omega)$ den linearen und stetigen (sogar eindeutigen) Spuroperator. Dann definieren wir

$$H^{1/2}(\partial\Omega) := \gamma(H^1(\Omega)) = \{v \in L^2(\partial\Omega) \mid \exists u \in H^1(\Omega) : \gamma(u) = v\}.$$

Versehen mit der Norm

$$\|v\|_{H^{1/2}(\partial\Omega)} := \inf \{ \|u\|_{H^1(\Omega)} \mid u \in H^1(\Omega) \wedge \gamma(u) = v \}$$

ist dieser sogar ein Banachraum.

Für eine genaue Definition des Spurooperators siehe [Jün20, Theorem 1.5]. Wir werden weiters Gebrauch des folgenden Begriffs machen:

Definition 2.3 (Gelfand Tripel). Seien H ein separabler Hilbertraum und V ein separabler, reflexiver Banachraum mit stetiger und dichter Einbettung $V \hookrightarrow H$. Dann bezeichnen wir mit $V \hookrightarrow H \hookrightarrow V'$ das sogenannte *Gelfand Tripel*.

Man kann zeigen, dass mit $V := H^1(\Omega)$ und $H := L^2(\Omega)$ die obigen Voraussetzungen erfüllt sind, womit wir typischerweise das Gelfand Tripel $H^1(\Omega) \hookrightarrow L^2(\Omega) \hookrightarrow (H^1(\Omega))'$ erhalten. Wir werden auf diesen Begriff zu einem späteren Zeitpunkt zu sprechen kommen. Um allgemein die Existenz von schwachen Lösungen zu zeigen, wiederholen wir zunächst die folgenden Begriffe:

Definition 2.4. Sei V ein Banachraum mit dazugehöriger Norm $\|\cdot\|_V$, dann bezeichnet $a : V \times V \rightarrow \mathbb{C}$ eine (symmetrische) *Sesquilinearform*, falls folgende Bedingungen für beliebige $\alpha \in \mathbb{C}$ und $u, v, w \in V$ erfüllt sind:

- linear im ersten Argument: $a(\alpha u + v, w) = \alpha a(u, w) + a(v, w)$
- konjugiert linear im zweiten Argument: $a(u, \alpha v + w) = \bar{\alpha} a(u, v) + a(u, w)$
- (symmetrisch): $a(u, v) = \overline{a(v, u)}$

Ist überdies von einer *stetigen Sesquilinearform* die Rede, dann erfüllt $a(\cdot, \cdot)$ zusätzlich

$$|a(u, v)| \leq A \|u\|_V \|v\|_V$$

für eine positive Konstante $A > 0$. Zuletzt bezeichnen wir eine Sesquilinearform als *koerziv*, falls eine Konstante $C > 0$ existiert, die nachfolgende Ungleichung für beliebiges $u \in V$ erfüllt ist:

$$|a(u, u)| \geq C \|u\|_V^2.$$

Für stetige, koerzive Sesquilinearformen erweist sich der nachfolgende Satz in Hinblick auf die Existenztheorie von partiellen Differentialgleichungen als unabdingbar, weshalb wir diesen ebenfalls wiederholen:

Satz 2.5 (Lemma von Lax Milgram). Sei H ein Hilbertraum und $a : H \times H \rightarrow \mathbb{R}$ eine stetige und koerzive Bilinearform, sowie F ein stetiges lineares Funktional $F : H \rightarrow \mathbb{C}$. Dann existiert ein eindeutiges $u \in H$, so dass

$$a(u, v) = F(v) \quad \text{für alle } v \in H.$$

Weiters gilt für die Lösung u die folgende Abschätzung

$$\|u\|_H \leq \alpha^{-1} \|F\|_{H'},$$

wobei $\alpha > 0$ die Koerzivitäts-Konstante bezeichnet.

Ein Beweis befindet sich zum Beispiel in [Jün19, Satz 5.18]. Vor allem in Kapitel 5.1 werden wir einige Sätze und Resultate aus der Komplexen Analysis verwenden. Um etwaige Missverständnisse zu vermeiden, führen wir deshalb auch noch einige Grundbegriffe und Notationen an. Diese werden hauptsächlich aus [Jän93] entnommen. Zunächst definieren wir den folgenden Begriff:

Definition 2.6. Sei $\gamma : [0, 1] \rightarrow U$ eine stetig differenzierbare Kurve, wobei $U \subset \mathbb{C}$ gilt. Weiters sei $f : U \rightarrow \mathbb{C}$ eine stetige Funktion, dann bezeichnen wir mit

$$\int_{\gamma} f(z) dz := \int_0^1 f(\gamma(t)) \dot{\gamma}(t) dt$$

das *komplexe Wegintegral* von f entlang γ .

Ein weiterer Begriff aus der komplexen Analysis ist die sogenannte Umlaufzahl $\nu_{\gamma}(a)$. Anschaulich gibt diese an, wie oft ein Punkt a von einer Kurve γ in positiver Richtung umlaufen wird. Für geschlossene Kurven kann folgende Charakterisierung angegeben werden, die wir der Vollständigkeit halber hier noch anführen:

Definition 2.7 (Umlaufzahl). Sei γ eine geschlossene Kurve in der komplexen Ebene, dann kann für $a \in \mathbb{C}$ die Umlaufzahl wie folgt berechnet werden:

$$\nu_{\gamma}(a) = \frac{1}{2\pi i} \int_{\gamma} \frac{1}{z - a} dz.$$

Ein weiterer wichtiger Begriff, den wir an dieser Stelle definieren wollen, ist der folgende:

Definition 2.8 (Holomorphie). Sei $U \subset \mathbb{C}$ offene Teilmenge der komplexen Ebene. Eine Funktion $f : U \rightarrow \mathbb{C}$ ist an z_0 *komplex differenzierbar*, wenn der Grenzwert

$$\lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0} =: f'(z_0)$$

existiert. Ist f an allen Punkten $z_0 \in U$ komplex differenzierbar, so nennen wir f *holomorph*.

Allgemein lässt sich sagen, dass Holomorphie eine sehr starke Eigenschaft ist, welche eine Vielzahl an Sätze impliziert. Zunächst wollen wir aber noch den folgenden Begriff definieren:

Definition 2.9 (Isolierte Singularität). Sei $U \subset \mathbb{C}$ offen und $f : U \rightarrow \mathbb{C}$ ein holomorphe Funktion. Betrachten wir nun das Komplement $\mathbb{C} \setminus U$, so bezeichnen wir mit deren isolierten Punkten die sogenannten *isolierten Singularitäten* von f .

Anschaulich gesprochen ist eine isolierte Singularität einer Funktion f also ein Punkt z_0 , um den eine Kreisscheibe $U := \{z \in \mathbb{C} : |z - z_0| < \varepsilon\}$ mit Radius $\varepsilon > 0$ gelegt werden kann, sodass f auf $U \setminus z_0$ definiert ist. Hat man also ein holomorphe Funktionen vorliegen, dann kann mit den obigen Überlegungen der nachfolgende Begriff definiert werden.

Definition 2.10 (Residuum). Sei z_0 isolierte Singularität einer holomorphen Funktion $f : U \rightarrow \mathbb{C}$. Dann ist der nachfolgende Begriff für $\varepsilon > 0$ hinreichend klein wohldefiniert und bezeichnet das sogenannte *Residuum* von f an z_0

$$\text{Res}_{z_0} f := \frac{1}{2\pi i} \int_{|z - z_0| = \varepsilon} f(z) dz$$

Ein Beweis, dass jene Formulierung wohldefiniert ist, kann in jedem Handbuch der Komplexen Analysis, wie auch zum Beispiel [Jän93] gefunden werden. Nun haben wir alle wichtigen Begriffe zusammen, um den sogenannten *Residuensatz* zu formulieren, welcher vor allem für die Berechnung komplexer Wegintegrale sehr hilfreich ist:

Satz 2.11 (Residuensatz). Sei $U \subset \mathbb{C}$ offen und f eine bis auf einer Menge aus isolierten Singularitäten S holomorphe Funktion $f : U \setminus S \rightarrow \mathbb{C}$. Weiters sei γ eine geschlossene Kurve in U , die S nicht trifft und keinen Punkt außerhalb von U umläuft. Dann umläuft die Kurve γ nur endliche viele Singularitäten aus S und es gilt die folgende Gleichheit:

$$\frac{1}{2\pi i} \int_{\gamma} f(z) dz = \sum_{a \in S} \nu_{\gamma}(a) \text{Res}_a f. \quad (2.1)$$

Die angeführten Resultate aus der komplexen Analysis werden wir vor allem in Kapitel 5.1 verwenden, um einen Algorithmus zur Berechnung von Eigenwerten „innerhalb“ einer Kurve γ zu präsentieren.

Abschließend führen wir noch einige grundlegende Resultate und Notationen aus der Numerik ein, die wir in dieser Arbeit verwenden werden. Da diese Arbeit ebenso die Finiten Elemente Methode verwendet, wollen wir zunächst den Begriff einer konformen Triangulierung sauber definieren:

Definition 2.12 (Triangulierung). Sei $\Omega \subset \mathbb{R}^2$ ein Gebiet, dann bezeichnet \mathcal{T} eine *Triangulierung* von Ω , falls die folgenden Bedingungen erfüllt sind:

- \mathcal{T} ist eine Menge aus nicht-degenerierten Dreiecken
- Der Abschluss von Ω wird durch \mathcal{T} überdeckt, es gilt also: $\bar{\Omega} = \bigcup \mathcal{T}$
- für paarweise verschiedene $T, T' \in \mathcal{T}$ mit $T \neq T'$ ist das Maß des Durchschnitts eine Nullmenge, es gilt also $|T \cap T'| = 0$

Die obige Definition bietet also die Grundlage der Finiten Elemente Methode. Darauf aufbauend definieren wir den sogenannten *Spline* Raum:

Definition 2.13. Sei \mathcal{T} eine Triangulierung von $\Omega \subset \mathbb{R}^2$, dann bezeichnen wir mit

$$\mathcal{S}^p(\mathcal{T}) := \{v_h \in C(\Omega) : \forall T \in \mathcal{T} : v_h|_T \text{ ist Polynom der Ordnung } p\}$$

den Raum der \mathcal{T} stückweise affinen und global stetigen Funktionen, kurz den *Spline* Raum erster Ordnung. Man kann zeigen, dass $\dim \mathcal{S}^p(\mathcal{T}) < \infty$ und ebenfalls $\mathcal{S}^p(\mathcal{T}) \subset H^1(\Omega)$ gilt. Analog dazu können wir den folgenden Raum definieren, der für das Lösen von Dirichlet Problemen verwendet wird:

$$\mathcal{S}_D^p(\mathcal{T}) := \{v_h \in \mathcal{S}^p(\mathcal{T}) : \forall z \in \mathcal{K} \cap \bar{\Gamma}_D : v_h(z) = 0\}$$

Dabei bezeichnet \mathcal{K} die Menge aller Ecken von Dreiecken T aus einer Triangulierung \mathcal{T} .

Jene diskreten Räume bieten die Grundlage vieler theoretischer Resultate von Finiten Elemente Diskretisierungen. Als nächstes möchten wir noch für eine gegebene Triangulierung \mathcal{T} die sogenannte *mesh-size* definieren:

Definition 2.14. Sei \mathcal{T} Triangulierung von Ω , dann bezeichnet die wie folgt definierte Funktion $h \in L^\infty(\Omega)$ die *mesh-size* der Triangulierung:

$$h|_T := \text{diam}(T), \quad T \in \mathcal{T}.$$

Diese obige Größe spielt, wie man erahnen kann in der Analyse der Fehlerabschätzung von Finiten Elemente Methoden eine wesentliche Rolle. Für den späteren Gebrauch wollen wir noch das nachfolgende Resultat angeben:

Satz 2.15 (Approximations Theorem). *Sei $u \in H^{p+1}(\Omega)$, dann gelten für $u_h \in \mathcal{S}^p(\mathcal{T})$ die Abschätzungen*

$$\min_{u_h \in \mathcal{S}^p(\mathcal{T})} \|u - u_h\|_{L^2(\Omega)} \leq Ch^{p+1} \|D^{p+1}u\|_{L^2(\Omega)}, \quad (2.2)$$

sowie

$$\min_{u_h \in \mathcal{S}^p(\mathcal{T})} \|u - u_h\|_{H^1(\Omega)} \leq D\sigma(\mathcal{T})h^p \|D^{p+1}u\|_{L^2(\Omega)}, \quad (2.3)$$

wobei die Konstanten C, D nur vom Durchmesser von Ω abhängen und $\sigma(\mathcal{T})$ die shape regularity Konstante ist, siehe auch [NFP21, Abschnitt 3.2.1].

Ein Beweis der obigen Aussage findet sich (für $p = 1$) in [NFP21, Korollar 3.6]. Dieser Satz lässt nun ebenfalls erkennen, dass die Güte der Approximation stark von der Regularität der Lösung u abhängt. Dieses Verhalten werden wir in den kommenden numerischen Experimenten ebenfalls erkennen bzw. nachweisen. Wir haben nun alle notwendigen Resultate und Notationen angeführt, um die eigentliche Problemstellung dieser Arbeit zu präsentieren.

3 Helmholtz Probleme

Da wir uns in dieser Arbeit hauptsächlich mit Lösungen der Helmholtz Gleichung beschäftigen, wollen wir zunächst deren Ursprung bzw. den physikalischen Hintergrund beleuchten. Dazu gehen wir vor wie in [Ihl98, Kapitel 1.1.2] bzw. [Nan22, Kapitel 1] und stellen zunächst fest, dass Helmholtz Probleme hauptsächlich in der Akustik auftauchen. Dabei werden Schallwellen als kleine Störungen in einem kompressiblen Medium, wie etwa Flüssigkeiten oder Gasen betrachtet. Diese können durch ein Geschwindigkeitsfeld $v(x, t)$, dem Druck $p(x, t)$ und der Dichte $\rho(x, t)$ beschrieben werden, wobei wir mit $x \in \mathbb{R}^3$ den Ort und $t \geq 0$ die Zeit bezeichnen. Überdies definieren wir den stationären Zustand als $v_0 = 0, p_0 = \text{const}$, sowie $\rho_0 = \rho_0(x)$. Deren Wechselwirkung wird durch die sogenannte linearisierte Euler Gleichung

$$\rho_0 \frac{\partial v}{\partial t} + \nabla p = 0,$$

die linearisierte Kontinuitätsgleichung

$$\frac{\partial \rho}{\partial t} + \text{div}(\rho_0 v) = 0,$$

sowie die linearisierte Zustandsgleichung

$$p = c^2 \rho$$

beschrieben, wobei $c(x)$ die lokale Schallgeschwindigkeit bezeichnet. Durch geeignete Kombination der drei Gleichungen erhalten wir schlussendlich für den Druck p die nachfolgende (Wellen-) Gleichung:

$$\frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} - \Delta p = 0. \tag{3.1}$$

Ein Ansatz, um Lösungen der obigen Gleichung zu erhalten, besteht darin, Wellen mit einer Kreisfrequenz κ und (ortsabhängigen) Amplitude $u(x)$, also

$$p(x, t) = \Re(u(x)e^{-i\kappa t}) \tag{3.2}$$

anzusetzen. Durch Einsetzen von (3.2) in (3.1) erhalten wir nun jene Gleichung:

$$-\Delta u - \frac{\kappa^2}{c^2} u = 0.$$

Falls zusätzlich ein homogenes Medium zu Grunde liegt, also $c(x) = \text{const}$ gilt, dann erhalten wir mit $\omega := \frac{\kappa}{c}$ schlussendlich die bekannte *Helmholtz Gleichung*

$$-\Delta u - \omega^2 u = 0. \tag{3.3}$$

Kennt man eine Lösung u der Helmholtz Gleichung, kann die physikalisch relevante Größe $p(x, t)$ durch (3.2) rückkonstruiert werden. Somit schränken wir die weitere Analyse auf Problem (3.3) ein, deren Lösbarkeit wir in den folgenden Abschnitten diskutieren wollen.

3.1 Helmholtz Gleichung auf beschränkten Gebieten

Wir starten diesen Abschnitt, indem wir Helmholtz Streuprobleme auf einem zunächst beschränkten Gebiet $\Omega \subset \mathbb{R}^d$ mit $d \in \{2, 3\}$ betrachten, wobei wir zusätzlich $\partial\Omega$ als Lipschitz stetig voraussetzen mit äußerem Normalenvektor η . Gesucht ist also eine Lösung $u : \Omega \rightarrow \mathbb{C}$ des folgenden Problems mit $\omega \in \mathbb{C}$:

$$\begin{aligned} -\Delta u - \omega^2 u &= f, & \text{in } \Omega \\ \frac{\partial u}{\partial \eta} &= g, & \text{auf } \partial\Omega. \end{aligned} \quad (3.4)$$

Dabei setzen wir zusätzlich die rechte Seite $f \in L^2(\Omega)$ und $g \in L^2(\partial\Omega)$ voraus. Mit diesen Voraussetzungen können wir zur schwachen Formulierung von (3.4) übergehen, wodurch wir nach partieller Integration die folgende Problemstellung erhalten:

Suche $u \in H^1(\Omega)$, sodass gilt:

$$\int_{\Omega} \nabla u \overline{\nabla v} \, dx - \omega^2 \int_{\Omega} u \bar{v} \, dx = \int_{\Omega} f \bar{v} \, dx + \int_{\partial\Omega} g \bar{v} \, ds, \quad \forall v \in H^1(\Omega). \quad (3.5)$$

Somit erhalten wir die auf $H^1(\Omega)$ stetigen, symmetrischen Sesquilinearformen

$$a(u, v) := \int_{\Omega} \nabla u \overline{\nabla v} \, dx \quad \text{und} \quad b(u, v) := \int_{\Omega} u \bar{v} \, dx \quad (3.6)$$

sowie die stetige Linearform

$$f(v) := \int_{\Omega} f \bar{v} \, dx + \int_{\partial\Omega} g \bar{v} \, ds.$$

Dadurch lässt sich (3.5) schreiben als

$$a(u, v) - \omega^2 b(u, v) = f(v), \quad \forall v \in H^1(\Omega).$$

Für den Rest des Abschnitts werden wir die (eindeutige) Lösbarkeit von Problem (3.5) diskutieren. Dafür benötigen wir die folgende Begriffsbildung: Wir nennen $(\lambda, u) \in \mathbb{C} \times H^1(\Omega) \setminus \{0\}$ Eigenpaare des negativen Laplace-Operators $-\Delta$, wenn gilt

$$a(u, v) = \lambda b(u, v), \quad \forall v \in H^1(\Omega). \quad (3.7)$$

Diese spielen (wie wir sehen werden) eine große Rolle bei der Frage nach eindeutigen Lösungen von (3.4). Zunächst führen wir das folgende Resultat an:

Satz 3.1. Sei $\epsilon \in \mathbb{C} \setminus \mathbb{R}_{\leq 0}$, dann ist die Sesquilinearform

$$s_\epsilon(u, v) := a(u, v) + \epsilon b(u, v), \quad u, v \in H^1(\Omega),$$

wobei a, b wie in (3.6) definiert, stetig und koerziv. Dies bedeutet, es existieren Konstanten $C, D > 0$, sodass

$$\begin{aligned} |s(u, v)| &\leq C \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} \\ |s(u, u)| &\geq D \|u\|_{H^1(\Omega)}^2 \end{aligned}$$

für alle $u, v \in H^1(\Omega)$ gilt.

Für einen Beweis verweisen wir auf [Nan22, Kapitel 2]. Obiges Resultat impliziert insbesondere, dass Problem (3.5) für $\omega \in \mathbb{C} \setminus \mathbb{R}_{\geq 0}$ nach dem Lemma von Lax-Milgram eine eindeutige Lösung besitzt. Wir wollen nun noch diskutieren, für welche $0 < \kappa \in \mathbb{R}$ wir dennoch eine eindeutige Lösung erwarten können. Dazu benötigen wir zunächst noch das folgende Hilfsresultat:

Satz 3.2. Sei $\epsilon > 0$. Dann existiert ein linearer, kompakter Operator $T_\epsilon : H^1(\Omega) \rightarrow H^1(\Omega)$, sodass

$$s_\epsilon(T_\epsilon u, v) = b(u, v), \quad \forall u, v \in H^1(\Omega).$$

Des Weiteren ist T_ϵ selbstadjungiert bezüglich, des von $s_\epsilon(\cdot, \cdot)$ induzierten Skalarproduktes, es gilt also $s_\epsilon(T_\epsilon u, v) = s_\epsilon(u, T_\epsilon v)$. Für Eigenpaare (λ, u) von T_ϵ gilt außerdem die folgende Äquivalenz:

$$(\lambda, u) \text{ ist Eigenpaar von (3.7)} \Leftrightarrow \left(\frac{1}{\lambda + \epsilon}, u\right) \text{ ist Eigenpaar von } T_\epsilon.$$

Ein Beweis für die obigen Aussagen findet sich unter anderem in [Nan22, Satz 2.6 und Satz 2.9]. Wir sehen also, dass das Eigenwertproblem (3.7) äquivalent zu jenem eines kompakten, selbstadjungierten Operators T_ϵ auf einem Hilbertraum ist. In diesem Setting greift die Spektraltheorie aus der Funktionalanalysis, womit wir den folgenden Existenzsatz formulieren können:

Korollar 3.3 (Existenz- und Eindeutigkeit von Helmholtz-Quellproblemen). Es gelten die Voraussetzungen von zu Beginn des Kapitels. Dann besitzt das Helmholtz Streuproblem (3.5) genau dann eine eindeutige Lösung $u \in \Omega$, wenn ω^2 kein Eigenwert des Eigenwertproblems (3.7) ist.

Wir halten an dieser Stelle fest, dass wir für Helmholtz Probleme auf beschränkten Gebieten mit der Spektraltheorie von kompakten Operatoren und dem Lemma von Lax-Milgram vergleichsweise einfach auf Existenz und Eindeutigkeit von Lösungen schließen können. Da diese Arbeit jedoch hauptsächlich um Helmholtz Probleme auf unbeschränkten Gebieten handelt, versuchen wir nun diese Frage auf unbeschränkten Gebieten zu klären. Der wesentliche Unterschied hierbei ist, dass $-\Delta$ dann nicht mehr kompakt ist und somit ein kontinuierliches Spektrum mit komplexen Eigenwerten aufweist. Auch beim Aufstellen einer schwachen Formulierung muss man vorsichtig sein, da die Unbeschränktheit des

Gebiets zu *uneigentlichen Integralen* führt. Wir werden im folgenden Abschnitt Techniken präsentieren, wie wir Probleme auf unbeschränkten Gebieten dennoch auf schwache Formulierungen auf beschränkten Gebieten zurückführen können. Dies wird in weiterer Hinsicht ebenfalls den Weg zur Verwendung von numerischen Verfahren ebnet.

3.2 Helmholtz Gleichung auf unbeschränkten Gebieten

Nachdem wir im vorangehenden Abschnitt relativ prägnant auf die Existenz und Eindeutigkeit von Helmholtz Problemen eingegangen sind, betrachten wir nun diese Probleme auf unbeschränkten Gebieten. Dazu sei $K \subset \mathbb{R}^d$ kompakt mit ∂K Lipschitz stetig und $d \in \{2, 3\}$ gegeben. Als zu Grunde liegende Geometrie betrachten wir nun das Komplement $\Omega := \mathbb{R}^d \setminus K$. Als Ausgangspunkt für die nachfolgenden Untersuchungen dient also die folgende Problemstellung:

$$\begin{aligned} -\Delta u - \omega^2 u &= 0, \quad \text{in } \Omega \\ \frac{\partial u}{\partial \eta} &= \phi, \quad \text{auf } \partial K. \end{aligned} \tag{3.8}$$

Dabei setzen wir überdies $\omega > 0$ voraus. Die Unbeschränktheit des Gebiets hat zunächst zur Folge, dass wir eine zusätzliche Bedingung an die Lösung u stellen müssen, sofern wir an eindeutiger Lösbarkeit interessiert sind. Typischerweise wird bei unbeschränkten Gebieten eine Bedingung in Form von $\lim_{|x| \rightarrow \infty} u(x)$ gestellt. Eine Möglichkeit ist die sogenannte *Sommerfeld'sche Ausstrahlungsbedingung*, deren Herleitung wir im folgenden Abschnitt präsentieren wollen.

3.2.1 Sommerfeld Bedingung

Wie bereits erwähnt, erhält man die Helmholtz Gleichung, indem man einen zeitharmonischen Ansatz $u(x, t) = \Re(u(x)e^{-i\omega t})$ in die Wellengleichung einsetzt. Somit repräsentiert die Lösung der Helmholtz Gleichung ebenfalls Wellenphänomene, wonach es physikalisch sinnvoll ist zu fordern, dass Wellen auf unbeschränkten Gebieten im „Unendlichen“ nicht reflektieren, sondern gewissermaßen absorbiert werden. Diese physikalische Eigenschaft gewährleistet uns, wie wir später sehen werden, eindeutige Lösungen. In diesem Abschnitt versuchen wir diese recht grob formulierte, physikalische Eigenschaft mathematisch präziser zu formulieren, dafür gehen wir vor wie in [Ih198, Kapitel 1.1.3]. Zunächst benötigen wir aber den folgenden Begriff:

Definition 3.4 (Green'sche Funktion). Sei $\Omega \subset \mathbb{R}^n$ und L ein linearer, formaler Differentialoperator der Form

$$L(u) := -\operatorname{div}(A(x)\nabla u) + b(x) \cdot \nabla u + c(x)u,$$

wobei $A(x) \in \mathbb{R}^{n \times n}$, $b(x) \in \mathbb{R}^n$ und $c(x) \in \mathbb{R}$, dann bezeichnen wir als *Green'sche Funktion* $G(x, \cdot)$ die Lösung von

$$L(G(x, \cdot)) = \delta_x, \quad \text{wobei } \delta_x(y) = \begin{cases} 0, & x \neq y \\ \infty, & x = y. \end{cases}$$

Sei nun $\Omega \subset \mathbb{R}^d$ wie zu Beginn des Kapitels definiert, dann kann man zeigen, dass für die Helmholtz Gleichung die Green'schen Funktionen wie folgt gegeben sind:

$$G(x, y) = \begin{cases} \frac{i\mathcal{H}_0^{(1)}(\omega \cdot |x-y|)}{4}, & d = 2 \\ \frac{e^{i\omega|x-y|}}{4\pi|x-y|}, & d = 3. \end{cases}$$

Dabei bezeichnet $\mathcal{H}_0^{(1)}$ die *Hankel Funktion* erster Ordnung, welche einfach gesagt als Lösung der *sphärischen Bessel'schen Differentialgleichung* für $n = 0$ charakterisiert ist. Wir verzichten an dieser Stelle auf eine genaue Definition, verweisen dazu auf [Th198, Abschnitt 2.1.2] für Details. Mit Hilfe der Green'schen Funktionen können wir nun ähnlich wie in [Kre14, Theorem 6.5] nachweisen, dass für Lösungen von (3.8) die folgende Integral-Repräsentation gilt:

$$u(x) = \int_{\partial\Omega} \left[u(y) \frac{\partial}{\partial\eta} G(x, y) - G(x, y) \frac{\partial}{\partial\eta} u(y) \right] dS(y).$$

Wählen wir nun eine Kugel B_r mit Radius $r > 0$, die K ganz umschließt, dann können wir $\Omega \cap B_r$ betrachten und sehen, dass $\partial\Omega = \lim_{r \rightarrow \infty} \partial K \cup \partial B_r$. Aus den vorausgehenden Überlegungen fordern wir, dass u im unendlichen absorbiert, also $\lim_{|x| \rightarrow \infty} u(x) = 0$ gelten soll und somit ebenfalls

$$\lim_{r \rightarrow \infty} \int_{\partial B_r} \left[u(y) \frac{\partial}{\partial\eta} G(x, y) - G(x, y) \frac{\partial}{\partial\eta} u(y) \right] dS(y) = 0. \quad (3.9)$$

Für $x \in \Omega$ wählen wir nun $r > 0$ hinreichend groß, sodass für $y \in \partial B_r$ die folgende Identität gilt:

$$r = |y| \sim |x - y| =: R \quad \text{sowie} \quad \frac{\partial}{\partial\eta} = \frac{\partial}{\partial r} \sim \frac{\partial}{\partial R}.$$

Betrachten wir zunächst den Fall $d = 3$, dann erhalten wir mit der Green'schen Funktion aus (3.9) die folgende Identität:

$$\int_{\partial B_r} \frac{1}{R} \left[i\omega u - \frac{u}{R} - \frac{\partial}{\partial R} u \right] \frac{e^{i\omega R}}{4\pi} dS(r).$$

Verwenden wir wie üblich Kugelkoordinaten (θ, φ) für das Randintegral, so ergibt sich mit einer simplen Rechnung für das Oberflächenintegral

$$dS = r^2 \sin \varphi \underbrace{d\varphi d\theta}_{=: d\hat{S}}.$$

Eingesetzt in die obige Gleichung, folgern wir also, dass für Lösungen u von (3.8) gelten soll:

$$\lim_{r \rightarrow \infty} \int_{\partial B_r} r \cdot \left[i\omega u - \frac{u}{r} - \frac{\partial}{\partial r} u \right] \frac{e^{i\omega r}}{4\pi} d\hat{S} = 0. \quad (3.10)$$

Betrachten wir nun den Fall $d = 2$, dann können wir mit einem ähnlichen Argument wie zuvor und den asymptotischen Verhalten der sogenannten *Hankel Funktionen* (siehe abermals [Ih198, Abschnitt 2.1.2]):

$$\mathcal{H}_n^{(1)}(r) \sim \frac{\exp(ir)}{\sqrt{r}}, \quad \mathcal{H}_n^{(1)'}(r) \sim i \frac{\exp(ir)}{\sqrt{r}}$$

die Gleichung (3.9) schreiben als:

$$\int_{\partial B_r} \frac{1}{\sqrt{r}} \left[i\omega u - \frac{\partial}{\partial r} u \right] e^{i\omega r} dS(r).$$

Analog wie zuvor gehen wir zu Polarkoordinaten über, um das Randintegral darzustellen. Wir erhalten also mit $dS = rd\hat{S}$

$$\int_{\partial B_r} \sqrt{r} \left[i\omega u - \frac{\partial}{\partial r} u \right] e^{i\omega r} d\hat{S}. \quad (3.11)$$

Fassen wir nun (3.11) und (3.10) zusammen, erhalten wir die folgenden Bedingungen an u :

$$u = \mathcal{O}\left(\frac{1}{r^{(d-1)/2}}\right), \quad (3.12)$$

sowie

$$\lim_{r \rightarrow \infty} r^{(d-1)/2} \left(\frac{\partial}{\partial r} u - i\omega u \right) = 0. \quad (3.13)$$

Mit zusätzlichem Aufwand kann wie in [CK98] gezeigt werden, dass für Lösungen von (3.8), welche Bedingung (3.13) erfüllen, auch automatisch Bedingung (3.12) folgt. Dies ermöglicht uns nun die folgende Ausstrahlungsbedingung aufzustellen.

Definition 3.5. Sei u Lösung von (3.8) wobei $\Omega \in \mathbb{R}^d$ mit $d \in \{2, 3\}$. Wir sagen, dass u die *Sommerfeld'sche Ausstrahlungsbedingung* erfüllt, wenn folgender Grenzwert gleichmäßig in $x/|x|$ gilt:

$$\lim_{|x| \rightarrow \infty} |x|^{\frac{d-1}{2}} \left(\frac{\partial}{\partial |x|} u(x) - i\omega u(x) \right) = 0. \quad (3.14)$$

Mit Hilfe der Sommerfeld'schen Ausstrahlungsbedingung können wir nun das folgende Modellproblem formulieren für das wir, wie wir sehen werden, eindeutige Lösbarkeit erwarten können. Mit den Voraussetzungen von zu Beginn dieses Kapitels suchen wir also die Lösung $u : \Omega \rightarrow \mathbb{C}$ der folgenden Gleichungen:

$$\begin{aligned} -\Delta u - \omega^2 u &= 0 \text{ in } \Omega \\ \frac{\partial u}{\partial \eta} &= \phi \text{ auf } \partial K \\ u &\text{ erfüllt (3.14).} \end{aligned} \quad (3.15)$$

In den weiteren Abschnitten werden wir Techniken vorstellen, die uns ermöglichen das obige Modellproblem numerisch zu lösen. Wesentlicher Bestandteil dabei spielt der sogenannte *Dirichlet-to-Neumann Operator*.

3.2.2 Dirichlet-to-Neumann Operator

Das zentrale Thema dieser Arbeit wird das Lösen von Problem (3.15) mittels Finite Elemente Diskretisierung sein. Dafür werden wir im folgenden das unbeschränkte Gebiet Ω in zwei Teilgebiete $\Omega_{\text{int}}, \Omega_{\text{ext}}$ unterteilen. Zusätzlich fordern wir $\Omega_{\text{int}} \cap \Omega_{\text{ext}} = \emptyset$ und $\overline{\Omega_{\text{int}}} \cup \overline{\Omega_{\text{ext}}} = \Omega$, sowie, dass Ω_{int} beschränkt ist. Somit erhalten wir zwei Teilprobleme

- „Innenraumproblem“:

$$\begin{aligned} -\Delta u_{\text{int}} - \omega^2 u_{\text{int}} &= 0, \quad \text{in } \Omega_{\text{int}} \\ \frac{\partial u_{\text{int}}}{\partial \eta} &= \phi, \quad \text{auf } \partial K \\ \frac{\partial u_{\text{int}}}{\partial \eta} &= f' \quad \text{auf } \Gamma_a \end{aligned} \tag{3.16}$$

- „Außenraumproblem“:

$$\begin{aligned} -\Delta u_{\text{ext}} - \omega^2 u_{\text{ext}} &= 0, \quad \text{in } \Omega_{\text{ext}} \\ u_{\text{ext}} &= f \quad \text{auf } \Gamma_a \\ u_{\text{ext}} &\text{ erfüllt (3.14) ,} \end{aligned} \tag{3.17}$$

die über die zusätzlichen Bedingungen $\frac{\partial u_{\text{int}}}{\partial \eta} = \frac{\partial u_{\text{ext}}}{\partial \eta}$ sowie $u_{\text{ext}} = u_{\text{int}}$ auf Γ_a miteinander gekoppelt sind. Wesentlich in der weiteren Herangehensweise ist die Annahme, dass für gegebenes $f \in H^{1/2}(\Gamma_a)$ das Problem (3.17) eindeutig lösbar ist, womit wir die Beziehung $f \mapsto f'$ betrachten können. Diese ist Gegenstand der folgenden Definition:

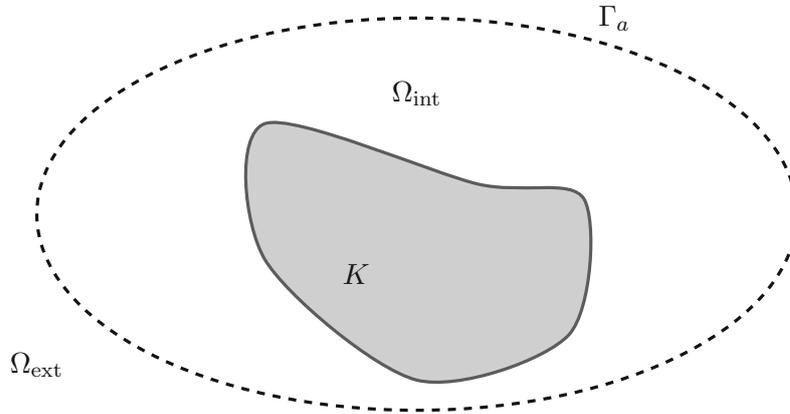


Abbildung 3.1: Zerlegung von unbeschränkten Gebieten mittels einem transparenten Rand

Definition 3.6 (Dirichlet-to-Neuman Operator). Mit den Voraussetzungen von zu Beginn des Kapitels, sei u_{ext} die eindeutige Lösung von (3.17), wobei $f \in H^{1/2}(\Gamma_a)$. Dann ist der

durch

$$\begin{aligned} \mathcal{Dt}\mathcal{N} : H^{1/2}(\Gamma_a) &\rightarrow H^{-1/2}(\Gamma_a) = (H^{1/2}(\Gamma_a))' \\ f &\mapsto \mathcal{Dt}\mathcal{N} f := \frac{\partial u_{\text{ext}}}{\partial \eta} \Big|_{\Gamma_a} \end{aligned}$$

definierte Operator linear und wohldefiniert.

Wir verzichten an der Stelle auf einen genauen Beweis, da die Wohldefiniertheit aus der eindeutigen Lösbarkeit und der Linearität aus der von Problem (3.17) unmittelbar folgt.

Bemerkung 3.7. Wir bemerken an dieser Stelle weiters, dass Definition 3.6 implizit durch die Koppelung zwischen Ω_{int} und Ω_{ext} gegeben ist. Für spezielle Geometrien Γ_a werden wir sehen, dass auch eine explizite Darstellung des Dirichlet-to-Neumann Operators möglich ist.

Mit Hilfe des Operators $\mathcal{Dt}\mathcal{N}$ können wir nun eine zu (3.15) äquivalente schwache Formulierung in einem beschränkten Gebiet angeben. Insbesondere erhalten wir nach Multiplikation mit einer Testfunktion $v \in H^1(\Omega_{\text{int}})$ und partieller Integration die folgende schwache Formulierung:

Suche $u \in H^1(\Omega_{\text{int}})$, sodass

$$\int_{\Omega_{\text{int}}} \nabla u \cdot \overline{\nabla v} \, dx - \omega^2 \int_{\Omega_{\text{int}}} u \bar{v} \, dx - \langle \mathcal{Dt}\mathcal{N}(u|_{\Gamma_a}), v|_{\Gamma_a} \rangle_{H^{-1/2}(\Gamma_a) \times H^{1/2}(\Gamma_a)} = \int_{\partial K} \phi \bar{v} \, ds \quad (3.18)$$

für alle $v \in H^1(\Omega_{\text{int}})$ gilt. Dabei bezeichnet $\langle \cdot, \cdot \rangle_{V' \times V}$ wie üblich das Dualitätsprodukt eines Banachraumes V und seines dualen Raumes V' , definiert durch $\langle u, v \rangle = u(v)$. Falls zusätzlich $\mathcal{Dt}\mathcal{N}(u|_{\Gamma_a}) \in L^2(\Gamma_a)$ gilt, so lässt sich das Dualitätsprodukt in der schwachen Formulierung durch das L^2 Skalarprodukt schreiben als:

$$\langle \mathcal{Dt}\mathcal{N}(u|_{\Gamma_a}), v|_{\Gamma_a} \rangle = \int_{\Gamma_a} \mathcal{Dt}\mathcal{N}(u|_{\Gamma_a}) \overline{v|_{\Gamma_a}} \, ds,$$

womit sich die Gleichung in (3.18) vereinfacht zu

$$s(u, v) := \int_{\Omega_{\text{int}}} \nabla u \cdot \overline{\nabla v} \, dx - \omega^2 \int_{\Omega_{\text{int}}} u \bar{v} \, dx - \int_{\Gamma_a} \mathcal{Dt}\mathcal{N}(u|_{\Gamma_a}) \overline{v|_{\Gamma_a}} \, ds = \int_{\partial K} \phi \bar{v} \, ds =: f(v).$$

Diese Darstellung ermöglicht nun zumindest theoretisch die Verwendung von numerischen Verfahren. Zunächst werden wir jedoch der Frage der eindeutigen Lösbarkeit nachgehen. Man sieht unmittelbar, dass s eine stetige Sesquilinearform auf $H^1(\Omega_{\text{int}})$ ist. Diese Sesquilinearform ist jedoch im allgemeinen nicht koerziv, was man an dem negativen Vorzeichen des zweiten Terms erahnen kann. Im nächsten Abschnitt werden wir andere Techniken und Resultate präsentieren, die uns unter gewissen Voraussetzungen ebenfalls eindeutige Lösbarkeit von nicht koerziven Problemen garantieren.

3.3 Lösbarkeit und Eindeutigkeit auf unbeschränkten Gebieten

Während die Koerzivität von stetigen Sesquilinearformen bereits hinreichend für die Existenz von Lösungen ist (vgl. Lemma von Lax-Milgram), muss für nicht koerzive Probleme eine zusätzliche Bedingung gefordert werden. Dies wird in dem folgenden Satz präzisiert:

Satz 3.8 (Satz von Babuska). *Sei V ein Hilbertraum und $s : V \times V \rightarrow \mathbb{C}$ eine stetige Sesquilinearform. Dann sind die beiden Aussagen äquivalent:*

1. Für jedes $f \in V'$ existiert genau ein $u \in V$, sodass:

$$s(u, v) = f(v), \quad \forall v \in V$$

2. Die Sesquilinearform $s(.,.)$ erfüllt die folgenden Bedingungen:

- *inf-sup Bedingung:*

$$0 < \beta := \inf_{0 \neq v \in V} \sup_{0 \neq w \in V} \frac{|s(v, w)|}{\|v\|_V \|w\|_V} \quad (3.19)$$

- *nicht degeneriert Bedingung:*

$$\forall 0 \neq v \in V \exists w \in V : |s(v, w)| > 0. \quad (3.20)$$

Ein Beweis findet sich unter anderem in [Ihl98, Theorem 2.15].

Ist zusätzlich der Hilbertraum V endlich dimensional, dann kann man nachweisen, dass bereits eine der beiden Bedingungen aus Satz 3.8 hinreichend für die Existenz eindeutiger Lösungen ist. Es gilt nämlich der folgende Satz:

Satz 3.9. *Sei V ein Hilbertraum mit $\dim V < \infty$ und $s : V \times V \rightarrow \mathbb{C}$ eine stetige Sesquilinearform. Dann erfüllt s genau dann die inf-sup Bedingung (3.19), wenn s nicht degeneriert ist, also (3.20) erfüllt.*

Dieser Satz findet sich unter anderem in [Hac92, Abschnitt 6.5.4] oder für reelle Banachräume ebenfalls in [NFP21, Proposition 6.4]. Die obigen beiden Sätze liefern uns somit bereits Kriterien für eindeutige Lösungen von indefiniten Problemen. In der Praxis ist es jedoch oft schwer, die inf-sup Bedingung nachzuweisen. Im folgenden wollen wir deshalb ebenfalls eine alternative Herangehensweise mit Hilfe der Riesz-Fredholm Theorie vorstellen. Dafür brauchen wir aber zunächst die folgende Definition:

Definition 3.10 (Garding Ungleichung). Sei $V \hookrightarrow H \hookrightarrow V'$ ein Gelfand Tripel. Wir sagen, dass eine Sesquilinearform $s : V \times V \rightarrow \mathbb{C}$ die *Garding Ungleichung* erfüllt, wenn Konstanten $C_1, C_2 > 0$ existieren, sodass gilt:

$$|s(u, u)| \geq C_1 \|u\|_V^2 - C_2 \|u\|_H^2 \quad \forall u \in V.$$

Ist zusätzlich die Einbettung $V \hookrightarrow H$ kompakt, so können wir das folgende, aus der Funktionalanalysis bekannte Resultat anführen:

Satz 3.11 (Fredholmalternative). Sei $V \hookrightarrow H \hookrightarrow V'$ Gelfand Tripel mit kompakter Einbettung $V \hookrightarrow H$ und $s : V \times V \rightarrow \mathbb{C}$ eine stetige Sesquilinearform, die die Garding Ungleichung 3.10 für $C_1, C_2 > 0$ erfüllt. Bezeichnet nun S den durch die Sesquilinearform $s(\cdot, \cdot)$ induzierten Operator, dann folgt aus der Injektivität von S , bzw äquivalent dazu

$$\langle Su, v \rangle = s(u, v) = 0 \quad \forall v \in V \Rightarrow u = 0,$$

bereits die Bijektivität von S .

Beweis. Da ad-hoc nicht klar ist, welche Rolle die Kompaktheit der Einbettung $V \hookrightarrow H$ spielt, wollen wir an dieser Stelle einen Beweis wie in [NFP21, Theorem 6.19] angeben. Nach Voraussetzung erfüllt $s(\cdot, \cdot)$ die Garding Ungleichung, es gilt also

$$s(u, u) \geq C_1 \|u\|_V^2 - C_2 \|u\|_H^2$$

für Konstanten $C_1, C_2 > 0$. Wir können also die Sesquilinearform $\tilde{s}(u, v) := s(u, v) + C_2(u, v)_H$ betrachten, welche offensichtlich stetig und koerziv ist. Somit folgt aus dem Lemma von Lax Milgram die Bijektivität von \tilde{s} bzw dem induzierten Operator $\tilde{S} : V \rightarrow V'$. Betrachten wir nun die Differenz $K := \tilde{S} - S$, so folgt nach Definition

$$\langle Ku, v \rangle_{V' \times V} = \langle (\tilde{S} - S)u, v \rangle = C_2 \langle u, v \rangle_H$$

und mit der kompakten Einbettung $V \hookrightarrow H$ bereits die Kompaktheit von K . Somit können wir S schreiben als

$$S = \tilde{S} - (\tilde{S} - S) = \tilde{S} - K = \tilde{S}(\text{Id} - \tilde{S}^{-1}K).$$

Da \tilde{S}^{-1} existiert und stetig ist, folgt mit der Kompaktheit von K auch die Kompaktheit von $\tilde{S}^{-1}K$. Nun ist aber S nach Voraussetzung injektiv, das bedeutet also insbesondere auch, dass $(\text{Id} - \tilde{S}^{-1}K)$ injektiv ist und somit 1 kein Eigenwert von $\tilde{S}^{-1}K$ ist. Die Theorie kompakter Operatoren besagt also, dass

$$1 \in \rho(\tilde{S}^{-1}K) := \{\lambda \in \mathbb{C} : (\tilde{S}^{-1}K - \lambda \text{id}) \text{ ist stetig invertierbar}\}.$$

Als Verknüpfung zweier bijektiver Operatoren ist somit S selbst auch bijektiv, was genau die Behauptung beweist. \square

Ein weiteres nützliches Werkzeug, um auf Eindeutigkeit von schwachen Lösungen zu schließen, ist der folgende Satz:

Satz 3.12. Sei $\Omega \subset \mathbb{R}^n$ ein beschränktes Gebiet, wobei $\partial\Omega$ hinreichend glatt ist. Weiters sei $\emptyset \neq \Gamma \subset \partial\Omega$ und $q \in L^\infty(\Omega)$. Sei $u \in H^2(\Omega)$ Lösung von

$$\begin{aligned} -\Delta u + qu &= 0 && \text{in } \Omega \\ u &= 0 && \text{auf } \Gamma \\ \partial_\eta u &= 0 && \text{auf } \Gamma. \end{aligned}$$

Dann gilt bereits $u = 0$ in Ω .

Der Beweis verwendet das *unique continuation principle*, welches für allgemeine elliptische Operatoren gilt. Jener Satz besagt im Wesentlichen folgendes: wenn Lösungen $u \in H^2(\Omega)$ in einer Umgebung $G \subset \Omega$ verschwinden, dann ist die Lösung bereits auf ganz Ω identisch null. Die Voraussetzung $H^2(\Omega)$ ist im Allgemeinen keine wesentliche Einschränkung, da dies in den meisten Fällen für elliptische Probleme ohnehin durch diverse Regularitätssätze für zum Beispiel konvexes Ω oder $\partial\Omega \in C^1$ hervorgeht. Ein genauer Beweis der Aussage findet sich in [Lei86, Kapitel 4.3].

Kehren wir nun zurück zu dem ursprünglich definierten Problem, dann ergibt sich wie bereits diskutiert für Helmholtz Probleme auf einem unbeschränkten Gebiet nach Einführung einer transparenten Randbedingung die folgende schwache Formulierung:

Suche $u \in H^1(\Omega_{\text{int}})$, sodass für alle $v \in H^1(\Omega_{\text{int}})$

$$s(u, v) := \int_{\Omega_{\text{int}}} \nabla u \cdot \overline{\nabla v} \, dx - \omega^2 \int_{\Omega_{\text{int}}} u \overline{v} \, dx - \int_{\Gamma_a} \mathcal{D}\mathcal{N}(u|_{\Gamma_a}) \overline{v|_{\Gamma_a}} \, ds = \int_{\partial K} \phi \overline{v} \, ds.$$

Können wir nachweisen, dass s die Garding Ungleichung im Sinne von Definition 3.10 erfüllt, dann ist nach Satz 3.11 die Eindeutigkeit von Lösungen bereits hinreichend für die Bijektivität, also die Existenz eindeutiger Lösungen. Wählen wir als Gelfand Triple $H^1(\Omega_{\text{int}}) \hookrightarrow L^2(\Omega_{\text{int}}) \hookrightarrow (H^1(\Omega_{\text{int}}))'$, so müssten wir demnach nachweisen, dass Konstanten $C_1, C_2 > 0$ existieren, sodass

$$|s(u, u)| \geq C_1 \|u\|_{H^1(\Omega_{\text{int}})}^2 - C_2 \|u\|_{L^2(\Omega_{\text{int}})}^2$$

gilt. Problematisch ist in diesem Fall der Dirichlet-to-Neumann Operator, da ja $\mathcal{D}\mathcal{N}(u|_{\Gamma})$ zunächst nur in $H^{-1/2}(\Gamma_a)$ liegt. Bei genauerem Betrachten des Beweises von Satz 3.11 haben wir allerdings ebenfalls gesehen, dass der entscheidende Punkt war, den induzierten Operator S als kompakte Störung eines Bijektiven Operators zu schreiben. Die implizite Schreibweise aus Definition 3.6 gibt uns jedoch kaum Möglichkeiten für geeignete Abschätzungen. In dem folgenden Abschnitt setzen wir also daran, den Dirichlet-to-Neumann Operator für konkrete Fälle explizit zu berechnen. Wir werden sehen, dass uns dies über eine unendliche Reihendarstellung gelingt. Für numerische Verfahren ist dies zwar auf den ersten Blick wenig nützlich, für die Theorie der Existenz einer eindeutigen (kontinuierlichen) Lösung jedoch unabdingbar.

3.3.1 Eindeutige Lösbarkeit auf dem Ganzraum

In diesem Abschnitt wollen wir nun die Lösbarkeit von (3.15) auf dem Komplement einer kompakten Menge $K \subset \mathbb{R}^3$ diskutieren. Dazu legen wir wie in Abbildung 3.2 veranschaulicht eine Kugel $B_a := \{(x, y, z)^T \in \mathbb{R}^3 : x^2 + y^2 + z^2 < a^2\}$ mit Radius $a > 0$ hinreichend groß (sodass zumindest $B_a \supset K$ gilt) um K . Daraus resultierend erhalten wir als beschränktes Gebiet $\Omega_{\text{int}} := B_a \setminus K$ und als transparenten Rand Γ_a die Kugeloberfläche ∂B_a . Die folgenden Überlegungen basieren weitestgehend auf [HN09, Kapitel 3.1.1]. Wie bereits

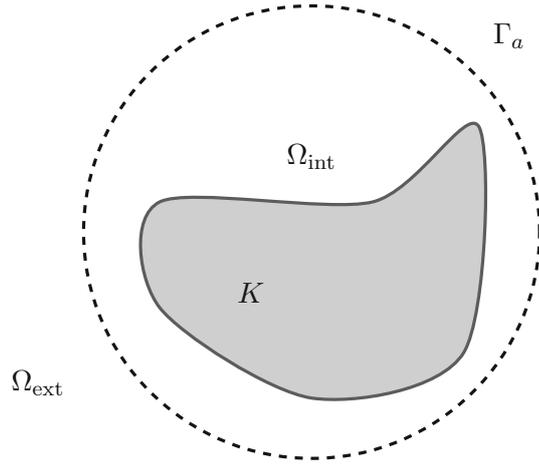


Abbildung 3.2: Verwendung einer Kugeloberfläche als transparenten Rand

erwähnt werden wir im folgenden die spezielle Geometrie der Kugeloberfläche ausnützen, um eine explizite Darstellung des Dirichlet-to-Neumann Operators zu erhalten. Um eine Lösung u_{ext} des Außenraumproblems

$$\begin{aligned} -\Delta u_{\text{ext}} - \omega^2 u_{\text{ext}} &= 0, \quad \text{in } \Omega_{\text{ext}} := \mathbb{R}^3 \setminus B_a \\ u_{\text{ext}} &= f \quad \text{auf } \partial B_a \\ u_{\text{ext}} &\text{ erfüllt (3.14)} \end{aligned} \quad (3.21)$$

zu erhalten, gehen wir zunächst in Polar- bzw. Kugelkoordinaten über. Dazu bezeichnen wir mit $a > 0$ den (verallgemeinerten) Radius $r := |x|/a \in \mathbb{R}^+$ und weiters $\hat{x} := x/r \in \Gamma_a$. Sei nun $x \in \mathbb{R}^3$ beliebig, so können wir dieses schreiben als $x = r\hat{x}$. Die Parametrisierung des transparenten Randes Γ_a erfolgt demnach durch

$$\begin{pmatrix} \varphi \\ \theta \end{pmatrix} \mapsto a \underbrace{\begin{pmatrix} \sin \theta \cos \varphi \\ \sin \theta \sin \varphi \\ \cos \theta \end{pmatrix}}_{=: \hat{x}(\varphi, \theta)}, \quad \varphi \in [0, 2\pi), \theta \in [0, \pi), \quad (3.22)$$

womit wir nach Koordinatentransformation die folgende Darstellung der Helmholtz-Gleichung (in Kugelkoordinaten) erhalten:

$$-\frac{1}{a^2} \left(\partial_r^2 u + \frac{2}{r} \partial_r u + \frac{1}{r^2 \sin \theta} \partial_\theta (\sin \theta \cdot \partial_\theta u) + \frac{1}{r^2 (\sin \theta)^2} \partial_\varphi^2 u \right) - \omega^2 u = 0. \quad (3.23)$$

Aus der Differentialgeometrie ist bekannt, dass für Untermannigfaltigkeiten des \mathbb{R}^3 und deren Parametrisierungen $\hat{x}(\xi)$, der Laplace-Beltrami Operator durch $\Delta_{\hat{x}} := \text{div grad}$ gegeben ist. Die folgenden Überlegungen sollen lediglich deren Darstellung motivieren, für Details sei der Leser auf Standardwerke der Differentialgeometrie verwiesen. Um die Gradienten bzw. Divergenz-Operatoren auf (2-dimensionalen) Untermannigfaltigkeiten definieren zu können, benötigen wir noch den Begriff des metrischen Tensors, gegeben durch

$$g = (g_{ij}) \in \mathbb{R}^{2 \times 2}, \quad \text{wobei } g_{ij} = \frac{\partial \hat{x}}{\partial \xi_i} \cdot \frac{\partial \hat{x}}{\partial \xi_j},$$

sowie den inversen metrischen Tensor $g^{-1} = (g^{ij})$, definiert durch $g^{-1}g = \text{id}_{\mathbb{R}^2 \times \mathbb{R}^2}$. Dadurch lassen sich die zuvor erwähnten Operatoren wie folgt definieren:

$$\text{grad} = \begin{pmatrix} g^{11} \partial_{\xi_1} + g^{12} \partial_{\xi_2} \\ g^{21} \partial_{\xi_1} + g^{22} \partial_{\xi_2} \end{pmatrix}, \quad \text{div} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \frac{1}{\sqrt{\det g}} \left(\partial_{\xi_1} (\sqrt{\det g} \cdot v_1) + \partial_{\xi_2} (\sqrt{\det g} \cdot v_2) \right).$$

Mit der Parametrisierung (3.22) erhalten wir mit $\xi = (\varphi, \theta)^T$ den bekannten metrischen Tensor

$$g = \begin{pmatrix} a^2 \sin^2 \theta & 0 \\ 0 & a^2 \end{pmatrix}$$

und daraus mit $\sqrt{\det g} = a^2 \sin \theta$ ebenfalls

$$\text{grad}_{\hat{x}} = \frac{1}{a^2} \begin{pmatrix} \frac{1}{\sin^2 \theta} \cdot \partial_{\varphi} \\ \partial_{\theta} \end{pmatrix}.$$

Somit können wir mit einer weiteren Rechnung den Laplace Beltrami Operator ablesen:

$$\Delta_{\hat{x}} = \frac{1}{a^2} \left(\frac{1}{(\sin \theta)^2} \partial_{\varphi}^2 + \frac{1}{\sin \theta} \partial_{\theta} (\sin \theta \cdot \partial_{\theta}) \right). \quad (3.24)$$

Mit diesen Vorbereitungen vereinfacht sich die Gleichung in (3.23) zu:

$$-\frac{1}{a^2} \left(\partial_r^2 u + \frac{2}{r} \partial_r u \right) - \frac{1}{r^2} \Delta_{\hat{x}} u - \omega^2 u = 0. \quad (3.25)$$

Bevor wir mit der weiteren Analyse der Helmholtz Gleichung starten, führen wir noch ein wichtiges Hilfsresultat für den oben definierten Laplace-Beltrami Operator an:

Satz 3.13. *Mit den Voraussetzungen von zu Beginn des Abschnitts sei $\Gamma_a := \partial B_a$ und $\Delta_{\hat{x}}$ der in (3.24) definierte Laplace Beltrami Operator für die Kugeloberfläche. Dann gelten die folgenden Aussagen:*

- $\Delta_{\hat{x}}$ besitzt auf $L^2(\Gamma_a)$ ein vollständiges Orthonormalsystem $(\Phi_n)_{n \in \mathbb{N}_0}$ aus Eigenfunktionen mit dazugehörigen Eigenwerten λ_n
- Die Eigenwerte λ_n haben die Vielfachheit $M_n := 2n + 1$ und berechnen sich mit der folgenden Formel:

$$\lambda_n = -\frac{n(n+1)}{a^2}.$$

Für einen Beweis verweisen wir auf [HN09, Lemma 3.4].

Wir haben nun alle Vorbereitungen getroffen, um die Helmholtz Gleichung in Kugelkoordinaten (3.25) explizit lösen zu können. Dafür verwenden wir einen Separationsansatz und erhalten mit Hilfe von Satz 3.13 die Darstellung

$$u(r\hat{x}) = \sum_{n=0}^{\infty} \sum_{m=1}^{M_n} \underbrace{\left(\int_{\Gamma_a} u(r\hat{x}) \overline{\Phi_{nm}(\hat{x})} d\hat{x} \right)}_{=: u_{nm}(r)} \Phi_{nm}(\hat{x}) = \sum_{n=0}^{\infty} u_{nm}(r) \Phi_{nm}(\hat{x}). \quad (3.26)$$

Setzen wir nun Identität (3.26) in (3.25) ein, so erhalten wir nach vertauschen von Summation und Differentiation

$$\sum_{n=0}^{\infty} \sum_{m=1}^{M_n} \left(-\frac{u''_{nm}(r)}{a^2} - \frac{2}{ra^2} u'_{nm}(r) + \left(\frac{\lambda_n}{r^2} - \omega^2 \right) u_{nm}(r) \right) \Phi_{nm}(\hat{x}) = 0,$$

womit wir nach Multiplikation mit $-a^2 r^2$ die sogenannte *sphärische Bessel'sche Differentialgleichung* (für $a\omega = 1$) erhalten:

$$r^2 u''_{nm}(r) + 2ru'_{nm}(r) - (a^2 \omega^2 r^2 + n(n-1)) u_{nm}(r) = 0, \quad \forall n \in \mathbb{N}_0. \quad (3.27)$$

Die allgemeine Lösung der obigen Gleichung kann, wie schon kurz in Abschnitt 3.2.1 erwähnt, mit Hilfe der *sphärischen Hankel Funktionen* erster und zweiter Ordnung $\mathcal{H}_n^{(1,2)}$ geschrieben werden als

$$u_{nm}(r) = c \mathcal{H}_n^{(1)}(a\omega r) + d \mathcal{H}_n^{(2)}(a\omega r).$$

Die Funktionen $\mathcal{H}_n^{(1,2)}$ lassen sich über eine unendliche Reihe darstellen und bilden, wie man ebenfalls nachweisen kann, zwei linear unabhängige Lösungen der Gleichung 3.27. Für Details verweisen wir auf ([Ih198, Abschnitt 2.1.2]), wo ebenfalls nachgewiesen wird, dass folgendes asymptotisches Verhalten für $r \rightarrow \infty$ gilt:

$$\begin{aligned} \mathcal{H}_n^{(1)}(r) &\sim \frac{\exp(ir)}{r}, & (\mathcal{H}_n^{(1)}(r))' &\sim i \frac{\exp(ir)}{r} \\ \mathcal{H}_n^{(2)}(r) &\sim \frac{\exp(-ir)}{r}, & (\mathcal{H}_n^{(2)}(r))' &\sim -i \frac{\exp(-ir)}{r}. \end{aligned}$$

Durch Bilden des Grenzwertes

$$\lim_{r \rightarrow \infty} r \left((\mathcal{H}_n^{(1,2)}(r))' \pm i \mathcal{H}_n^{(1,2)}(r) \right) = 0$$

schließen wir, dass Funktionen $u(r\hat{x}) = a_{nm} \mathcal{H}_n^1(a\omega r) \Phi_{nm}(\hat{x})$ die Identität (3.14) erfüllen und somit *ausstrahlende* Lösungen der Helmholtz Gleichung sind. Entwickelt man die Randdaten $f \in L^2(\Gamma_a)$ ebenfalls in eine Fourierreihe der Form

$$f(\hat{x}) = \sum_{n=0}^{\infty} \sum_{m=1}^{M_n} \underbrace{\left(\int_{\Gamma_a} f(\hat{x}) \overline{\Phi_{nm}(\hat{x})} d\hat{x} \right)}_{=: f_{nm}} \Phi_{nm}(\hat{x}) = \sum_{n=0}^{\infty} \sum_{m=1}^{M_n} f_{nm} \Phi_{nm}(\hat{x}),$$

so lassen sich die Konstanten a_{nm} eindeutig bestimmen und wir können durch Koeffizientenvergleich die (sogar eindeutige) Lösung von (3.21) darstellen mittels:

$$u(r\hat{x}) = \sum_{n=0}^{\infty} \sum_{m=1}^{M_n} \frac{\int_{\Gamma_a} f(\hat{x}) \overline{\Phi_{nm}(\hat{x})} d\hat{x}}{\mathcal{H}_n^{(1)}(\omega a)} \mathcal{H}_n^{(1)}(\omega r) \Phi_{nm}(\hat{x}). \quad (3.28)$$

Mit etwas mehr Aufwand lässt sich zeigen, dass diese Reihe absolut und gleichmäßig auf kompakten Mengen $K \subset \Omega_{\text{ext}}$ konvergiert. Selbiges gilt für die Reihe der Ableitungen (komponentenweise) bezüglich r . Für einen Beweis verweisen wir abermals auf weiterführende Literatur, wie etwa [CK98, Theorem 2.14]. Somit können wir die Darstellung in (3.28) verwenden, um eine explizite Darstellung für den DtN Operator auf der Kugeloberfläche zu erhalten:

$$\mathcal{DtN}^B f = \frac{\partial u_{\text{ext}}}{\partial \eta} \Big|_{\partial B_a} \stackrel{(3.28)}{=} \sum_{n=0}^{\infty} \sum_{m=1}^{M_n} \frac{\int_{\Gamma_a} f(\hat{x}) \overline{\Phi_{nm}(\hat{x})} d\hat{x}}{\mathcal{H}_n^{(1)}(\omega a)} \omega \cdot \mathcal{H}_n^{(1)'}(\omega a) \Phi_{nm}. \quad (3.29)$$

Diese explizite Darstellung erweist sich als zentraler Bestandteil, um Eindeutigkeit von Lösungen auf dem Ganzraum \mathbb{R}^3 zu zeigen. Bevor wir zum Hauptresultat dieses Abschnitts kommen, benötigen wir noch das folgende Hilfslemma:

Lemma 3.14. Mit den Voraussetzungen von zu Beginn des Kapitels, sei α_n definiert durch:

$$\alpha_n := \omega \frac{\mathcal{H}_n^{(1)'}(\omega a)}{\mathcal{H}_n^{(1)}(\omega a)}. \quad (3.30)$$

Dann gilt für den Imaginärteil:

$$\Im(\alpha_n) > 0, \forall n \in \mathbb{N}_0, \quad \Im(\alpha_n) \rightarrow 0$$

sowie für den Realteil die folgende Ungleichung:

$$\max\left(\frac{1}{a}, \frac{n+1}{a} - \omega\right) \leq -\Re(\alpha_n) \leq \frac{n+1}{a} - \omega.$$

Ein Beweis dafür findet sich in [Ih198, Lemma 3.3].

Wir haben nun alle notwendigen Techniken und Resultate beisammen, um den Hauptsatz dieses Abschnitts formulieren und beweisen zu können:

Satz 3.15. Sei $\omega \in \mathbb{R}$ und $\Omega_{\text{int}} := B_a \setminus K$ wie zu Beginn des Kapitels definiert. Dann existiert für $\phi \in L^2(\partial K)$ eine eindeutige Lösung $u \in H^1(\Omega_{\text{int}})$, die

$$\int_{\Omega_{\text{int}}} \nabla u \cdot \overline{\nabla v} dx - \omega^2 \int_{\Omega_{\text{int}}} u \overline{v} dx - \int_{\partial B_a} \mathcal{DtN}^B(u|_{\partial B_a}) \overline{v|_{\partial B_a}} ds = \int_{\partial K} \phi \overline{v} ds \quad (3.31)$$

erfüllt für alle $v \in H^1(\Omega_{\text{int}})$.

Beweis. Der Beweis orientiert sich zu Teilen an [DI01] sowie [Néd01, Theorem 2.6.5] und zielt darauf ab, die Fredholm'sche Alternative anzuwenden, wonach die Eindeutigkeit von Lösungen gleichbedeutend mit der Existenz ist. Dafür setzen wir wie gewohnt den transparenten Rand als $\partial B_a =: \Gamma_a$ und definieren

$$s(u, v) := \int_{\Omega_{\text{int}}} \nabla u \cdot \overline{\nabla v} \, dx + \omega^2 \int_{\Omega_{\text{int}}} u \bar{v} \, dx - \int_{\Gamma_a} \mathcal{D}t\mathcal{N}^B(u|_{\Gamma_a}) \overline{v|_{\Gamma_a}} \, ds$$

sowie

$$k(u, v) := 2\omega^2 \int_{\Omega_{\text{int}}} u \bar{v} \, dx \quad \text{und} \quad f(v) := \int_{\partial K} \phi \bar{v} \, ds.$$

Mit Hilfe von (3.28), (3.30) und der Tatsache, dass Φ_{nm} eine Orthonormalbasis auf Γ_a bildet, vereinfachen sich die Randintegrale zu:

$$\int_{\Gamma_a} \mathcal{D}t\mathcal{N}^B(u|_{\Gamma_a}) \overline{v|_{\Gamma_a}} \, ds = \sum_{n=0}^{\infty} \alpha_n \sum_{m=0}^{M_n} (u|_{\Gamma_a}, \Phi_{nm})_{L^2(\Gamma_a)} \cdot \overline{(v|_{\Gamma_a}, \Phi_{nm})_{L^2(\Gamma_a)}}.$$

Mit Hilfe von Lemma 3.14 folgern wir somit, dass

$$-\Re \left(\int_{\Gamma_a} \mathcal{D}t\mathcal{N}^B(u|_{\Gamma_a}) \overline{u|_{\Gamma_a}} \, ds \right) \geq \frac{1}{a} \|u\|_{L^2(\Gamma_a)}^2$$

und somit $s(\cdot, \cdot)$ auf $H^1(\Omega_{\text{int}})$ koerziv ist, da nach Voraussetzung $\Re(\omega^2) > 0$ und somit gilt:

$$\begin{aligned} |s(u, v)| &\geq \Re(s(u, u)) = (\nabla u, \nabla u)_{L^2(\Omega_{\text{int}})} + \Re(\omega^2)(u, u)_{L^2(\Omega_{\text{int}})} - \Re \left(\int_{\Gamma_a} \mathcal{D}t\mathcal{N}^B(u|_{\Gamma_a}) \overline{u|_{\Gamma_a}} \, ds \right) \\ &\geq \min\{1, \Re(\omega^2)\} (\|\nabla u\|_{L^2(\Omega_{\text{int}})}^2 + \|u\|_{L^2(\Omega_{\text{int}})}^2) + \underbrace{\frac{1}{a} \|u\|_{L^2(\Gamma_a)}^2}_{\geq 0} \\ &\geq \min\{1, \Re(\omega^2)\} \|u\|_{H^1(\Omega_{\text{int}})}^2. \end{aligned}$$

Das Lemma von Lax-Milgram sichert uns nun die Bijektivität des durch s induzierten Operators $S : H^1(\Omega_{\text{int}}) \rightarrow H^{-1}(\Omega_{\text{int}})$ definiert durch $Su := s(u, \cdot)$. Durch die kompakte Einbettung $H^1(\Omega_{\text{int}}) \hookrightarrow L^2(\Omega_{\text{int}})$ folgt überdies, dass auch der durch k induzierte Operator K kompakt ist. Somit ist die Formulierung (3.31) äquivalent zu der Operatorgleichung (mit $Fv := f(v)$)

$$(S - K)u = F, \quad u \in H^1(\Omega_{\text{int}}).$$

Somit befinden wir uns gerade im Setting von Satz 3.11, wonach Injektivität bereits Bijektivität impliziert. Es bleibt somit lediglich die Eindeutigkeit von Lösungen von (3.31) zu zeigen. Angenommen $u_1 \neq u_2$ sind Lösungen von (3.31), dann gilt mit $w := u_1 - u_2$ einerseits $\partial_\eta w|_{\partial K} \equiv 0$ und andererseits (durch wählen von w als Testfunktion) die folgende Gleichung:

$$\int_{\Omega_{\text{int}}} |\nabla w|^2 \, dx - \omega^2 \int_{\Omega_{\text{int}}} |w|^2 \, dx = \int_{\Gamma_a} \mathcal{D}t\mathcal{N}^B(w|_{\Gamma_a}) \overline{w|_{\Gamma_a}} \, ds$$

Da die linke Seite reell ist, folgt mit Hilfe der Darstellung (3.29) sofort, dass ebenfalls

$$0 = \Im \left(\int_{\Gamma_a} \mathcal{D}t\mathcal{N}^B(w|_{\Gamma_a}) \overline{w|_{\Gamma_a}} ds \right) = \sum_{n=0}^{\infty} \Im(\alpha_n) \sum_{m=0}^{M_n} |(w|_{\Gamma_a}, \Phi_{nm})_{L^2(\Gamma_a)}|^2$$

gelten muss. Dies kann jedoch nach Satz 3.14 nur der Fall sein, wenn $(w|_{\Gamma_a}, \Phi_{nm})_{L^2(\Gamma_a)} = 0$ gilt $\forall m, n$. Daraus wiederum folgt aus der Darstellung von w und $\mathcal{D}t\mathcal{N}^B w|_{\Gamma}$ unmittelbar:

$$w = 0 \text{ auf } \Gamma_a, \quad \frac{\partial w}{\partial \eta} \Big|_{\Gamma_a} = 0.$$

Da überdies ∂K und somit $\partial\Omega_{\text{int}}$ als hinreichend glatt vorausgesetzt wurden, können wir mit klassischen Regularitätssätzen ebenfalls auf $w \in H^2(\Omega_{\text{int}})$ schließen. Somit erfüllt w aber gerade die Voraussetzungen von Satz 3.12 und wir erhalten, dass $w \equiv 0$ in Ω_{int} gilt. Dies beweist die Eindeutigkeit von Lösungen von (3.31) und somit die Behauptung. \square

Tatsächlich deckt die Wahl einer Kugeloberfläche als transparenten Rand bereits eine Vielzahl an Problemstellungen ab. Natürlich lässt sich obige Argumentation ebenfalls auf den zweidimensionalen Fall anwenden, siehe dazu [HN09, Satz 3.12]. Interessant wird es jedoch, falls es die zu untersuchende Geometrie nicht ermöglicht, einen Kreis als „künstlichen Rand“ zu wählen. Im folgenden Kapitel wollen wir ein einfaches Beispiel dazu anführen.

3.3.2 Eindeutige Lösbarkeit in Wellenleitern

In diesem Abschnitt werden wir Streuprobleme in Wellenleitern diskutieren, da diese ebenfalls Gegenstand der numerischen Simulationen in den folgenden Kapiteln sein werden. Ausgangspunkt dafür ist der *halboffene Wellenleiter* $W := (0, \infty) \times (0, c)$ mit $c > 0$ sowie eine kompakte Menge $K \subset W$, welcher aus physikalischer Sicht den Streukörper darstellt. Als zu Grunde liegender Geometrie erhalten wir somit $\Omega := W \setminus K$, siehe auch Abbildung 3.3. Für den Rest des Abschnitts beschäftigen wir uns also mit der Lösbarkeit der



Abbildung 3.3: Wellenleiter mit Streukörper

Helmholtz Gleichung auf Ω . Um die nachfolgende Präsentation zu vereinfachen, setzen wir überdies $\Gamma := [0, \pi]$, $\Gamma_D := \{0\} \times \Gamma$ und $\Gamma_N := \partial\Omega \setminus \Gamma_D$, wodurch sich das Modellproblem wie folgt schreiben lässt:

$$\begin{aligned} -\Delta u - \omega^2 u &= 0 & \text{in } \Omega \\ u &= u_0 & \text{auf } \Gamma_D \\ \frac{\partial u}{\partial \eta} &= 0 & \text{auf } \Gamma_N. \end{aligned} \tag{3.32}$$

Da wir vor allem an schwachen Lösungen interessiert sind, setzen wir die Randdaten $u_0 \in H^{1/2}(\Gamma_D)$ voraus.

Bemerkung 3.16. Die Ausstrahlungsbedingung für den Wellenleiter kann zunächst nicht ohne weiteres aus (3.14) übernommen werden, da wir in der dortigen Herleitung Probleme auf dem Ganzraum \mathbb{R}^d betrachtet haben. Tatsächlich kann man (sogar für allgemeine elliptische Differentialoperatoren) für Lösungen in Wellenleitern eine Ausstrahlungsbedingung formulieren, sodass das Problem (3.32) eindeutig lösbar wird. Für Details verweisen wir abermals auf weiterführende Literatur wie [Hoa11b] oder [SCH22].

Die folgenden Überlegungen basieren auf dem sogenannten *Limiting Absorbion Principle*. Dieses besteht darin eine Lösung von (3.32) zu konstruieren, indem zunächst für $\delta > 0$ und den selben Voraussetzungen wie zu Beginn des Abschnitts das gedämpfte Problem

$$-\Delta u - (\omega^2 + i\delta)u = 0 \quad \text{in } \Omega \tag{3.33}$$

ermittelt wird. Der Parameter $i\delta$ spiegelt dabei, wie wir später sehen werden, eine Dämpfung im Lösungsverhalten wieder. Existieren für δ hinreichend klein (eindeutige) Lösungen u_δ von (3.33), so kann man versuchen den Grenzwert $\lim_{\delta \rightarrow 0} u_\delta =: \hat{u}$ zu betrachten. Dieser Grenzwert (sofern existent) beschreibt nach dem Limiting-Absorbion Prinzip die gesuchte und physikalisch relevante Lösung von (3.32).

Analog zum vorherigen Abschnitt zerlegen wir zunächst das Gebiet Ω in zwei Teilgebiete Ω_{int} und Ω_{ext} und versuchen eine analytische Lösung auf Ω_{ext} zu ermitteln. Durch die

Kompaktheit von K existiert jedenfalls ein $a > 0$, sodass $(a, \infty) \times (0, \pi) \cap K = \emptyset$ gilt, siehe auch Abbildung 3.3. Somit können wir mit $\Omega_{\text{ext}} := (a, \infty) \times (0, \pi)$ und $\Omega_{\text{int}} := ((0, a) \times (0, \pi)) \setminus K$ die zu Grunde liegende Geometrie schreiben als $\Omega = \overline{\Omega_{\text{int}} \cup \Omega_{\text{ext}}}$. Das dazugehörige Außenraumproblem mit transparenten Rand $\Gamma_a := \{a\} \times [0, \pi]$ lautet demnach wie folgt:

$$\begin{aligned} -\Delta u_{\text{ext}} - \omega^2 u_{\text{ext}} &= 0 & \text{in } \Omega_{\text{ext}} \\ u_{\text{ext}} &= f & \text{auf } \Gamma_a \\ \frac{\partial u_{\text{ext}}}{\partial \eta} &= 0 & \text{auf } \partial\Omega_{\text{ext}} \setminus \Gamma_a. \end{aligned} \tag{3.34}$$

Für dieses Problem können wir nun wieder eine explizite Lösung mit Hilfe des Separationsansatz konstruieren. Parametrisieren wir nämlich den transparenten Rand Γ_a durch $\hat{x}(y) = (a, y)^T$, so erhalten wir analog zu den Überlegungen in Kapitel 3.3.1 den negativen Laplace Beltrami Operator $-\Delta_{\hat{x}}$ auf Γ_a durch ∂_y^2 . Dies bedeutet insbesondere, dass $-\Delta$ auf Ω bereits separiert und keine weitere Koordinatentransformation mehr verwendet werden muss. Eine Orthonormalbasis aus Eigenfunktionen von $-\Delta_{\hat{x}}$ (Neumann Rand) auf $L^2(\Gamma_a)$ bzw. $L^2(0, \pi)$ können wir überdies explizit durch

$$\Phi_n(y) := \sqrt{\frac{2}{\pi}} \cos(ny)$$

und den dazugehörigen Eigenwerten $\lambda_n := n^2$ mit $n \in \mathbb{N}_0$ angeben. Lösungen von (3.34) können demnach in eine Fourier Reihe entwickelt werden:

$$u(x, y) = \sum_{n=0}^{\infty} \underbrace{\left(\int_{\Gamma_a} u(x, y) \Phi_n(y) dy \right)}_{=: u_n(x)} \Phi_n(y). \tag{3.35}$$

Durch Einsetzen von (3.35) in 3.34 erhalten wir für $\omega \notin \mathbb{Z}$ die folgende Gleichung:

$$u_n''(x) + (\omega^2 - n^2)u_n(x) = 0, \quad \forall n \in \mathbb{N}_0. \tag{3.36}$$

Als gewöhnliche Differentialgleichung zweiter Ordnung erhalten wir Lösungen der folgende Darstellung:

$$u_n(x) = a_n \exp(i\sqrt{\omega^2 - n^2} \cdot x) + b_n \exp(-i\sqrt{\omega^2 - n^2} \cdot x). \tag{3.37}$$

Wir müssen nun die folgenden Fälle separat betrachten:

- $n < \omega$:

Da der Ausdruck unter der Wurzel positiv ist, spiegeln die Lösungen aus (3.37) zwei „klassische Schwingungen“ wieder. Um eine der beiden Lösungen zu eliminieren, bedienen wir uns im folgenden wie bereits angesprochen dem Limiting-Absorption Prinzip. Dieses besteht im wesentlichen daraus der „pur schwingenden“ Lösungen eine künstliche Dämpfung hinzuzufügen. Mathematisch können wir dies fassen, indem wir

den Übergang ($\omega^2 \mapsto \omega^2 + i\delta$) für $\delta > 0$ durchführen. Ähnlich wie zuvor erhalten wir somit die folgenden allgemeinen Lösungen des „gedämpften Problems“:

$$u_n^\delta(x) = a_n^\delta \exp(i\sqrt{\omega^2 - n^2 + i\delta} \cdot x) + b_n^\delta \exp(-i\sqrt{\omega^2 - n^2 + i\delta} \cdot x). \quad (3.38)$$

Da wir an einer physikalisch sinnvollen Lösung interessiert sind, liegt es nahe zu fordern, dass Lösungen von (3.33) bzw. (3.32) zumindest beschränkt sein sollen, also $u_n^\delta \in L^\infty(\Omega)$ gelten soll. Nun gilt für allgemeine komplexe Zahlen $z = u + iw$ die folgende Identität

$$|e^{iz}| = |e^{iu-w}| = |e^{iu}e^{-w}| = e^{-\Im z}. \quad (3.39)$$

Da jedoch $\delta > 0$ vorausgesetzt wurde, können wir nun einerseits sehen (wie man elementar mit Hilfe von Polarkoordinaten nachrechnen kann), dass

$$\Im(\sqrt{\omega^2 - n^2 + i\delta}) > 0$$

und andererseits, dass deshalb $b_n^\delta = 0$ gelten muss, da sonst mit (3.39) folgen würde, dass

$$\lim_{x \rightarrow \infty} |b_n^\delta| \cdot e^{\Im z \cdot x} = \infty$$

und somit u_n^δ ebenfalls unbeschränkt ist. Somit ist klar, welche der beiden Lösungen in (3.38) physikalisch relevant ist. Führen wir den Grenzübergang $\delta \rightarrow 0$ durch, dann ist nun ebenfalls geklärt, welche der beiden Fundamentallösungen in (3.37) zu nehmen ist.

- $n > \omega$:

In diesem Fall müssen wir die Wurzel einer negativen reellen Zahl bilden, womit wir klarerweise zunächst die komplexe Wurzel definieren müssen:

$$\sqrt[\mathcal{C}]{\cdot} : \mathbb{C} \setminus \mathbb{R}_{<0} \rightarrow \mathbb{C}, z \mapsto \sqrt[\mathcal{C}]{z} := \sqrt{|z|} \exp\left(\frac{i \arg z}{2}\right). \quad (3.40)$$

Dabei bezeichnet $\arg z$ den Winkel (Polarkoordinaten!), für den wir den folgenden Definitionsbereich wählen:

$$-\pi < \arg z \leq \pi.$$

Falls nicht anders angeführt, verwenden wir stets diese Definition der komplexen Wurzel, für die wir im folgenden $\sqrt{\cdot}$ schreiben. Mit jener Wahl können wir feststellen, dass

$$\sqrt{\omega^2 - n^2} = i\sqrt{n^2 - \omega^2}$$

gilt. Somit ist ebenfalls klar, dass $b_n = 0$ in (3.37) gelten muss, sofern wir wieder $u \in L^\infty(\Omega)$ verlangen.

- $n = \omega$:

Wir haben zwar bisher $\omega \notin \mathbb{Z}$ vorausgesetzt, jedoch kann man auch für ganzzahlige Werte eine Ausstrahlungsbedingung festlegen. Zunächst reduziert sich die Gleichung in (3.36) zu

$$u_n''(x) = 0,$$

deren Lösungen klarerweise durch

$$u_n(x) = a + bx$$

für Konstanten $a, b \in \mathbb{R}$ gegeben sind. Beschränkt bleibt die Lösung klarerweise für $b = 0$, wobei die gewünschte Lösung $u \equiv a$ entlang des Welleneiters ist.

Zusammengefasst erhalten wir für das Problem (3.34) also Lösungen der Form

$$u_n(x, y) = a_n \cdot \exp(i\sqrt{\omega^2 - n^2} \cdot x) \Phi_n(y).$$

Zur Bestimmung der Konstanten a_n , entwickeln wir die Randdaten f ebenfalls in eine Fourier-Reihe

$$f = \sum_{n=0}^{\infty} (f, \Phi_n)_{L^2(\Gamma_a)} \Phi_n$$

und erhalten nach erneutem Koeffizientenvergleich die gesuchte *ausstrahlende Lösung* des Außenraumproblems (3.34):

$$u(x, y) = \sum_{n=0}^{\infty} \exp(i\sqrt{\omega^2 - n^2} \cdot (x - a)) \cdot (f, \Phi_n)_{L^2(\Gamma_a)} \Phi_n(y). \quad (3.41)$$

Bemerkung 3.17. Das Limiting-Absorption Prinzip sorgt also dafür, dass eine der beiden Fundamentallösungen in (3.37) eliminiert wird. Vergleicht man diese Auswahl mit der Bedingung in (3.14), so haben wir uns in der (durch Separation entstandenen) eindimensionalen Gleichung (3.37) für jene Lösung entschieden, die

$$u_n'(x) - i\sqrt{\omega^2 - n^2} \cdot u_n(x) = 0$$

erfüllt. Diese Wahl entspricht einer sogenannten *ausstrahlenden Lösung*, während die „zweite Fundamentallösung“ die obige Gleichung mit positiven Vorzeichen erfüllt. Man spricht in diesem Fall von *einfallenden Lösungen*. Typischerweise verwendet man die anhand von $\omega \in \mathbb{R}$ oben definierten Ausstrahlungsbedingungen ebenfalls für komplexe $\omega \in \mathbb{C}$. Hier muss man in jedem Fall die genaue Definition der komplexen Wurzel berücksichtigen, die leider nicht einheitlich ist und von der jeweiligen Literatur abhängt. Allgemein lässt sich eine beliebige komplexe Wurzel mit Polarkoordinaten schreiben als

$$\sqrt[\theta]{z} : \mathbb{C} \setminus (\cos \theta + i \sin \theta) \mathbb{R}_{>0} \rightarrow \mathbb{C}, z \mapsto \sqrt[\theta]{z} := \sqrt{|z|} \exp\left(\frac{i \arg_{\theta} z}{2}\right). \quad (3.42)$$

Dabei bestimmt θ die Wahl des Winkel $\arg_\theta z$, bzw. genauer gesagt den Definitionsbereich

$$\theta - 2\pi < \arg_\theta z \leq \theta.$$

Hat man allgemein für $\theta \in [0, 2\pi)$ eine komplexe Wurzel gewählt, kann man zunächst wieder mit reellen Frequenzen ω starten und dann analog zu oben jeweils eine der beiden in (3.37) auftretenden Lösungen eliminieren, indem man Beschränktheit fordert.

Aus der Darstellung in (3.41) können wir nun den Dirichlet to Neumann Operator ablesen, welcher für $f \in H^{1/2}(\Gamma_a)$ über die eindeutige Lösung u von (3.34) wohldefiniert ist:

$$\mathcal{DtN}^W f = \frac{\partial u}{\partial y} \Big|_{x=a} = \sum_{n=0}^{\infty} i\sqrt{\omega^2 - n^2} \cdot (f, \Phi_n)_{L^2(\Gamma_a)} \Phi_n \quad (3.43)$$

Mit diesen Vorbereitungen können wir letztlich zur schwachen Formulierung von (3.32) auf Ω_{int} übergehen. Diese lautet wie folgt:

Suche $u \in H^1(\Omega_{\text{int}})$ mit $u|_{\Gamma_D} = u_0$, sodass gilt:

$$\int_{\Omega_{\text{int}}} \nabla u \overline{\nabla v} - \omega^2 u \overline{v} \, dx - \int_{\Gamma_a} \mathcal{DtN}^W(u|_{\Gamma_a}) \overline{v|_{\Gamma_a}} \, ds = 0, \quad \forall v \in H^1(\Omega_{\text{int}}). \quad (3.44)$$

Bemerkung 3.18. Die Randbedingung $u|_{\Gamma_D} = u_0$ wird in der Praxis über eine sogenannte Homogenisierung der Randdaten behandelt. Dazu sei \hat{u}_0 eine beliebige Fortsetzung aus $H^1(\Omega_{\text{int}})$ von u_0 mit $\hat{u}_0|_{\Gamma_D} = u_0$. Setzen wir nun $w := u - \hat{u}_0$, dann erfüllt $w \in H^1(\Omega_{\text{int}})$ einerseits $w|_{\Gamma_D} = 0$ und andererseits die folgende Gleichung:

$$\int_{\Omega_{\text{int}}} \nabla w \overline{\nabla v} - \omega^2 w \overline{v} \, dx - \int_{\Gamma_a} \mathcal{DtN}^W(w|_{\Gamma_a}) \overline{v|_{\Gamma_a}} \, ds = \int_{\Omega_{\text{int}}} \hat{u}_0 \overline{v} \, dx, \quad \forall v \in H^1(\Omega_{\text{int}}^{\Gamma_D}) \quad (3.45)$$

Kennen wir eine Lösung von (3.45), so erhalten wir die Lösung des ursprünglichen Problems (3.44) durch Setzen von $u := w + \hat{u}_0$.

Allgemein erweist die Theorie über die eindeutige Lösbarkeit von (3.44) bzw (3.45) für beliebige Ω und Streukörper $K \subset \mathbb{R}^2$ als durchaus komplex. Wir wollen aber trotzdem für den Fall $K = \emptyset$ bzw $\Omega := W$ (dies entspricht einem homogenen Wellenleiter ohne Streukörper) die Wohldefiniertheit von Problem (3.44) für $\omega \in \mathbb{R} \setminus \mathbb{Z}$ zeigen. Dafür brauchen wir zunächst das folgende Hilfsresultat:

Satz 3.19. Sei $\Omega := (0, \infty) \times (0, \pi)$ und $\Gamma_D := \{0\} \times [0, \pi]$ gegeben. Weiters seien die Randdaten $u_0 \in H^{1/2}(\Gamma_D)$, sowie $f \in H^{-1}(\Omega)$ mit kompaktem Träger und $\omega > 0$ beliebig. Dann existiert eine Lösung $u \in H_{\text{loc}}^1(\Omega)$ des folgenden Problems:

$$\begin{aligned} -\Delta u - \omega u &= f && \text{in } \Omega \\ u &= u_0 && \text{auf } \Gamma_D, \\ \frac{\partial u}{\partial \eta} &= 0 && \text{auf } \partial\Omega \setminus \Gamma_D. \end{aligned} \quad (3.46)$$

Überdies erfüllt die Lösung mit $\Omega_m := (m, m+1) \times (0, \pi)$ die folgende Abschätzung:

$$\sup_{m \in \mathbb{N}} \|u\|_{H^1(\Omega_m)} < \infty.$$

Ein Beweis für obige Behauptung befindet sich in sehr abstrakter Form (sogar für allgemeine elliptische Differentialoperatoren mit periodischen Koeffizienten) in [Hoa11a] bzw. ebenfalls in [Gol82, Kapitel 2]. Bei genauerem Betrachten der Literatur, stellten wir fest, dass die Existenz für Lösungen von (3.46) nur für sogenannte *reguläre Frequenzen* gilt. Dabei bezeichnen wir ω als nicht regulär, also singulär, wenn ω im Punktspektrum von $-\Delta$ auf $L^2(\Omega)$ (mit Neumann bzw Dirichlet Randbedingungen) liegt (siehe dazu auch [BG99]). Bekannterweise ist jedoch das Punktspektrum für einen homogenen Wellenleiter W die leere Menge, wodurch der Satz insbesondere für alle $\omega > 0$ anwendbar ist. Wir sind nun in der Lage den folgenden Existenzsatz zu formulieren und zu beweisen:

Satz 3.20 (Existenz und Eindeutigkeit). *Mit den Voraussetzungen von zu Beginn des Kapitels sei $\omega \in \mathbb{R} \setminus \mathbb{Z}$ und $u_0 \in H^{1/2}(\Gamma_D)$ beliebig. Dann existiert für Problem (3.44) mit $\Omega := W$ bzw $\Omega_{\text{int}} := (0, a) \times (0, \pi)$ eine eindeutige Lösung $u \in H^1(\Omega_{\text{int}})$.*

Beweis. Ähnlich wie im Beweis in Kapitel 3.3.1 wollen wir die Fredholm'sche Alternative anwenden, wonach Existenz von Lösungen zur Eindeutigkeit äquivalent ist. Aus der Darstellung des Dirichlet-to-Neumann Operators (3.43) erhalten wir mit $M := \lceil \omega \rceil$

$$\begin{aligned} \Re\left(\int_{\Gamma_a} \mathcal{DtN}^W(u|_{\Gamma_a})\overline{u|_{\Gamma_a}} ds\right) &= \Re\left(\sum_{n=0}^{\infty} i\sqrt{\omega^2 - n^2} \cdot |(u|_{\Gamma_a}, \Phi_n)_{L^2(\Gamma_a)}|^2 \|\Phi_n\|_{L^2(\Gamma_a)}^2\right) \\ &= -\sum_{n=M}^{\infty} \sqrt{n^2 - \omega^2} \cdot |(u|_{\Gamma_a}, \Phi_n)_{L^2(\Gamma_a)}|^2 \\ &\leq -\sqrt{M^2 - \omega^2} \underbrace{\sum_{n=M}^{\infty} |(u|_{\Gamma_a}, \Phi_n)_{L^2(\Gamma_a)}|^2}_{< \infty}. \end{aligned}$$

Damit können wir den Term in der schwachen Formulierung wie folgt abschätzen:

$$-\Re\left(\int_{\Gamma_a} \mathcal{DtN}^W(u|_{\Gamma_a})\overline{v|_{\Gamma_a}} ds\right) > 0, \quad \forall u, v \in H^1(\Omega_{\text{int}}).$$

Wir gehen nun wie in Bemerkung 3.18 erwähnt zur äquivalenten schwachen Formulierung (3.45) über und definieren die Bilinearform $s(., .)$ durch

$$s(u, v) := \int_{\Omega_{\text{int}}} \nabla u \overline{\nabla v} + \omega^2 u \overline{v} dx - \int_{\Gamma_a} \mathcal{DtN}^W(u|_{\Gamma_a})\overline{v|_{\Gamma_a}} ds$$

auf $H^1(\Omega_{\text{int}})$, welche nach dem gerade gezeigten die folgende Ungleichung erfüllt:

$$|s(u, u)| \geq \Re(s(u, v)) \geq \min\{1, \Re(\omega^2)\} \|u\|_{H^1(\Omega_{\text{int}})}^2.$$

Dies entspricht genau der Koerzivität von $s(., .)$ auf $H^1(\Omega_{\text{int}})$. Analog zum vorherigen Kapitel definieren wir nun

$$k(u, v) := 2\omega^2 \int_{\Omega_{\text{int}}} u \overline{v} dx \quad \text{sowie} \quad f(v) := \int_{\Omega_{\text{int}}} \hat{u}_0 \overline{v} dx$$

und sehen, dass (3.45) mit den durch $s(\cdot, \cdot)$, $k(\cdot, \cdot)$ und $f(\cdot)$ induzierten Operatoren wie folgt lautet:

$$(S - K)u = F, u \in H^1(\Omega_{\text{int}}).$$

Dabei ist nun abermals S bijektiv und K kompakt, womit es nach der Fredholm'schen Alternative nun reicht, die Eindeutigkeit von Lösungen zu zeigen. Um diese zu beweisen, verwenden wir im Wesentlichen eine Idee aus [Hoa11a, Kapitel 4]. Seien also u_1, u_2 Lösungen von (3.45), dann erfüllt die Differenz $w := u_1 - u_2$ einerseits $w|_{\Gamma_D} = 0$ und andererseits die folgende Integralgleichung:

$$\int_{\Omega_{\text{int}}} |\nabla w|^2 - \omega^2 |w|^2 dx = \int_{\Gamma_a} \mathcal{DtN}^W(w|_{\Gamma_a}) \overline{w|_{\Gamma_a}} ds, \quad \forall w \in H^1(\Omega_{\text{int}}).$$

Daraus folgt unmittelbar

$$\Im \left(\int_{\Gamma_a} \mathcal{DtN}^W(w|_{\Gamma_a}) \overline{w|_{\Gamma_a}} ds \right) = 0.$$

Leider können wir daraus nicht wie im letzten Kapitel unmittelbar auf $w|_{\Gamma_a}, \partial_\eta w|_{\Gamma_a} \equiv 0$ schließen, wie man sich mit der Darstellung aus (3.43) leicht überlegt. Es gilt jedoch

$$\Im \left(\int_{\Gamma_a} \mathcal{DtN}^W(w|_{\Gamma_a}) \overline{w|_{\Gamma_a}} ds \right) = \sum_{n=0}^M |(w|_{\Gamma_a}, \Phi_n)_{L^2(\Gamma_a)}|^2 \underbrace{\sqrt{\omega^2 - n^2}}_{=1} \|\Phi_n\|_{L^2(\Gamma_a)}^2$$

und wir können folgern $(w|_{\Gamma_a}, \Phi_n)_{L^2(\Gamma_a)} = 0$ für alle $n \leq M$. Das Ziel besteht nun darin ebenfalls die restlichen Terme zu eliminieren. Dazu setzen wir nun $w_{\text{int}} := w(x, y)$ auf Ω_{ext} fort, indem wir die Lösung auf Ω_{ext} in eine Fourier-Reihe entwickeln. Wir erhalten somit eine Fortsetzung

$$z(x, y) = \begin{cases} w_{\text{int}}(x, y), & \text{in } \Omega_{\text{int}} \\ w_{\text{ext}}(x, y) := \sum_{n=M+1}^{\infty} e^{i\sqrt{\omega^2 - n^2} \cdot (x-a)} (w_{\text{int}}|_{\Gamma_a}, \Phi_n)_{L^2(\Gamma_a)} \Phi_n(y), & \text{in } \Omega_{\text{ext}} \end{cases}$$

Die Fortsetzung w ist so gewählt, dass einerseits $w_{\text{int}}|_{\Gamma_a} = w_{\text{ext}}|_{\Gamma_a}$ gilt und andererseits (wie man nach partieller Integration und der Darstellung der allgemeinen Lösung von (3.34) sieht) für kompaktes $K \subset \Omega$ die Fortsetzung $z \in H^1(K)$ Lösung von

$$\int_K \nabla z \overline{\nabla v} - \omega^2 z \overline{v} dx = \int_{\partial K} \frac{\partial z}{\partial \eta} \overline{v} ds \quad \forall v \in H^1(K) \quad (3.47)$$

ist. Sei nun eine kompakte Menge $K \subset \Omega$ gegeben. Weiters sei $v \in H_{\text{loc}}^1(\Omega)$ eine Lösung des Problems (diese existiert wegen Satz 3.19):

$$\begin{aligned} -\Delta v - \omega^2 v &= \mathbb{1}_K z & \text{in } \Omega \\ v &= 0 & \text{auf } \Gamma_D, \\ \frac{\partial v}{\partial \eta} &= 0 & \text{auf } \partial\Omega \setminus \Gamma_D \end{aligned}$$

In Integralschreibweise bedeutet dies für beschränktes $B \supset K$ mit Hilfe partieller Integration gerade

$$\int_B \nabla v \overline{\nabla u} - \omega^2 v \overline{u} \, dx - \int_{\partial B} \frac{\partial v}{\partial \eta} \overline{u} \, ds = \int_K z \overline{u} \, dx \quad \forall u \in H^1(B).$$

Setzen wir nun $\Omega_r := (0, r) \times (0, \pi)$ für $r > 0$ hinreichend groß, dann folgt mit dem gerade gezeigten

$$\begin{aligned} \int_K |z|^2 \, dx &= \int_K z (\overline{\mathbb{1}_K z}) \, dx \\ &= \int_{\Omega_r} \nabla v \overline{\nabla z} \, dx - \int_{\Omega_r} \omega^2 v \overline{z} \, dx - \int_{\partial \Omega_r} \frac{\partial v}{\partial \eta} \overline{z} \, ds. \end{aligned}$$

Verwenden wir nun ebenfalls (3.47), dann können wir den mittleren Term wie folgt schreiben:

$$- \int_{\Omega_r} \omega^2 \overline{z} v \, dx = \int_{\partial \Omega_r} \frac{\partial \overline{z}}{\partial \eta} v \, ds - \int_{\Omega_r} \overline{\nabla z} \nabla v \, dx.$$

Einsetzen in die vorherige Gleichung ergibt nun mit $\Gamma_r := \{r\} \times [0, \pi]$ die Identität:

$$\int_K |z|^2 \, dx \stackrel{(1)}{=} \int_{\Gamma_D} \frac{\partial z}{\partial \eta} \overline{v} - z \frac{\partial \overline{v}}{\partial \eta} \, ds - \int_{\Gamma_r} \frac{\partial z}{\partial \eta} \overline{v} - z \frac{\partial \overline{v}}{\partial \eta} \, ds.$$

Dabei gilt die Gleichheit (1), da die Randintegrale über $\partial \Omega \setminus \Gamma_D$ verschwinden. Da nun aber auch $z, v \equiv 0$ auf Γ_D gilt, verschwindet ebenfalls der erste Term. Für $r \rightarrow \infty$ geht überdies mit Hilfe der Darstellung der Fortsetzung für $(x, y) \in \Omega_{\text{ext}}$,

$$z(x, y) = \sum_{n=M+1}^{\infty} e^{-\sqrt{n^2 - \kappa^2} \cdot (x-a)} (w|_{\Gamma_a}, \Phi_n)_{L^2(\Gamma_a)} \Phi_n(y)$$

auch z und $\frac{\partial z}{\partial x}$ für $x \rightarrow \infty$ gegen 0, womit zusammen mit $\sup_{m \in \mathbb{N}} \|v\|_{H^1(\Omega_m)} < \infty$ insgesamt für beliebiges kompaktes $K \subset \Omega$ gilt

$$\int_K |z|^2 = 0 \Rightarrow z \equiv 0 \Rightarrow u_1 = u_2.$$

Dies beweist die Eindeutige Lösbarkeit von (3.44) und somit die Behauptung. □

3.4 Konvergenz von diskreten Methoden

Nachdem wir in den vergangenen Abschnitten die eindeutige Lösbarkeit von schwachen Lösungen des kontinuierlichen Problem gezeigt haben, werden wir in diesem Kapitel versuchen, jene Problemstellung in ein diskretes Setting zu bringen und erste Einblicke in die

Konvergenz von diskreten Verfahren zu geben. Dazu starten wir abermals mit der schwachen Formulierung des Modellproblems (3.15), dessen eindeutige Lösung $u \in H^1(\Omega_{\text{int}})$ die nachfolgende Gleichung für alle $v \in H^1(\Omega_{\text{int}})$ erfüllt:

$$\int_{\Omega_{\text{int}}} \nabla u \cdot \overline{\nabla v} \, dx - \omega^2 \int_{\Omega_{\text{int}}} u \overline{v} \, dx - \langle \mathcal{DtN}(u|_{\Gamma_a}), v|_{\Gamma_a} \rangle = \int_{\partial K} \phi \overline{v} \, ds. \quad (3.48)$$

Wie üblich setzen wir für ein diskretes (FEM) Verfahren $u_h \in V_h \subset H^1(\Omega_{\text{int}})$ an, wobei jedenfalls $\dim(V_h) < \infty$ gelten soll. Für eine mögliche Implementierung erweist sich abermals der Dirichlet-to-Neumann Operator als problematisch, da wir aus Definition 3.6 für allgemeine Probleme keine explizite Darstellung erhalten und somit der Ausdruck $\langle \mathcal{DtN}(u|_{\Gamma_a}), v|_{\Gamma_a} \rangle$ keine berechenbare Größe ist. Selbst die in den vergangenen Kapiteln behandelten Spezialfälle (3.29) und (3.43) schließen auf Grund einer unendlichen Reihendarstellung eine exakte Implementierung aus, wonach wir jedenfalls für ein diskretes Verfahren eine Approximation

$$\text{DtN}^{(h)} \approx \mathcal{DtN}$$

verwenden müssen. Wir klären an dieser Stelle nicht wie diese Approximation aussehen soll, da dies ohnehin Gegenstand des nächsten Kapitels sein wird. Somit ergibt sich mit den vorausgehenden Überlegungen das diskrete Problem wie folgt:

Suche $u_h \in V_h$, sodass

$$\int_{\Omega_{\text{int}}} \nabla u_h \cdot \overline{\nabla v_h} \, dx - \omega^2 \int_{\Omega_{\text{int}}} u_h \overline{v_h} \, dx - \langle \text{DtN}^{(h)}(u_h|_{\Gamma_a}), v_h|_{\Gamma_a} \rangle = \int_{\partial K} \phi \overline{v_h} \, ds \quad (3.49)$$

gilt für alle $v_h \in V_h$.

Für die Konvergenzanalyse von „gestörten Problemen“ verwenden wir das folgende Lemma als Ausgangspunkt:

Lemma 3.21 (Strang Lemma). Sei V Banachraum und $a : V \times V \rightarrow \mathbb{C}$ eine stetige Sesquilinearform, sowie $f : V \rightarrow \mathbb{C}$ stetig. Weiters bezeichne u die eindeutige Lösung des Variationsproblems:

$$a(u, v) = f(v), \quad \forall v \in V. \quad (3.50)$$

Dies bedeutet nach Satz 3.8 insbesondere, dass $a(\cdot, \cdot)$ die *inf-sup* Bedingung erfüllt

$$\inf_{0 \neq v \in V} \sup_{0 \neq w \in V} \frac{|a(v, w)|}{\|v\|_V \|w\|_V} =: \alpha > 0.$$

Weiters seien $(V_h)_{h>0}$ Unterräume $V_h \subset V$ und die dazugehörigen Sesquilinearformen $a_h : V_h \times V_h \rightarrow \mathbb{C}$ erfüllen ebenfalls gleichmäßig die *diskrete inf-sup* Bedingung, also

$$\exists h_0 > 0 : \forall h \leq h_0 : \inf_{0 \neq v_h \in V_h} \sup_{0 \neq w_h \in V_h} \frac{|a_h(v_h, w_h)|}{\|v_h\|_V \|w_h\|_V} =: \alpha_h \geq \tilde{\alpha} > 0. \quad (3.51)$$

Dann gilt für stetige $f_h : V_h \rightarrow \mathbb{C}$ und Lösungen u_h der dazugehörigen variationellen Problemen

$$a_h(u_h, v_h) = f_h(v_h) \quad \forall v_h \in V_h \quad (3.52)$$

die folgende Fehlerabschätzung:

$$\|u - u_h\|_V \leq C \left[\inf_{v_h \in V_h} \left(\|u - v_h\|_V + \sup_{w_h \in V_h} \frac{|a(v_h, w_h) - a_h(v_h, w_h)|}{\|w_h\|_V} \right) + \sup_{w_h \in V_h} \frac{|f(w) - f_h(w_h)|}{\|w_h\|_V} \right],$$

mit $\tilde{C} > 0$ unabhängig von h .

Beweis. Für einen Beweis orientieren wir uns hauptsächlich an [DI01, Abschnitt 2.4]. Wir stellen zunächst fest, dass für festes $v_h \in V_h$ mit Hilfe der Dreiecksungleichung gilt:

$$\|u - u_h\|_V \leq \|u - v_h\|_V + \|u_h - v_h\|_V. \quad (3.53)$$

Wir betrachten nun den zweiten Term separat und erhalten nach der Voraussetzung (3.51) für hinreichend kleine $h > 0$

$$\begin{aligned} \tilde{\alpha} \|u_h - v_h\| &\leq \sup_{w_h \in V_h} \frac{|a_h(u_h - v_h, w_h)|}{\|w_h\|_V} \\ &= \sup_{w_h \in V_h} \frac{|a_h(u_h - v_h, w_h) \pm a(u - v_h, w_h)|}{\|w_h\|_V}. \end{aligned}$$

Da nun aber u bzw. u_h Lösungen von (3.50) bzw (3.52) sind, erhalten wir mit Hilfe der Dreiecksungleichung die folgende Identität:

$$\tilde{\alpha} \|u_h - v_h\|_V \leq \sup_{w_h \in V_h} \left[\frac{|f_h(w_h) - a_h(v_h, w_h) + a(u - v_h, w_h) - f(w_h) + a(v_h, w_h)|}{\|w_h\|_V} \right].$$

Nun gilt aber mit der Stetigkeit $|a(u, v)| \leq \|a\| \|u\|_V \|v\|_V$ die folgende Ungleichung:

$$\tilde{\alpha} \|u_h - v_h\|_V \leq \|a\| \|u - v_h\|_V + \sup_{w_h \in V_h} \left[\frac{|f_h(w_h) - f(w_h)|}{\|w_h\|_V} + \frac{|a(v_h, w_h) - a_h(v_h, w_h)|}{\|w_h\|_V} \right].$$

Somit folgt mit Hilfe von (3.53) und $\tilde{C} := \max\{\frac{\|a\|}{\tilde{\alpha}} + 1, 1\}$ nach Bildung des Infimums über $v_h \in V_h$ die gewünschte Behauptung:

$$\|u - u_h\|_V \leq \tilde{C} \left[\inf_{v_h \in V_h} \left(\|u - v_h\|_V + \sup_{w_h \in V_h} \frac{|a(v_h, w_h) - a_h(v_h, w_h)|}{\|w_h\|_V} \right) + \sup_{w_h \in V_h} \frac{|f(w) - f_h(w_h)|}{\|w_h\|_V} \right].$$

□

Wir können nun Lemma 3.21 auf die Sesquilinearformen in (3.48) und (3.49) anwenden und setzen

$$a(u, v) := \int_{\Omega_{\text{int}}} \nabla u \cdot \overline{\nabla v} \, dx - \omega^2 \int_{\Omega_{\text{int}}} u \overline{v} \, dx - \langle \mathcal{DtN}(u|_{\Gamma_a}), v|_{\Gamma_a} \rangle$$

sowie

$$a_h(u_h, v_h) := \int_{\Omega_{\text{int}}} \nabla u_h \cdot \overline{\nabla v_h} \, dx - \omega^2 \int_{\Omega_{\text{int}}} u_h \overline{v_h} \, dx - \langle \text{DtN}^{(h)}(u_h|_{\Gamma_a}), v_h|_{\Gamma_a} \rangle.$$

Gehen wir nun davon aus, dass (3.48) eindeutig lösbar ist, so folgt insbesondere nach Satz 3.8, dass $a(\cdot, \cdot)$ auf $H^1(\Omega_{\text{int}})$ die (kontinuierliche) *inf-sup* Bedingung erfüllt. Setzen wir nun aber eine Lösung u aus einem endlichdimensionalen Raum $V_h \subset H^1(\Omega)$ an, so überträgt sich diese Eigenschaft nicht unmittelbar auf das diskrete Setting. Unter gewissen Voraussetzungen an $a(\cdot, \cdot)$ und V_h gilt jedoch das folgende Resultat:

Lemma 3.22. Seien V, H Hilberträume mit kompakter Einbettung $V \hookrightarrow H$. Weiters erfüllt die stetige Sesquilinearform $a : V \times V \rightarrow \mathbb{C}$ die Garding Ungleichung (3.10) und sei der durch $a(\cdot, \cdot)$ induzierte Operator $A : V \rightarrow V'$, $Au := a(u, \cdot)$ bijektiv. Überdies sei $(V_h)_{h>0}$ eine Familie von endlichdimensionalen Unterräumen $V_h \subset V$ mit

$$\lim_{h \rightarrow 0} \inf_{v \in V_h} \|u - v\|_V = 0 \quad \forall u \in V.$$

Dann existiert $h_0 > 0$ und $\tilde{\alpha}$, sodass für alle $h \in (0, h_0)$ die *diskrete inf-sup* Bedingung gleichmäßig erfüllt ist:

$$\inf_{0 \neq v_h \in V_h} \sup_{0 \neq w_h \in V_h} \frac{|a(v_h, w_h)|}{\|v_h\|_V \|w_h\|_V} \geq \tilde{\alpha} > 0. \quad (3.54)$$

Ein Beweis der Aussage findet sich zum Beispiel in [NFP21, Theorem 6.18]. Der obige Satz sichert nun ebenfalls die Lösbarkeit von (3.48) auf V_h , sofern die endlich-dimensionalen Unterräume passend gewählt werden. Es bleibt jedoch noch die Frage zu klären, unter welchen Voraussetzungen das diskrete Problem (3.49) eindeutig lösbar ist. Dazu reicht es nach Satz 3.9, dass $a_h(\cdot, \cdot)$ die *inf-sup* Bedingung auf V_h erfüllt. Um dies nachzuweisen erweist sich das folgende Lemma als nützlich:

Lemma 3.23. Seien $a(\cdot, \cdot)$ und $a_h(\cdot, \cdot)$ Sesquilinearform auf V bzw. $V_h \subset V$. Weiters setzen wir voraus, dass $a(\cdot, \cdot)$ für h hinreichend klein auf V_h gleichmäßig die *inf-sup* Bedingung, also (3.54) erfüllt. Gilt die Konvergenz von a_h gegen a im folgenden Sinne

$$\lim_{h \rightarrow \infty} E_h = 0 \quad \text{mit} \quad E_h := \sup_{v_h, w_h \in V_h \setminus \{0\}} \frac{|a(v_h, w_h) - a_h(v_h, w_h)|}{\|v_h\|_V \|w_h\|_V}, \quad (3.55)$$

dann erfüllt $a_h(\cdot, \cdot)$ für hinreichend kleine h ebenfalls gleichmäßig die *inf-sup* Bedingung:

$$\exists \tilde{\alpha} \exists h_0 : \forall h \in (0, h_0) \forall v_h, w_h \in V_h \setminus \{0\} : \inf_{0 \neq v_h \in V_h} \sup_{0 \neq w_h \in V_h} \frac{|a_h(v_h, w_h)|}{\|v_h\|_V \|w_h\|_V} =: \alpha_h \geq \tilde{\alpha} > 0$$

Beweis. Wir verwenden die gleiche Beweistechnik wie in [NFP21, Proposition 5.1]. Aufgrund der Voraussetzung (3.55) können wir $\tilde{\alpha} > \epsilon > 0$ und $h_0 > 0$ wählen, sodass für $h \in (0, h_0)$ gilt:

$$\sup_{v_h, w_h \in V_h \setminus \{0\}} \frac{|a(v_h, w_h) - a_h(v_h, w_h)|}{\|v_h\|_V \|w_h\|_V} \leq \epsilon. \quad (3.56)$$

Nach der Voraussetzung an $a(\cdot, \cdot)$ existiert ebenfalls ein $0 < h_1 < h_0$, sodass für alle $h \in (0, h_1)$ gilt

$$\begin{aligned} \alpha \|v_h\| &\leq \sup_{w_h \in V_h \setminus \{0\}} \frac{|a(v_h, w_h)|}{\|w_h\|_V} \\ &\leq \sup_{w_h \in V_h \setminus \{0\}} \frac{|a(v_h, w_h) - a_h(v_h, w_h)| + |a_h(v_h, w_h)|}{\|w_h\|_V} \\ &\stackrel{(3.56)}{\leq} \epsilon \|v_h\| + \sup_{w_h \in V_h \setminus \{0\}} \frac{|a_h(v_h, w_h)|}{\|w_h\|_V}. \end{aligned}$$

Somit ergibt sich für $0 < h < h_1$

$$\tilde{\alpha} := (\alpha - \epsilon) \leq \sup_{w_h \in V_h \setminus \{0\}} \frac{|a_h(v_h, w_h)|}{\|w_h\|_H \|v_h\|_H},$$

womit nach Bildung des Infimums schlussendlich die Behauptung folgt:

$$\tilde{\alpha} \leq \inf_{v_h \in V_h \setminus \{0\}} \sup_{w_h \in V_h \setminus \{0\}} \frac{|a_h(v_h, w_h)|}{\|w_h\|_H \|v_h\|_H}.$$

□

Mit jenen Vorbereitungen können wir nun zumindest theoretisch auf die Konvergenz von Lösungen des „voll-diskreten Verfahrens“ (3.49) auf die des kontinuierliche Problem schließen. Um die Argumentationen zu veranschaulichen, schließen wir dieses Kapitel mit einem Beispiel.

Beispiel 3.24. Wir betrachten die Helmholtz Gleichung in einem (homogenen) Wellenleiter (siehe auch Kapitel 3.3.2) und verwenden die schwache Formulierung (3.45) als Ausgangspunkt:

$$\int_{\Omega_{\text{int}}} \nabla w \overline{\nabla v} - \omega^2 w \bar{v} \, dx - \int_{\Gamma_a} \mathcal{DtN}^W(w|_{\Gamma_a}) \overline{v|_{\Gamma_a}} \, ds = \int_{\Omega_{\text{int}}} \hat{u}_0 \bar{v} \, dx, \quad \forall v \in H^1(\Omega_{\text{int}}^{\Gamma_D}).$$

Dabei bezeichnet \hat{u}_0 eine beliebige Fortsetzung aus $H^1(\Omega_{\text{int}})$ mit $\hat{u}_0|_{\Gamma_D} = u_0$. Wir haben im Beweis von Satz 3.20 bereits nachgewiesen, dass $a(\cdot, \cdot)$ die Garding Ungleichung erfüllt und der induzierte Operator A bijektiv ist, da ja die schwache Formulierung (für $\omega \in \mathbb{R} \setminus \mathbb{Z}$)

$$a(u, v) = f(v)$$

eine eindeutige Lösung besitzt. Für das diskrete Problem wählen wir den Spline Raum einer Triangulierung \mathcal{T} von Ω_{int} , setzen also $V_h := \mathcal{S}^p(\mathcal{T})$ und sehen mit Satz 2.15, dass die Voraussetzungen von Lemma 3.22 erfüllt sind. Somit erfüllt $a(\cdot, \cdot)$ auf V_h für h hinreichend klein ebenfalls die diskrete inf-sup Bedingung

$$\inf_{0 \neq v_h \in V_h} \sup_{0 \neq w_h \in V_h} \frac{|a(v_h, w_h)|}{\|v_h\|_V \|w_h\|_V} =: \alpha_h \geq \tilde{\alpha} > 0.$$

Können wir überdies die Approximation an den DtN Operator so wählen, dass gilt

$$\lim_{h \rightarrow 0} \|\mathcal{DtN}^W - \text{DtN}^{(h)}\|_{H^{1/2}(\Gamma_a) \rightarrow H^{-1/2}(\Gamma_a)} = 0,$$

dann sehen wir, dass $a_h(\cdot, \cdot)$ definiert durch

$$a_h(u_h, v_h) := \int_{\Omega_{\text{int}}} \nabla u_h \overline{\nabla v_h} - \omega^2 u_h \overline{v_h} dx - \int_{\Gamma_a} \text{DtN}^{(h)}(u_h|_{\Gamma_a}) \overline{v_h|_{\Gamma_a}} ds$$

nach Lemma 3.23 ebenfalls die diskrete inf-sup Bedingung (sogar gleichmäßig) erfüllt. Lemma 3.9 garantiert uns nun die Existenz einer eindeutigen Lösung $u_h \in \mathcal{S}^p(\mathcal{T})$ von

$$a_h(u_h, v_h) = f(v_h), \quad \forall v_h \in \mathcal{S}^p(\mathcal{T}).$$

Diese Lösung konvergiert nach Lemma 3.21 überdies gegen die Lösung u des kontinuierlichen Problems

$$a(u, v) = f(v), \quad \forall v \in H^1(\Omega).$$

Abschließend bemerken wir, dass die obige Argumentation sich der Annahme bedient, dass eine „gute Approximation“ an den Dirichlet-to-Neumann Operator existiert. Die Frage nach der Konstruktion einer solchen Approximation erweist sich als durchaus komplex, da einerseits eine unendliche Reihe approximiert werden muss und andererseits die Koeffizienten Integrale über den gesamten Rand Γ beinhalten, womit bei einer „naiven Implementierung“ (durch zB Taylor Reihe) die resultierenden Matrizen ihre *sparse* Struktur verlieren würden.

4 Diskretisierung des DtN Operators

In den bisherigen Abschnitten konnten wir bereits teilweise klären, unter welchen Voraussetzungen wir eindeutige Lösungen von kontinuierlichen Problemen erwarten können. Für die Verwendung eines numerischen Verfahrens hilft diese Information alleine jedoch nicht sonderlich weiter, weshalb in diesem Kapitel besonderes Augenmerk in Hinblick einer praktischen Implementierung gelegt wird. Dazu präsentieren wir zunächst eine Methode um den Dirichlet-to-Neumann Operator effizient zu diskretisieren und werden daran anschließend diverse Problemstellungen numerisch lösen und deren Ergebnisse anführen.

4.1 Learned Infinite Elemente Verfahren

In diesem Abschnitt werden wir eine vergleichsweise neue Diskretisierungsmethode des Dirichlet-to-Neumann Operators, die sogenannte *Learned Infinite Elemente Methode*, vorstellen. Die Idee stammt ursprünglich aus [HLP21] und besteht im Wesentlichen daraus, die Finite Elemente Matrix über externe Freiheitsgrade so zu erweitern, dass deren Schur Komplement den exakten DtN Operator „bestmöglich“ approximiert. Dabei werden die erweiterten Einträge über ein gewisses Optimierungsproblem (*Learning process*) bestimmt, was auch der Grund für die Namensgebung des Verfahrens ist. Wir präsentieren im Folgenden nur die Ideen, für genaue Beweise sei auf [HLP21] verwiesen. Essentiell für die nachfolgenden Überlegungen ist, dass sich der Dirichlet-to-Neumann Operator aus Definition 3.6 wie folgt darstellen lässt:

$$\mathcal{DtN}u_0 = \sum_{n=1}^{\infty} dtn(\lambda_n)(u_0, v_n)_{L^2(\Gamma)} v_n. \quad (4.1)$$

Um die Notation zu vereinfachen, bezeichnen wir fortan mit Γ den transparenten Rand und mit (λ_n, v_n) die Eigenpaare von $-\Delta_\Gamma$, wobei die Eigenfunktionen v_n orthonormal bezüglich dem L^2 Skalarprodukt auf Γ gewählt sind. Die sogenannten Dirichlet-to-Neumann Nummern dtn stellen also eine zunächst nur auf dem Spektrum $\sigma(-\Delta_\Gamma)$ definierte Funktion dar.

Bemerkung 4.1. Die Darstellung in (4.1) besagt im Falle der Helmholtz Gleichung im Wesentlichen, dass das Außenraumproblem (3.17) separierbar bezüglich einer Orthonormalbasis v_n aus Eigenfunktionen des negativen Laplace-Beltrami Operators $-\Delta_\Gamma$ ist. In den zuvor diskutierten Beispielen aus Kapitel 3.3.1 und 3.3.2 haben wir nach einigem Aufwand eine explizite Darstellung wie in (3.29) bzw. (3.43) angeben können, welche genau der obigen Form entspricht.

Um den exakten Operator \mathcal{DtN} schließlich zu approximieren, werden wir zunächst diskutieren, wie die FEM Matrix zu erweitern ist. In einem weiteren Schritt werden wir dann zeigen, dass das Schur Komplement der erweiterten Matrix bzw. Freiheitsgraden einen Operator (den sogenannten diskreten Dirichlet-to-Neumann Operator), der nachfolgenden Form darstellt:

$$\text{DtN}u_0 = \sum_{n=1}^N \text{dtn}(\lambda_n^{(h)}) \langle u_0, v_n^{(h)} \rangle_M \cdot v_n^{(h)}. \quad (4.2)$$

Dabei sind $(\lambda_n^{(h)}, v_n^{(h)})$ die Eigenpaare des diskreten Laplace-Beltrami Operators $-\Delta_{\Gamma,h}$ (bezüglich finiter Elemente Diskretisierung) und $\langle \cdot, \cdot \rangle_M$ das induzierte Skalarprodukt einer Matrix M , welche wir ebenfalls noch genauer spezifizieren werden. Die approximierten Dirichlet-to-Neumann Nummern $\text{dtn}(\lambda_n^{(h)})$ bilden somit per Definition ebenfalls eine auf dem diskreten Spektrum definierte Funktion. Wir nehmen an dieser Stelle gleich vorweg, dass die diskreten Eigenwerte nur für theoretische Argumentationen verwendet werden, nicht aber für die Implementierung selbst. Dies ist insofern von Bedeutung, da eine zusätzliche Berechnung der diskreten Eigenwerte im Allgemeinen mit einem hohen Aufwand verbunden ist. Aus der Darstellung aus (4.2) ist nun auch klar (siehe auch Kapitel 3.4), dass der diskrete Operator 4.2 nur dann eine gute Approximation des exakten DtN Operators darstellt, sofern wir die Fehler in den Eigenfunktionen $\|v_n - v_n^{(h)}\|_{L^2(\Gamma)}$, den Eigenwerten $|\lambda_n - \lambda_n^{(h)}|$ und den Dirichlet-to-Neumann Nummern $|\text{dtn}(\lambda_n) - \text{dtn}(\lambda_n^{(h)})|$ kontrollieren können.

Wir starten also damit zu erklären, wie die Darstellung in (4.2) gerechtfertigt werden kann. Um die Präsentation zu vereinfachen, beschränken wir uns auf die Lösung des nachfolgenden Modellproblems:

Suche $u \in H^1(\Omega_{\text{int}})$, sodass

$$\int_{\Omega_{\text{int}}} \nabla u \overline{\nabla v} \, dx - \omega^2 \int_{\Omega_{\text{int}}} u \overline{v} \, dx - \int_{\Gamma} \mathcal{DtN}(u|_{\Gamma}) \overline{v|_{\Gamma}} \, ds = \int_{\Omega_{\text{int}}} f \overline{v} \, dx \quad (4.3)$$

für alle $v \in H^1(\Omega_{\text{int}})$ gilt. Dabei ist $f \in L^2(\Omega_{\text{int}})$ zusätzlich vorausgesetzt, damit die Formulierung überhaupt wohldefiniert ist. Geht man zur diskreten Formulierung zu obigem Problem über, setzt also $u_h, v_h \in X_h$ mit $\dim X_h =: n < \infty$ und einer dazugehörigen Basis $\{\varphi_m : m = 1, \dots, n\}$, so sind zunächst die Matrizen

$$P_{ij} := \int_{\Omega_{\text{int}}} \nabla \varphi_i \overline{\nabla \varphi_j} \, dx, \quad Q_{ij} := \int_{\Omega_{\text{int}}} \varphi_i \overline{\varphi_j} \, dx, \quad P, Q \in \mathbb{C}^{n \times n}$$

sowie

$$K_{ij} := \int_{\Gamma} \nabla \varphi_i \overline{\nabla \varphi_j} \, ds, \quad M_{ij} := \int_{\Gamma} \varphi_i \overline{\varphi_j} \, ds, \quad K, M \in \mathbb{C}^{n_{\Gamma} \times n_{\Gamma}}$$

und ebenfalls einen „rechte Seite“

$$F_i := \int_{\Omega_{\text{int}}} f \overline{\varphi_i} \, dx, \quad F \in \mathbb{C}^n$$

aufzustellen. Um eine geeignete Matrix Formulierung zu erhalten, unterteilen wir noch die Freiheitsgrade von $u_h = \sum_n \hat{u}_n \varphi_n$ in $\hat{u} = (u_I, u_\Gamma)^T$, wobei Γ die n_Γ Freiheitsgrade zugehörig zu dem (transparenten) Rand Γ und I die n_I inneren Freiheitsgrade (also alle anderen) bezeichnet. Klarerweise gilt dabei $n = n_\Gamma + n_I$. Die diskrete Formulierung von (4.3) kann nun wie folgt geschrieben werden:

$$\left[\begin{pmatrix} P_{II} & P_{\Gamma I} \\ P_{I\Gamma} & P_{\Gamma\Gamma} \end{pmatrix} - \omega^2 \begin{pmatrix} Q_{II} & Q_{\Gamma I} \\ Q_{I\Gamma} & Q_{\Gamma\Gamma} \end{pmatrix} \right] \begin{pmatrix} u_I \\ u_\Gamma \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & M \mathcal{DtN} \end{pmatrix} \begin{pmatrix} u_I \\ u_\Gamma \end{pmatrix} = \begin{pmatrix} F_I \\ F_\Gamma \end{pmatrix}.$$

Bemerkung 4.2. Streng genommen müsste man an dieser Stelle noch klarstellen, wie der Operator \mathcal{DtN} für $u_\Gamma \in \mathbb{C}^{n_\Gamma}$ definiert ist. Dies können wir an jener Stelle am einfachsten mit Definition 3.6 bewerkstelligen, nach dieser der Operator lediglich den Zusammenhang zwischen Dirichletdaten u_Γ und Neumanndaten u_N beschreibt, wobei die Neumanndaten aus der Lösung des Außenraumproblems bei gegebenen Dirichletdaten berechnet werden. Somit ist in diesem Fall $\mathcal{DtN}(u_\Gamma) := u_N$ als Definition zu verwenden. Die Definition findet also wieder implizit statt, stört uns aber nicht weiter, da wir sowieso nur an einer Approximation interessiert sind, die wir schlussendlich auch numerisch implementieren können.

Durch setzen von $S := P - \omega^2 Q$ erhalten wir eine etwas kompaktere Formulierung welche wie folgt lautet:

$$\begin{pmatrix} S_{II} & S_{\Gamma I} \\ S_{I\Gamma} & S_{\Gamma\Gamma} - M \mathcal{DtN} \end{pmatrix} \begin{pmatrix} u_I \\ u_\Gamma \end{pmatrix} = \begin{pmatrix} F_I \\ F_\Gamma \end{pmatrix}. \quad (4.4)$$

Problematisch an dieser Formulierung ist jedoch immer noch der Dirichlet-to-Neumann Operator \mathcal{DtN} , den wir ohne weiteres nicht numerisch implementieren können. An dieser Stelle kommt nun die Learned Infinite Elements Methode ins Spiel. Bei dieser Methode wird, wie bereits erwähnt, die Matrix in (4.4) durch zusätzliche Freiheitsgrade erweitert. Diese werden über den transparenten Rand Γ mit der FEM Matrix gekoppelt, sodass wir das folgende erweiterte Gleichungssystem erhalten

$$\begin{pmatrix} S_{II} & S_{\Gamma I} & 0 \\ S_{I\Gamma} & S_{\Gamma\Gamma} + L_{\Gamma\Gamma} & L_{\Gamma E} \\ 0 & L_{E\Gamma} & L_{EE} \end{pmatrix} \begin{pmatrix} u_I \\ u_\Gamma \\ u_E \end{pmatrix} = \begin{pmatrix} F_I \\ F_\Gamma \\ 0 \end{pmatrix}. \quad (4.5)$$

Zunächst wollen wir die erweiternden Matrizen noch nicht näher spezifizieren, halten jedoch fest, dass mit dem Schur Komplement der Matrix L_{EE} , welches wie folgt definiert ist

$$L_{\Gamma\Gamma} - L_{\Gamma E} L_{EE}^{-1} L_{E\Gamma},$$

der exakte Operator aus Formulierung (4.4) in der erweiterten Matrix aus (4.5) wie folgt ersetzt bzw. approximiert wird:

$$\mathcal{DtN} \approx \text{DtN}^{(LIE)} := M^{-1} (L_{\Gamma\Gamma} - L_{\Gamma E} L_{EE}^{-1} L_{E\Gamma}). \quad (4.6)$$

Es bleibt an dieser Stelle noch die Frage offen, wie die Matrizen $L_{\Gamma\Gamma}$, $L_{\Gamma E}$, $L_{E\Gamma}$ und L_{EE} gewählt werden müssen, damit die Darstellung über das Schur Komplement (4.6) überhaupt

sinnvoll ist. Um diese Frage beantworten zu können, erinnern wir daran, wie das Tensorprodukt zweier Matrizen $A \in \mathbb{C}^{m \times n}$ und $B \in \mathbb{C}^{p \times q}$ definiert ist:

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix} \in \mathbb{C}^{mp \times nq}.$$

Um die zusätzlichen Matrizen in (4.5) schlussendlich aufzustellen, verwenden wir nun eben jenen Tensorprodukt-Ansatz, setzen also

$$\begin{pmatrix} L_{\Gamma\Gamma} & L_{\Gamma E} \\ L_{E\Gamma} & L_{EE} \end{pmatrix} = \underbrace{\begin{pmatrix} A_{\Gamma\Gamma} & A_{\Gamma E} \\ A_{E\Gamma} & A_{EE} \end{pmatrix}}_{=:A} \otimes M + \underbrace{\begin{pmatrix} B_{\Gamma\Gamma} & B_{\Gamma E} \\ B_{E\Gamma} & B_{EE} \end{pmatrix}}_{=:B} \otimes K \quad (4.7)$$

für gewisse Matrizen $A, B \in \mathbb{C}^{N_{\text{ext}} \times N_{\text{ext}}}$ mit $A_{\Gamma\Gamma}, B_{\Gamma\Gamma} \in \mathbb{C}$. Dabei stellen N_{ext} die Anzahl der externen Freiheitsgrade dar und wir nehmen überdies an, dass N_{ext} in der Größenordnung 10 ist, da die Einträge der Matrizen A, B im Endeffekt über einen „Learning Process“ bestimmt werden müssen. Bevor wir das Hauptresultat dieses Abschnitts anführen, wollen wir noch den schon zuvor angesprochenen, diskreten Laplace-Beltrami Operator $-\Delta_{\Gamma,h}$ etwas näher spezifizieren. Wir sagen, dass $(\lambda_n^{(h)}, v_n^{(h)})$ Eigenpaar von $-\Delta_{\Gamma,h}$ ist, wenn gilt $(v_n^{(h)} = \sum_i \hat{v}_n^i \varphi_i)$:

$$K \hat{v}_n = \lambda_n^{(h)} M \hat{v}_n, \quad 0 \neq \hat{v}_n \in \mathbb{C}^{n_\Gamma}. \quad (4.8)$$

Dies entspricht einem verallgemeinerten Eigenwertproblem. In der Literatur wird deshalb auch oft der diskrete Laplace-Beltrami Operator durch die Matrizen der Oberflächenelemente direkt über die Identität $-\Delta_{\Gamma,h} = M^{-1}K$ beschrieben. Mit jenen Vorbereitungen können wir somit den Hauptsatz dieses Abschnitts formulieren.

Satz 4.3. *Mit den Voraussetzungen von zu Beginn des Kapitels sei $(\lambda_n^{(h)}, v_n^{(h)})$ ein Eigenpaar von $-\Delta_{\Gamma,h}$, dann erfüllt der in (4.6) definierte approximierte DtN Operator mit den in (4.7) definierten Matrizen die folgende Eigenschaft:*

$$DtN^{(LIE)} \hat{v}_n = dtn^{(LIE)}(\lambda_n^{(h)}) \hat{v}_n. \quad (4.9)$$

Dabei stellt $dtn^{(LIE)}(\lambda)$ für $\lambda \in \mathbb{C}$ eine rationale Funktion der folgenden Form dar:

$$dtn^{(LIE)}(\lambda) := A_{\Gamma\Gamma} + \lambda B_{\Gamma\Gamma} - (A_{\Gamma E} + \lambda B_{\Gamma E})(A_{EE} + \lambda B_{EE})^{-1}(A_{E\Gamma} + \lambda B_{E\Gamma}). \quad (4.10)$$

Für einen Beweis sei auf [HLP21, Theorem 2.3] verwiesen. Wir wollen nun die Wichtigkeit des obigen Satzes erläutern. Dazu stellen wir zunächst fest, dass durch die diskreten Eigenfunktionen $\{\hat{v}_n : n = 1, \dots, n_\Gamma\}$ eine Basis des \mathbb{C}^{n_Γ} gegeben ist. Des weiteren ist durch

$$\langle \cdot, \cdot \rangle_M : \mathbb{C}^{n_\Gamma} \times \mathbb{C}^{n_\Gamma} \rightarrow \mathbb{C}, (v, w) \mapsto v^H M w$$

ein Skalarprodukt definiert. Durch geeignete Skalierung können wir $\{\hat{v}_n : n = 1, \dots, n_\Gamma\}$ sogar als Orthonormalbasis von \mathbb{C}^{n_Γ} bezüglich des durch M induzierten Skalarproduktes wählen. Dies bedeutet, dass wir beliebiges $u_\Gamma \in \mathbb{C}^{n_\Gamma}$ schreiben können als

$$u_\Gamma = \sum_{n=1}^{n_\Gamma} \langle u_\Gamma, \hat{v}_n \rangle_M \hat{v}_n.$$

Somit wissen wir insbesondere, wie der approximierte DtN Operator aus (4.6) operiert, da wir nach (4.9) wissen wie dieser auf einer bestimmten Basis agiert. Zusammengefasst gilt für beliebiges $u_\Gamma \in \mathbb{C}^{n_\Gamma}$ also:

$$\text{DtN}^{(\text{LIE})} u_\Gamma = \sum_{n=1}^{n_\Gamma} \text{dtn}^{(\text{LIE})}(\lambda_n^{(h)}) \langle u_\Gamma, \hat{v}_n \rangle_M \hat{v}_n.$$

Dies entspricht nun exakt der zu Beginn des Kapitels behaupteten Darstellung (4.2). Die vorausgehenden Überlegungen fassen nun im Wesentlichen die Ideen der Learned Infinite Elemente Methode zusammen. In praktischer Hinsicht fehlt jedoch immer noch die genaue Bestimmung der Matrizen A und B aus (4.10). Dies geschieht in einem sogenannten „Learning Prozess“, den wir im folgenden Abschnitt erklären. Grundsätzlich kann man jedoch sehen, dass für $N_{\text{ext}} \in \mathbb{N}$ externe Freiheitsgrade genau $2N_{\text{ext}}^2$ Einträge bestimmt werden sollen. Tatsächlich kann man wie in [HLP21, Kapitel 4] argumentieren, dass es ausreicht Matrizen der folgenden Bauart zu betrachten:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1N_{\text{ext}}} \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ a_{N_{\text{ext}}1} & 0 & \cdots & 0 & a_{N_{\text{ext}}N_{\text{ext}}} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} & \cdots & \cdots & b_{1N_{\text{ext}}} \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

womit nur mehr $4N_{\text{ext}} - 2$ Einträge berechnet werden müssen. Dieser reduzierte Ansatz resultiert weiters durch Einsetzen in Formel (4.10) in den nachfolgenden approximierten Dirichlet-to-Neumann Nummern:

$$\text{dtn}^{(\text{LIE})}(\lambda) = a_{00} + \lambda \cdot b_{00} + \sum_{j=1}^{N_{\text{ext}}} \frac{(a_{0j} + \lambda \cdot b_{0j})(a_{j0} + \lambda)}{a_{jj} + \lambda}. \quad (4.11)$$

Man sieht unmittelbar, dass dies einer rationale Funktion entspricht, die sich asymptotisch linear verhält und Pole der Ordnung 1 an

$$\bar{\lambda}_j := -a_{jj}, \quad j = 1, \dots, N_{\text{ext}}$$

besitzt. Vieles deutet nun auf ein Optimierungsproblem zur Bestimmung der Koeffizienten hin. Um dieses sinnvoll aufsetzen zu können müssen wir natürlich beachten, dass die exakten Dirichlet-to-Neumann Nummern \mathcal{dtn} zunächst nur auf dem Spektrum $-\Delta_\Gamma$ definiert sind. Weiters ist die in (4.11) definierte Funktion außerdem asymptotisch linear, wonach

wir selbstverständlich nur einen kleinen Teil des Spektrums (wenn überhaupt) hinreichend gut approximieren können. Dies wirkt zwar auf den ersten Blick wenig zufriedenstellend, jedoch ist durch die Darstellung des Operators \mathcal{DtN} als unendliche Reihe in (4.1) ohnehin klar, dass nur endlich viele Summanden zur Lösung beitragen (die Koeffizienten der Reihe müssen ja exponentiell fallen, damit diese konvergiert) und somit nur diese hinreichend gut approximiert werden müssen. Mit diesen Vorbereitungen können wir nun das Optimierungsproblem wie folgt definieren:

Definition 4.4 (Learning Prozess). Bezeichne mit λ_n für $n \in \mathbb{N}_0$ die Eigenwerte des kontinuierlichen Laplace Beltrami Operators $-\Delta_\Gamma$. Weiters seien $\mathcal{dtn}(\lambda_n)$ die kontinuierlichen DtN Nummern aus (4.1) und $\text{dtn}^{(\text{LIE})}(\lambda)$ die approximierten DtN Nummern wie in (4.11) beschrieben. Dann bezeichnen wir für positive Gewichte $\omega_n \geq 0$ den gewichteten Fehler mit

$$E(A, B) := \frac{1}{2} \sum_{n \in \mathbb{N}} \left| \omega_n (\mathcal{dtn}(\lambda_n) - \text{dtn}^{(\text{LIE})}(\lambda_n)) \right|. \quad (4.12)$$

Das Optimierungsproblem besteht nun darin $A, B \in \mathbb{C}^{N_{\text{ext}} \times N_{\text{ext}}}$ zu finden, sodass der gewichtete Fehler $E(A, B)$ minimiert wird.

Bemerkung 4.5. Wir bemerken an dieser Stelle, dass im Optimierungsproblem keine Information bezüglich der diskreten Eigenwerte $\lambda_n^{(h)}$ verwendet wird. In Bezug auf eine numerische Implementierung ist somit nur die Kenntnis der exakten Eigenwerte von $-\Delta_\Gamma$ und im weiteren Sinne von $\mathcal{dtn}(\lambda_n)$ nötig. Dies ist durch die Annahme, dass das Außenraumproblem (3.17) analytisch lösbar ist, zumindest theoretisch gerechtfertigt. Die Gewichte ω_n bestimmen überdies wie genau die jeweiligen Eigenwerte λ_n approximiert werden sollen. Da die Koeffizienten $\mathcal{dtn}(\lambda_n)$ des exakten DtN Operators ohnehin exponentiell fallen, kann man somit einen relevanten Bereich festlegen, womit die Fehlerfunktion $E(.,.)$ in (4.12) ebenfalls eine endliche Summe darstellt.

Zusammengefasst benötigt die Learned Infinite Elemente Methode also einen „pre-processing“ Schritt, bei der die Koeffizienten der Matrizen A, B „gelernt“ werden. Diese werden schließlich durch die Freiheitsgrade auf dem transparenten Rand Γ mit dem aus der FEM Diskretisierung resultierenden Gleichungssystem gekoppelt. Die eigentliche Approximation des DtN Operators geschieht dann über das Schur Komplement dieser externen Freiheitsgrade. Der nachfolgende Abschnitt beinhaltet einige numerische Beispiele, an denen die vorausgehende Theorie veranschaulicht werden soll.

4.2 Numerische Resultate

Dieses Kapitel beinhaltet die ersten numerischen Experimente, welche die zuvor präsentierte Learned Infinite Elemente Methode veranschaulichen sollen. Dabei beschränken wir uns hauptsächlich auf das Lösen der Helmholtz Gleichung in Wellenleitern, siehe auch Kapitel 3.3.2. Um die nachfolgende Präsentation zu vereinfachen, verwenden wir, falls nicht anders

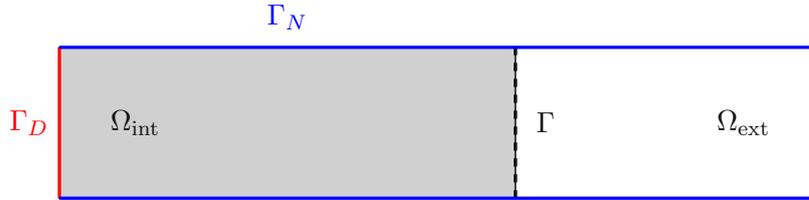


Abbildung 4.1: Darstellung der Geometrie für nachfolgende numerische Experimente

spezifiziert, Wellenleiter der Breite π , setzen also $\Omega := (0, \infty) \times (0, \pi)$. Wir interessieren uns somit an der schwachen Lösung des folgenden Problems:

$$\begin{aligned} -\Delta u - \omega^2 u &= 0 & \text{in } \Omega \\ u &= u_0 & \text{auf } \Gamma_D \\ \frac{\partial u}{\partial \eta} &= 0 & \text{auf } \Gamma_N. \end{aligned} \quad (4.13)$$

Wie bereits mehrmals durchgeführt, erhalten wir nach Setzen eines transparenten Randes $\Gamma := \{a\} \times (0, \pi)$ mit $a > 0$ die folgende schwache Formulierung, siehe auch (3.45):

Suche $u \in H^1(\Omega_{\text{int}}^{\Gamma_D})$, sodass

$$\int_{\Omega_{\text{int}}} \nabla u \overline{\nabla v} - \omega^2 u \overline{v} \, dx - \int_{\Gamma} \mathcal{DtN}(u|_{\Gamma}) \overline{v|_{\Gamma}} \, ds = \int_{\Omega_{\text{int}}} \hat{u}_0 \overline{v} \, dx, \quad (4.14)$$

für alle $v \in H^1(\Omega_{\text{int}}^{\Gamma_D})$ gilt. Dabei bezeichnet $\Omega_{\text{int}} := (0, a) \times (0, \pi)$ und $\hat{u}_0 \in H^1(\Omega_{\text{int}})$ eine beliebige Fortsetzung von u_0 mit $\hat{u}_0|_{\Gamma} = u_0$ im Sinne des Spurooperators. In den nächsten Kapiteln werden wir ausgehend von dieser Formulierung einige Experimente durchführen und deren Resultate präsentieren.

4.2.1 Einführendes Beispiel

Wir starten diesen Abschnitt, indem wir die Formulierung aus (4.14) mittels Finiter Elemente diskretisieren. Weiters erinnern wir an die in (3.43) präsentierte, analytische Darstellung des DtN Operators:

$$\mathcal{DtN}(u|_{\Gamma}) = \sum_{n=0}^{\infty} \underbrace{i\sqrt{\omega^2 - \lambda_n}}_{\text{dtn}(\lambda_n)} \cdot (u, v_n)_{L^2(\Gamma)} v_n$$

Dabei haben wir verwendet, dass die Eigenwerte des negativen Laplace-Beltrami Operators $-\Delta_\Gamma$ auf $\Gamma := \{a\} \times (0, \pi)$ durch $\lambda_n := n^2$ gegeben sind mit dazugehörigen Eigenfunktionen $v_n = \cos(n \cdot)$. Die analytischen DtN Nummern lassen sich also durch die Funktion

$$dtn : \sigma(-\Delta_\Gamma) = \{n^2 : n \in \mathbb{N}_0\} \rightarrow \mathbb{C}, \lambda \mapsto i\sqrt{\omega^2 - \lambda}$$

beschreiben. Mit jenen Vorbereitungen können wir nun ebenfalls das Optimierungsproblem (4.12) aufstellen. Es bleibt noch die Frage offen, welche Gewichtsfunktion ω_n sinnvoll ist. Die Beantwortung jener Frage ist im Allgemeinen leider nicht-trivial, da man die Gewichte stets an das zu Grunde liegende Problem anpassen muss. Für das Modellproblem 4.13 sind wir jedenfalls in der Lage eine analytische Lösung von (4.14) analog zu (3.41) angeben zu können mittels

$$u(x, y) = \sum_{n=0}^{\infty} \exp(i\sqrt{\omega^2 - n^2} \cdot x) \cdot (u_0, v_n)_{L^2(\Gamma)} \cdot v_n(y). \quad (4.15)$$

Interessant ist nun vor allem das Verhalten der Lösung bezüglich der x Komponente. Insbesondere sehen wir mit der in (3.40) definierten komplexen Wurzel durch

$$\left| \exp(i\sqrt{\omega^2 - n^2} \cdot x) \right| = \left| \exp(-x \cdot \Im(\sqrt{\omega^2 - n^2})) \right|, \quad (4.16)$$

dass die Koeffizienten für $n > |\Re(\omega)|$ und $x > 0$ exponentiell gegen 0 konvergieren. Physikalisch bedeutet dieser Sachverhalt, dass die jeweiligen Frequenzen λ_n für $n > |\Re(\omega)|$ entlang des Wellenleiters exponentiell gedämpft werden. Es macht also Sinn den Fokus auf jene Frequenzen λ_n mit $n \leq |\Re(\omega)|$ zu legen, womit die folgende Wahl der Gewichte gerechtfertigt werden kann:

$$\omega_n = \begin{cases} 1, & n \leq |\Re(\omega)| \\ \left| \exp(i\sqrt{\omega^2 - n^2} \cdot a) \right|, & n > |\Re(\omega)|. \end{cases}$$

In einem „pre-processing“ Schritt müssen wir nun mit den obigen Überlegungen die folgende Funktion minimieren:

$$E(A, B) = \frac{1}{2} \sum_{n=0}^{L_{\max}} \left| \omega_n (dtn(\lambda_n) - dtn^{(\text{LIE})}(\lambda_n)) \right|. \quad (4.17)$$

Dabei bezeichnet L_{\max} die Anzahl der zu approximierenden DtN Nummern. Wir bedienen uns in allen in dieser Arbeit präsentierten, numerischen Experimenten des in [HLP21] vorgestellten GitLab Moduls **Ceres-Dtn** [Jan21], welches bereits eine Routine bereitstellt, die für gegebene Gewichtsfunktion ω_n , Eigenwerte λ_n und dtn Nummern das Optimierungsproblem (4.17) löst. Somit sind zunächst die folgenden Funktionen aufzustellen:

```
#exact dtn numbers
def dtn_ref(lam, omega):
    result= -1j*cmath.sqrt(omega**2-lam)
    return result
```

```

#get reference eigenvalues of negative Laplace-Beltrami Operator
def get_ref_eigenvalues(L):
    lam = np.array([n**2 for n in range(L)])
    return lam

#obtain weights for misfit function
def omega_n(lam, omega):
    if lam <= abs((omega**2).real):
        weight = 1
    else:
        weight = abs(exp(1j*cmath.sqrt(omega**2-lam)*a))
    return weight

```

Die obige Wahl ist dabei an das zu untersuchende Modellproblem (4.13) angepasst. Im Allgemeinen hängt die Wahl natürlich an dem zu betrachtenden Problem ab, wonach man jene Funktionen natürlich erneut anpassen bzw. definieren muss. Daran anschließend können wir das Optimierungsproblem wie folgt aufsetzen:

```

def setup_opt_problem(omega, Lmax):
    lams = get_ref_eigenvalues(Lmax)
    dtn_nr = np.array([ dtn_ref(lami, omega) for lami in lams ])
    weights = np.array([ omega_n(lami, omega) for lami in lams ])
    return lams, dtn_nr, weights

```

Dieser „Output“ kann somit an die oben genannte Routine übergeben werden, womit wir schlussendlich die „gelernten Matrizen“ A und B erhalten. Für weitere Details, auch in Hinblick auf den Algorithmus, der für die Minimierung des Fehlers verwendet wird, verweisen wir auf weiterführende Literatur. Wir wollen aber noch anmerken, dass für die Optimierung selbst der `Ceres Solver` aus [AMT22] verwendet wird. Wir entscheiden uns

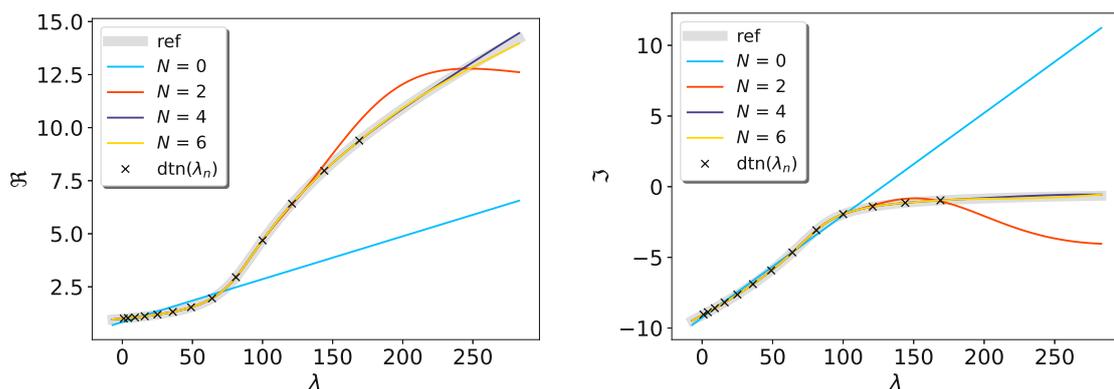


Abbildung 4.2: Approximation der $dtn(\cdot)$ Funktion mittels gelernter DtN Nummern, Darstellung des Realteils links sowie des Imaginärteils rechts für $\lambda \in (0, 300)$

in den nachfolgenden Experimenten $L_{\max} := \lceil |\Re(\omega)| \rceil + 2$ mit einer komplexen Frequenz $\omega := 9.1 + i$ zu setzen, außerdem dient als Geometrie $\Omega_{\text{int}} := (0, 2\pi) \times (0, \pi)$. Die obige Abbildung fasst die Ergebnisse des Optimierungsprozesses zusammen. Die Regularität der exakten *dtN* Funktionen sorgt dafür, dass bereits für eine geringe Anzahl an externen Freiheitsgraden N_{ext} die Approximationen (für den relevanten Teilbereich) sehr nahe an den analytischen *dtN* Nummern liegen. Diese diskreten DtN Nummern werden also, wie bereits diskutiert, in den Matrizen A, B im Sinne von Definition (4.7) gespeichert. Der nachfolgende Code beschreibt eine mögliche Implementierung in Python unter Verwendung von NetGen/NGSolve und dient lediglich zu Illustrationszwecken.

```
def SolveProblem(order, mesh, u0, A_IE, B_IE, omega):
    #create FESpaces
    fes_inner = H1(mesh, complex=True, order=order, dirichlet="bnd_dir")
    fes_surf = Compress(H1(mesh, order=order, complex=True,
        definedon=mesh.Boundaries("transp"), dirichlet="bnd_dir|bnd_neu"))
    #tensor product space for outer discretization
    inf_outer = [fes_surf for i in range(A_IE.shape[0]-1)]
    #create joined FE space
    X = FESpace([fes_inner]+inf_outer)
    #set up trial and test function
    gfu_X = GridFunction(X)
    uX, vX = X.TnT()
    #set-up Linearform
    f_X = LinearForm(X)
    #set up weak formulation
    aX = BilinearForm(X, symmetric=False)
    aX += (grad(uX[0])*grad(vX[0]) - omega**2*uX[0]*vX[0]) * dx
    #restrict to transparent boundary
    ds_outbnd = ds(definedon=mesh.Boundaries("transp"))
    for i in range(A_IE.shape[0]):
        for j in range(A_IE.shape[0]):
            if abs(A_IE[j,i]) > 1e-14 or abs(B_IE[j,i]) > 1e-14:
                aX += ( B_IE[j,i]*grad(uX[i]).Trace()*grad(vX[j]).Trace()
                    + A_IE[j,i]*uX[i]*vX[j] ) * ds_outbnd

    aX.Assemble()
    #declare dirichlet boundary
    gfu_X.components[0].Set(u0, definedon=mesh.Boundaries("bnd_dir"))
    #solve BVP
    r = f_X.vec.CreateVector()
    r.data = -aX.mat * gfu_X.vec
    gfu_X.vec.data += aX.mat.Inverse(freedofs=X.FreeDofs()) * r
    return gfu_X.components[0]
```

Interessant ist an dieser Stelle die Erweiterung der FEM Matrix durch die gelernten Matrizen A, B . Dies ist in NetGen/NGSolve sehr elegant möglich, indem man durch den Be-

fehl **FESpace** verschiedene FEM Räume kombinieren bzw. erweitern kann. Hervorheben an dieser Stelle wollen wir überdies, dass das Schur Komplement selbst nicht implementiert werden muss, da dieses lediglich aus Sicht der theoretischen Argumentation von Nöten war. Um das Modellproblem (4.14) schließlich lösen zu können, spezifizieren wir noch die Randdaten wie folgt:

$$u_0 = \sum_{n=1}^{10} \cos(n \cdot).$$

Mit dieser Wahl erhalten wir überdies eine Referenzlösung, welche sich wie folgt schreiben lässt:

$$u(x, y) = \sum_{n=1}^{10} \exp(i\sqrt{\omega^2 - n^2} \cdot x) \cdot \cos(ny).$$

Wie schon zuvor ist dabei $\omega := 9.1 + i$ gewählt worden. Links in Abbildung 4.3 visualisieren wir den L^2 Fehler $\|u - u_h\|_{L^2(\Omega_{int})}$ in Abhängigkeit der Anzahl der externen Freiheitsgrade N_{ext} des approximierten DtN Operators für fixe mesh size h . Rechts in Abbildung 4.3 zeigen wir dann den Fehler für fixes N_{ext} in Abhängigkeit der Anzahl an Freiheitsgraden $n := \dim X_h$. Zusammengefasst liefert das Learned Infinite Elemente Verfahren sehr gute

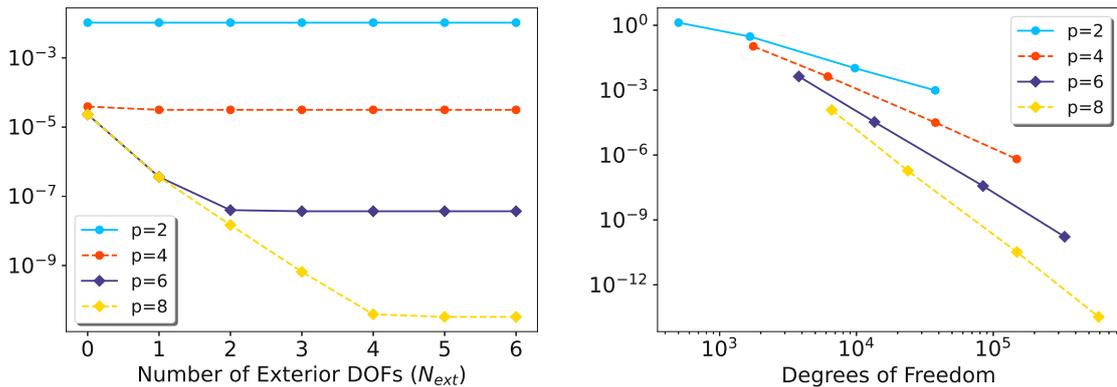


Abbildung 4.3: Links: $L^2(\Omega_{int})$ -Fehler für $h = 0.1$, Rechts: $L^2(\Omega_{int})$ -Fehler für $N_{ext} = 4$

Konvergenzraten, wie in den Abbildungen ersichtlich ist. Dies hat einerseits mit der glatten *dtN* Funktion und andererseits mit der Regularität der exakten Lösung u (aufgrund der Konvexität von Ω_{int}) zu tun, welche wie bereits diskutiert, höhere Konvergenzordnungen für Polynome höheren Grades garantiert. In den folgenden Abschnitten werden wir untersuchen, wie wir die Diskretisierung des DtN Operators auf kompliziertere Problemstellungen anwenden können.

4.2.2 Inhomogenität in der Wellenzahl

Während wir bislang $\omega \in \mathbb{C}$ als konstant angenommen haben, legen wir in diesem Abschnitt den Fokus auf das Lösen der Helmholtz Gleichung mit einer stückweise konstanten Wellen-

zahl $\omega(x)$. Genauer interessieren wir uns für die Lösung von (4.14) wobei wir annehmen, dass die Wellenzahl einen Sprung an $x = x^* > a$ außerhalb der FEM Diskretisierung des Gebietes Ω_{int} hat:

$$\omega(x) = \begin{cases} \omega_1 & , \quad x \leq x^* \\ \omega_2 & , \quad x > x^*. \end{cases}$$

Zuerst wollen wir klären, was sich im Vergleich zu dem Beispiel aus Abschnitt 4.2.1 geändert hat. Zunächst ist klar, dass die Eigenwerte λ_n des negativen Laplace-Beltrami Operators $-\Delta_\Gamma$ gleich bleiben. Tatsächlich ändert sich durch den Sprung jedoch die exakte Lösung des Außenraumproblems und somit in weiterer Konsequenz ebenfalls die exakten DtN Nummern dt_n . Diese gilt es nun zu berechnen. Wir gehen dazu ähnlich wie in Kapitel 3.3.2 vor und sehen, dass sich mit einem Separationsansatzes $u(x, y) = w(x)v(y)$ die Lösung des Problems auf die folgende gewöhnliche Differentialgleichung reduziert:

$$w_n''(x) + (\omega^2 - \lambda_n)w_n(x) = 0, \quad w_n(a) = 1.$$

Die allgemeine Lösung, ohne Berücksichtigung von $w_n(a) = 1$, ist gegeben durch

$$w_n(x) = a_n \exp(i\sqrt{\omega^2 - \lambda_n}x) + b_n \exp(-i\sqrt{\omega^2 - \lambda_n}x), \quad (4.18)$$

wonach wir auf Grund des Sprunges der Wellenzahl die gesuchte Lösung wie folgt ansetzen:

$$w_n(x) = \begin{cases} a_n \exp(i\sqrt{\omega_1^2 - \lambda_n}x) + b_n \exp(-i\sqrt{\omega_1^2 - \lambda_n}x) & , \quad x \leq x^* \\ c_n \exp(i\sqrt{\omega_2^2 - \lambda_n}x) & , \quad x > x^* \end{cases}$$

Da wir auf dem Außenraum klarerweise nach stetig differenzierbaren Lösungen suchen, können wir die Koeffizienten wie man leicht sieht, durch Lösen des nachfolgenden Gleichungssystems bestimmen: (Wir setzen $\kappa_n^i := \sqrt{\omega_i^2 - \lambda_n}$ zur besseren Lesbarkeit)

$$\begin{pmatrix} \exp(i\kappa_n^1 a) & \exp(-i\kappa_n^1 a) & 0 \\ \exp(i\kappa_n^1 x^*) & \exp(-i\kappa_n^1 x^*) & -\exp(i\kappa_n^2 x^*) \\ i\kappa_n^1 \exp(i\kappa_n^1 x^*) & -i\kappa_n^1 \exp(-i\kappa_n^1 x^*) & -i\kappa_n^2 \exp(i\kappa_n^2 x^*) \end{pmatrix} \begin{pmatrix} a_n \\ b_n \\ c_n \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad (4.19)$$

Natürlich müssen wir an dieser Stelle noch diskutieren, unter welchen Voraussetzungen das Gleichungssystem (4.19) überhaupt eindeutig lösbar ist. Aufgrund der endlichen Dimension reicht es nach dem Rangsatz der Linearen Algebra zu zeigen, dass die Spalten der Matrix linear unabhängig sind. Nehmen wir also an, dass die Spalten linear abhängig wären, dann müsste einerseits $c \in \mathbb{C}$ mit $c \neq 0$ existieren, sodass:

$$\exp(i\kappa_n^1 x^*) = c \cdot \exp(-i\kappa_n^1 x^*),$$

womit ebenfalls folgen würde:

$$\begin{aligned} i\kappa_n^1 \exp(i\kappa_n^1 x^*) &\stackrel{!}{=} -c \cdot i\kappa_n^1 \exp(-i\kappa_n^1 x^*) \\ \Leftrightarrow i\kappa_n^1 \exp(i\kappa_n^1 x^*) &= -i\kappa_n^1 \cdot \exp(i\kappa_n^1 x^*) \\ \Leftrightarrow -\kappa_n^1 &= \kappa_n^1 \Leftrightarrow \kappa_n^1 = 0. \end{aligned}$$

Somit wissen wir, dass für

$$\kappa_n^1 \neq 0 \Leftrightarrow \sqrt{\omega_1^2 - \lambda_n} \neq 0 \Leftrightarrow \omega_1^2 \neq \lambda_n, \quad n \in \mathbb{N}_0$$

eine eindeutige Lösung von (4.19) existiert. Dies rechtfertigt nun ebenfalls, das Gleichungssystem für allgemeine $\kappa_\lambda^i := \sqrt{\omega_i^2 - \lambda}$ zu betrachten, was ebenfalls die Fortsetzung der dtn Funktion auf $\mathbb{C} \setminus \{\omega_1^2\}$ ermöglicht. Wir sind nun überdies in der Lage den exakten DtN Operator $\mathcal{D}t\mathcal{N}$ anzugeben, dieser lautet nach dem bereits Gezeigten wie folgt:

$$\mathcal{D}t\mathcal{N}u_0 = \sum_{n=0}^{\infty} \left(a_n i \kappa_n^1 \exp(i \kappa_n^1 \cdot a) - b_n i \kappa_n^1 \exp(-i \kappa_n^1 \cdot a) \right) (u_0, v_n)_{L^2(\Gamma)} v_n. \quad (4.20)$$

Dabei sind die Konstanten a_n, b_n durch das Gleichungssystem (4.19) eindeutig bestimmt. Wie bereits diskutiert, lassen sich die exakten dtn Nummern auf $\mathbb{C} \setminus \{\omega_1^2\}$ fortsetzen mittels:

$$dtn(\lambda) := i \sqrt{\omega_1^2 - \lambda} (a_\lambda \exp(i \sqrt{\omega_1^2 - \lambda} \cdot a) - b_\lambda \exp(-i \sqrt{\omega_1^2 - \lambda} \cdot a)). \quad (4.21)$$

Dabei bezeichnen a_λ, b_λ und c_λ die Lösung des Gleichungssystems (4.19) für $\kappa_\lambda^i := \sqrt{\omega_i^2 - \lambda}$. An dieser Stelle bleibt noch die Frage offen, wie die Gewichte ω_n (wir verzichten an dieser Stelle die Notation zu ändern) für eine möglichst gute Approximation des Operators $\mathcal{D}t\mathcal{N}$ zu wählen sind. Um diese Frage zu beantworten, werfen wir abermals einen Blick auf die Koeffizienten der nachfolgenden Reihendarstellung der expliziten Lösung auf Ω_{int} :

$$u(x, y) = \sum_{n=0}^{\infty} \left(a_n \exp(i \sqrt{\omega_1^2 - n^2} x) + b_n \exp(-i \sqrt{\omega_1^2 - n^2} x) \right) (u_0, v_n)_{L^2(0, \pi)} v_n(y) \quad (4.22)$$

Ohne Kenntnis der Koeffizienten a_n und b_n können wir in diesem Fall noch keine Aussagen über das genau Verhalten der Lösung treffen, da im Gegensatz zu Kapitel 4.2.1 auch ein exponentiell steigender Teil in den Koeffizienten in Darstellung (4.22) vorhanden ist. Somit bleibt uns nichts anderes übrig als die Matrix in (4.19) zu invertieren (diese bezeichnen wir im Folgenden mit M) um folgende Gleichung zu erhalten:

$$\begin{pmatrix} a_n \\ b_n \\ c_n \end{pmatrix} = M^{-1} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} m_{11} \\ m_{21} \\ m_{31} \end{pmatrix} \quad (4.23)$$

Dabei bezeichnen wir mit m_{ij} die Koeffizienten der inversen Matrix M^{-1} . Eine explizite Darstellung dieser ist am einfachsten mit Hilfe der Cramer'schen Regel anzugeben, welche wie folgt lautet:

$$m_{ij} = (-1)^{i+j} \frac{\det M_{ji}}{\det M}. \quad (4.24)$$

Dabei wiederum bezeichnet ganz allgemein M_{kl} die (Unter-) Matrix, die durch Streichen der k ten Zeile und l ten Spalte von M entsteht. Nach einer weiteren Rechnung erhalten wir die Koeffizienten

$$a_n = -\frac{i(\kappa_n^1 + \kappa_n^2)}{\det M} \exp(i(\kappa_n^2 - \kappa_n^1)x^*)$$

sowie

$$b_n = \frac{i(\kappa_n^2 - \kappa_n^1)}{\det M} \exp(i(\kappa_n^1 + \kappa_n^2)x^*).$$

Es bleibt noch die Determinante von M zu berechnen, welche nach der Regel von Sarrus wie folgt lautet:

$$\det M = i \exp(i\kappa_n^2 x^*) \left((\kappa_n^2 - \kappa_n^1) \exp(i\kappa_n^1(x^* - a)) - (\kappa_n^2 + \kappa_n^1) \exp(i\kappa_n^1(a - x^*)) \right). \quad (4.25)$$

Um die Koeffizienten in (4.22) zu berechnen, müssen wir im Gleichungssystem (4.19) $a = 0$ setzen, womit sich der Ausdruck in (4.25) ebenfalls etwas vereinfacht. Das Ziel ist nun zu zeigen, dass der Term bezüglich a_n dominant ist, da wir dann analog zu Kapitel 4.2.1 die Gewichtsfunktion definieren können. Zunächst ergibt sich durch Einsetzen von (4.25) in die Gleichungen der Koeffizienten

$$a_n = - \frac{(\kappa_n^1 + \kappa_n^2) \exp(-i\kappa_n^1 x^*)}{(\kappa_n^2 - \kappa_n^1) \exp(i\kappa_n^1 x^*) - (\kappa_n^1 + \kappa_n^2) \exp(-i\kappa_n^1 x^*)},$$

sowie

$$b_n = \frac{(\kappa_n^2 - \kappa_n^1) \exp(i\kappa_n^1 x^*)}{(\kappa_n^2 - \kappa_n^1) \exp(i\kappa_n^1 x^*) - (\kappa_n^1 + \kappa_n^2) \exp(-i\kappa_n^1 x^*)}.$$

Weiters nützen wir, dass für $n \geq |\Re(\omega_1)|$ die Terme $\exp(i\kappa_n^1(x^* - z))$ für $z < x^*$ exponentiell gegen Null streben, womit wir im Grenzübergang $n \rightarrow \infty$ das folgende Verhalten bekommen

$$a_n \rightarrow 1, \quad b_n \rightarrow 0.$$

Um den Koeffizienten bezüglich b_n in (4.22) vernachlässigen zu können, reicht die obige Tatsache allein aber nicht aus. Mit einer weiteren Rechnung sehen wir jedoch, dass

$$b_n \exp(-i\sqrt{\omega_1^2 - n^2} \cdot x) = \frac{(\kappa_n^2 - \kappa_n^1) \exp(i\kappa_n^1(x^* - x))}{(\kappa_n^2 - \kappa_n^1) \exp(i\kappa_n^1 x^*) - (\kappa_n^1 + \kappa_n^2) \exp(-i\kappa_n^1 x^*)},$$

womit wir zusammengefasst ebenso

$$b_n \exp(-i\sqrt{\omega_1^2 - n^2} \cdot x) \approx 0 \quad \text{für} \quad n \geq |\Re(\omega_1)|, 0 < x < x^* \quad (4.26)$$

erhalten. Betrachten wir nun die Darstellung der analytischen Lösung auf dem beschränkten Gebiet $\Omega_{\text{int}} := (0, a) \times (0, \pi)$

$$u(x, y) = \sum_{n=0}^{\infty} \left(a_n \exp(i\sqrt{\omega_1^2 - n^2}x) + b_n \exp(-i\sqrt{\omega_1^2 - n^2}x) \right) (u_0, v_n)_{L^2(0, \pi)} v_n(y), \quad (4.27)$$

so haben wir schlussendlich gezeigt, dass für $n > |\Re(\omega_1)|$ die Koeffizienten der Reihe entlang des Wellenleiters (für $x \rightarrow \infty$) exponentiell gedämpft werden. Dies rechtfertigt nun die folgende Wahl der Gewichtsfunktion zur Bestimmung des Optimierungsproblems:

$$w_n = \begin{cases} 1, & n \leq |\Re(\omega_1)| \\ |a_n \exp(i\sqrt{\omega_1^2 - n^2} \cdot a) + b_n \exp(-i\sqrt{\omega_1^2 - n^2} \cdot a)|, & n > |\Re(\omega_1)|. \end{cases}$$

Überdies wissen wir ebenfalls auf welche Frequenzen der Fokus für die Approximation des Operators \mathcal{DtN} gelegt werden muss. Wie bereits erklärt, entscheiden wir uns somit aufgrund des exponentiell fallenden Verhaltens für $n \geq m := \lceil \Re(\omega_1) \rceil$ den Fokus auf die ersten $L_{\max} := \lfloor m \rfloor + 2$ Koeffizienten zu legen. Die Ergebnisse der approximierten DtN Nummern befinden sich in der nachfolgenden Abbildung: Wir sehen also, dass die entstandene \mathcal{dtn}

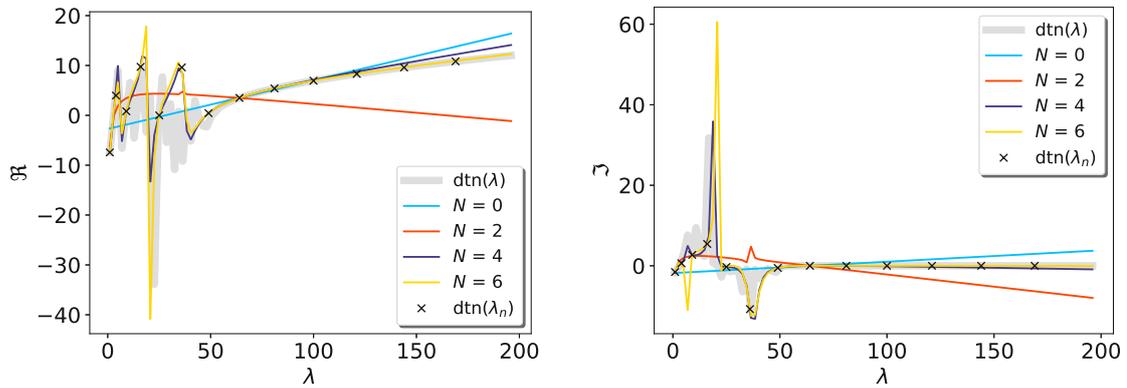


Abbildung 4.4: Approximation der $\mathcal{dtn}(\cdot)$ Funktion mittels gelernter DtN Nummern, Darstellung des Realteils links sowie des Imaginärteils rechts für $\lambda \in (0, 200)$, Wellenzahlen $\omega_1 = 7.2$ und $\omega_2 = 2.5$

Funktion einer komplexeren Bauart ist, als in dem Modellproblem zuvor, da für $\lambda < |\Re(\omega_1)|$ der Koeffizient b_n nicht verschwindet. Andererseits gilt für $\lambda > |\Re(\omega_1)|$ wie in (4.26) gezeigt gerade $\mathcal{dtn}(\lambda) \approx i\sqrt{\omega_1^2 - \lambda}$ womit wir in jenem Bereich im Wesentlichen die gleiche \mathcal{dtn} Funktion wie im Modellproblem 4.2.1 erhalten. In Abbildung 4.4 ist nun ebenfalls zu erkennen, dass erst für $N_{\text{ext}} \geq 4$ die approximierten DtN Nummern „brauchbare“ Resultate liefern. Dies ist natürlich mit der komplexeren Bauart der \mathcal{dtn} Funktion zu begründen. Es ist also davon auszugehen, dass der L^2 Fehler in den nachfolgenden Experimenten für kleine N_{ext} groß sein wird. Mit den gelernten Matrizen A, B gehen wir nun analog zum letzten Abschnitt vor und betrachten Lösungen von (4.14) wobei wir die stückweise konstante Wellenzahl wie folgt definieren:

$$\omega(x) = \begin{cases} 7.2 & , \quad x \leq 3\pi \\ 2.5 & , \quad x > 3\pi. \end{cases}$$

Zusätzlich verwenden wir die selben Randdaten wie zuvor, setzen also

$$u_0 = \sum_{n=1}^{10} \cos(n \cdot), \quad (4.28)$$

womit wir als Referenzlösung auf $\Omega_{\text{int}} := (0, 2\pi) \times (0, \pi)$ die folgende Lösung erhalten:

$$u(x, y) = \sum_{n=1}^{10} \left(a_n \exp(i\sqrt{(7.2)^2 - n^2} \cdot x) + b_n \exp(-i\sqrt{(7.2)^2 - n^2} \cdot x) \right) \cos(ny). \quad (4.29)$$

Die Konstanten a_n und b_n sind wie bereits erklärt über das Gleichungssystem (4.19) mit $a = 0$ und $x^* = 3\pi$ zu berechnen. Analog zu dem vorausgehenden Abschnitt untersuchen wir in den numerischen Experimenten den Fehler $\|u - u_h\|_{L^2(\Omega_{\text{int}})}$ in Abhängigkeit der Anzahl an externen Freiheitsgraden N_{ext} und $n := \dim X_h$. Die Resultate befinden sich in Abbildung 4.5. Wie bereits aus dem Verhalten der approximierten DtN Nummern in Ab-

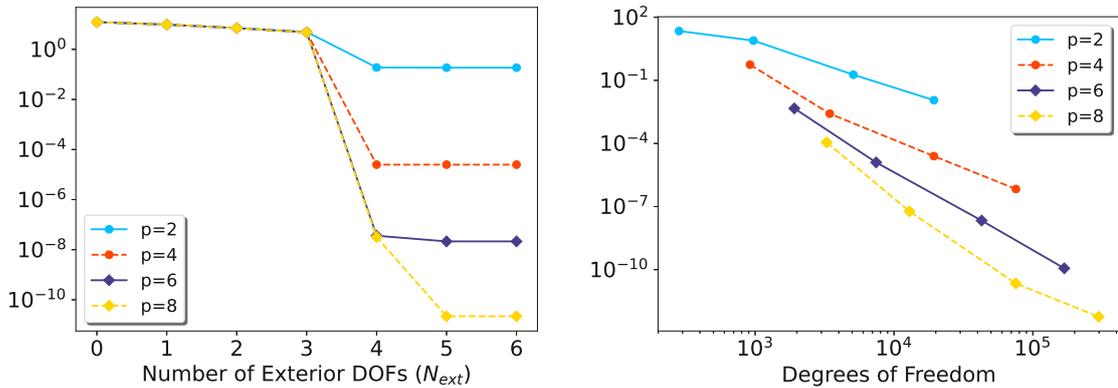


Abbildung 4.5: Links: $L^2(\Omega_{\text{int}})$ -Fehler für $h = 0.1$, Rechts: $L^2(\Omega_{\text{int}})$ -Fehler für $N_{\text{ext}} = 5$

bildung 4.4 zu erkennen war, konvergiert die approximierte Lösung u_h erst für $N_{\text{ext}} \geq 4$ gegen die Referenzlösung. Für N_{ext} hinreichend groß erhalten wir aber ähnliche gute Konvergenzraten wie in dem Beispiel zuvor mit einer konstanten Wellenzahl. Dieses Verhalten ist nicht sonderlich überraschend, da wir aufgrund der Regularität der exakten Lösung für steigende Polynomordnungen p auch bessere Konvergenzraten erwarten konnten, sofern die Approximation des DtN Operators als hinreichend gut vorausgesetzt ist.

Natürlich kann man die obigen Argumentationen für beliebige stückweise konstanten Wellenzahlen $\omega(x)$ mit etwas mehr Aufwand verallgemeinern. Aus physikalischen Gründen wollen wir abschließend noch Wellenzahlen der folgenden Bauart betrachten

$$\omega(x) = \begin{cases} \omega_1 & , \quad x \leq x^* \\ \omega_2 & , \quad x^* < x < \hat{x} \\ \omega_3 & , \quad x > \hat{x} \end{cases}$$

wobei wir wieder annehmen, dass die Sprünge der Wellenzahlen außerhalb des Gebietes $\Omega_{\text{int}} := (0, a) \times (0, \pi)$ stattfinden, insbesondere gilt also $a < x^* < \hat{x}$. Wie zuvor, suchen wir also erneut Lösungen des folgenden (eindimensionalen) Problems:

$$w_n(x) = \begin{cases} a_n \exp(i\sqrt{\omega_1^2 - \lambda_n}x) + b_n \exp(-i\sqrt{\omega_1^2 - \lambda_n}x) & , \quad x \leq x^* \\ c_n \exp(i\sqrt{\omega_2^2 - \lambda_n}x) + d_n \exp(-i\sqrt{\omega_2^2 - \lambda_n}x) & , \quad x^* < x \leq \hat{x} \\ e_n \exp(i\sqrt{\omega_3^2 - \lambda_n}x) & , \quad x > \hat{x} \end{cases} \quad (4.30)$$

Um eine stetig differenzierbare Referenzlösung u zu erhalten, muss für die Koeffizienten

das folgende Gleichungssystem gelten:

$$\begin{pmatrix} e^{i\kappa_n^1 a} & e^{-i\kappa_n^1 a} & 0 & 0 & 0 \\ e^{i\kappa_n^1 x^*} & e^{-i\kappa_n^1 x^*} & -e^{i\kappa_n^2 x^*} & -e^{i\kappa_n^2 x^*} & 0 \\ i\kappa_n^1 e^{i\kappa_n^1 x^*} & -i\kappa_n^1 e^{-i\kappa_n^1 x^*} & -i\kappa_n^2 e^{i\kappa_n^2 x^*} & i\kappa_n^2 e^{-i\kappa_n^2 x^*} & 0 \\ 0 & 0 & e^{i\kappa_n^2 \hat{x}} & e^{-i\kappa_n^2 \hat{x}} & -e^{i\kappa_n^3 \hat{x}} \\ 0 & 0 & i\kappa_n^2 e^{i\kappa_n^2 \hat{x}} & -i\kappa_n^2 e^{-i\kappa_n^2 \hat{x}} & -i\kappa_n^3 e^{i\kappa_n^3 \hat{x}} \end{pmatrix} \begin{pmatrix} a_n \\ b_n \\ c_n \\ d_n \\ e_n \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Dabei bezeichnet wie schon zuvor aus Gründen der besseren Lesbarkeit $\kappa_n^i := \sqrt{\omega_i^2 - \lambda_n}$. Die lineare Unabhängigkeit der Spalten kann analog zu dem von (4.19) argumentiert werden. Wir folgern, dass für $\omega_1^2 \neq \lambda_n$ und $\omega_2^2 \neq \lambda_n$ eine eindeutige Lösung existiert. Der exakte DtN Operator kann somit analog durch (4.20) beschrieben werden, wobei die Koeffizienten a_n, b_n als Lösungen des obigen Gleichungssystems zu entnehmen sind. Daraus sind die exakten dtn Nummern ebenfalls gegeben durch

$$dtn(\lambda) := i\sqrt{\omega_1^2 - \lambda}(a_\lambda \exp(i\sqrt{\omega_1^2 - \lambda} \cdot a) - b_\lambda \exp(-i\sqrt{\omega_1^2 - \lambda} \cdot a)), \quad (4.31)$$

wobei wir dieses Mal $\lambda \in \mathbb{C} \setminus \{\omega_1^2, \omega_2^2\}$ voraussetzen müssen, damit die Funktion wohl-

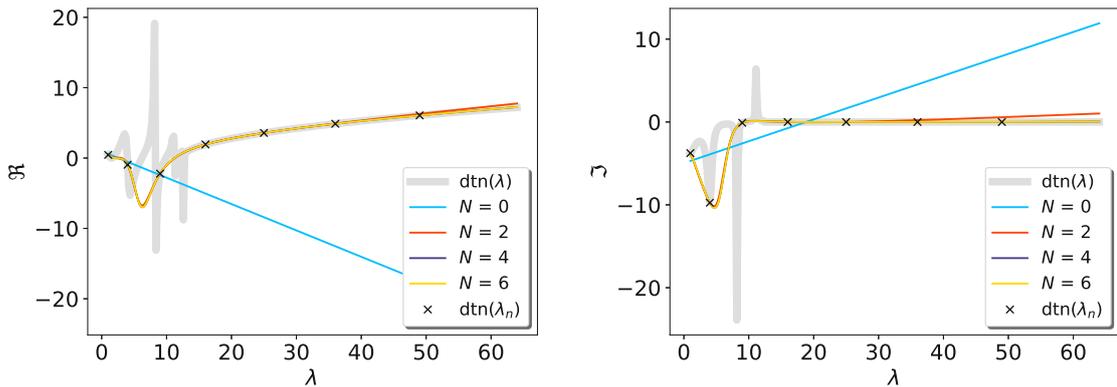


Abbildung 4.6: Approximation der $dtn(\cdot)$ Funktion mittels gelernter DtN Nummern, Darstellung des Realteils links sowie des Imaginärteils rechts für $\lambda \in (0, 60)$, Wellenzahlen $\omega_1 = 3.5$ und $\omega_2 = 7.1$

definiert ist. Wie man schon erahnen kann, kompliziert sich die exakte Berechnung der Koeffizienten a_n, b_n um Einiges, da für 5×5 Matrizen die Darstellung mittels Kofaktoren sehr unübersichtlich wird. Wir ersparen uns deswegen langwierige Rechnungen und definieren ausgehend zu den vorangehenden Überlegungen die folgende Gewichtsfunktion:

$$w_n = \begin{cases} 1, & n \leq |\Re(\omega_1)| \\ |a_n \exp(i\sqrt{\omega_1^2 - \lambda_n} \cdot a) + b_n \exp(-i\sqrt{\omega_1^2 - \lambda_n} \cdot a)|, & n > |\Re(\omega_1)| \end{cases}$$

Um rigoros vorzugehen, müsste man an dieser Stelle natürlich zeigen, dass

$$b_n \exp(-i\sqrt{\omega_1^2 - \lambda_n} \cdot a) \approx 0, \quad n > |\Re(\omega_1)|$$

gilt. Tatsächlich kann man dieses Verhalten aber in Abbildung 4.6 numerisch überprüfen, da wie man ablesen kann für $\lambda > |\Re(\omega_1)|$ die exakten dtN Nummern approximativ durch $dtN(\lambda) \approx i\sqrt{\omega_1 - \lambda}$ gegeben sind. Wir entscheiden uns daher erneut, die ersten $L_{\max} := \lfloor m \rfloor + 2$ Frequenzen mit $m := |\Re(\omega_1)|$ zu approximieren. In diesem Fall liegen die approximierten DtN Nummern bereits für wenige Freiheitsgrade N_{ext} nahe an der exakten dtN Funktion. Da in den Experimenten die Wellenzahl ω_1 kleiner als zuvor gewählt worden ist, resultiert dies auch in einem kleineren Bereich Λ , bei dem b_λ nicht verschwindet. Anschaulich gesprochen muss somit nur ein kleinerer Bereich der dtN Funktion hinreichend gut approximiert werden, wonach ebenfalls die Anzahl an externen Freiheitsgrade N_{ext} kleiner gewählt werden kann.

Für die Lösung des Helmholtz Problems entscheiden wir uns wieder für die selben Randbedingungen wie zuvor, setzen also

$$u_0 = \sum_{n=1}^{10} \cos(n \cdot).$$

Weiters verwenden wir die stückweise konstante Wellenzahl wie in Abbildung 4.6 vermerkt

$$\omega(x) = \begin{cases} 3.5 & , \quad x \leq 2\pi \\ 7.1 & , \quad 2\pi < x < 3\pi , \\ 3.5 & , \quad x > 3\pi \end{cases}$$

womit wir als Referenzlösung auf $\Omega_{\text{int}} := (0, \pi) \times (0, \pi)$ die folgende Lösung erhalten:

$$u(x, y) = \sum_{n=1}^{10} \left(a_n \exp(i\sqrt{(3.5)^2 - n^2} \cdot x) + b_n \exp(-i\sqrt{(3.5)^2 - n^2} \cdot x) \right) \cos(ny).$$

Wie bereits argumentiert, konvergieren die ermittelten Approximationen bereits für eine geringe Anzahl an äußeren Freiheitsgraden gegen die exakte Lösung. Weiters erkennen

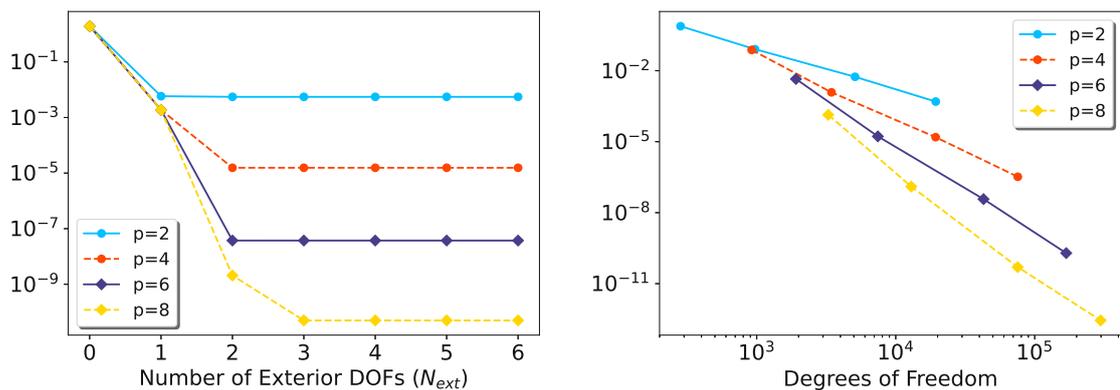


Abbildung 4.7: Links: $L^2(\Omega_{\text{int}})$ -Fehler für $h = 0.1$, Rechts: $L^2(\Omega_{\text{int}})$ -Fehler für $N_{\text{ext}} = 4$

wir hohe Konvergenzraten bei Verwendung Polynome hohen Grades. Zusammengefasst

ist somit hervorzuheben, dass auch für stückweise konstante Wellenzahlen mit Hilfe der Learned Infinite Elemente Methode sehr gute Ergebnisse erzielt wurden. Dabei gestaltete sich insbesondere die Berechnung der exakten dtN Nummern ein wenig komplizierter als zuvor. Im nächsten Abschnitt werden wir schlussendlich noch Probleme betrachten, bei denen wir in Ω_{int} keine expliziten Lösungen mehr angeben können.

4.2.3 Wellenleiter mit Streukörpern

Nachdem wir im letzten Abschnitt die Learned Infinite Elemente Methode auf die Helmholtz Gleichung mit stückweise konstanter Wellenzahlen $\omega(x)$ angewendet haben, wollen wir nun die zu Grunde liegende Geometrie verändern. Genauer platzieren wir einen sogenannten Streukörper K innerhalb des Wellenleiters. Dies hat zur Folge, dass das zu betrachtende Gebiet Ω_{int} dann nicht mehr konvex ist. Je nach der Form des Streukörpers hat dies ebenso Konsequenzen in Bezug auf die Regularität der analytischen Lösung u , welche maßgeblich für die hohen Konvergenzraten in den bisherigen Beispielen verantwortlich war. Wir können also ad-hoc nicht die gleichen Konvergenzraten wie in den vorausgehenden Beispielen erwarten. In den nachfolgenden Experimenten „skalieren“ wir außerdem den zu



Abbildung 4.8: Wellenleiter mit Streukörper inkl. Setzen mehrerer transparenter Ränder

Grunde liegenden Wellenleiter, indem wir die Breite auf 1 reduzieren, bzw. genauer gesagt $\Gamma := \{a\} \times (0, 1)$ setzen. Dies hat zur Folge, dass die Eigenfunktionen von $-\Delta_{\Gamma}$ durch $v_n = \cos(n\pi \cdot)$ gegeben sind, womit sich die Eigenwerte um den Faktor π^2 ändern und sich das Spektrum somit wie folgt berechnet:

$$\sigma(-\Delta_{\Gamma}) = \{n^2\pi^2 : n \in \mathbb{N}_0\}.$$

Wir bemerken weiters, dass wir durch den Streukörper K keine explizite Darstellung der analytischen Lösung u auf ganz Ω mehr angeben können. Diese war jedoch für die Wahl der Gewichtsfunktion ω_n in (4.12) entscheidend. Bei genauerem Hinsehen ist dies aber nicht weiter störend, denn an der Darstellung der Lösung des Außenraumproblems auf Ω_{ext} ändert der Streukörper zunächst nichts. Wir wissen weiterhin, dass die Lösung u_{ext} auf Ω_{ext} gegeben ist durch:

$$u_{\text{ext}}(x, y) = \sum_{n=0}^{\infty} \exp(i\sqrt{\omega^2 - \pi^2 n^2} \cdot x) \cdot c_n \cdot v_n(y), \quad (4.32)$$

wobei die Koeffizienten c_n von den „Randdaten“ $u|_{\Gamma}$ abhängen. Führen wir nun wie in Abbildung 4.8 veranschaulicht gedanklich ein einen weiteren Rand $\tilde{\Gamma}$ ein, so können wir jede beliebige Funktion aus $L^2(\tilde{\Gamma})$ schreiben als

$$u|_{\tilde{\Gamma}} = \sum_{n \in \mathbb{N}_0} (u|_{\tilde{\Gamma}}, v_n)_{L^2(0,1)} v_n \quad (4.33)$$

und wissen durch (4.32), wie die einzelnen Moden zu Γ „transportiert“ werden. Je nach Lage des transparenten Randes erwarten wir also selbiges exponentielles Verhalten wie schon in den vorherigen Abschnitten diskutiert, ohne jedoch die genauen Koeffizienten aus (4.33) zu kennen. Setzen wir überdies $\Gamma := \{a\} \times (0, 1)$ und $\tilde{\Gamma} := \{b\} \times (0, 1)$, können wir ausgehend von den obigen Überlegungen die Gewichtsfunktion für das Optimierungsproblem wie folgt definieren

$$\omega_n = \begin{cases} 1, & n^2\pi^2 \leq |\Re(\omega^2)| \\ |\exp(i\sqrt{\omega^2 - n^2\pi^2} \cdot (a - b))|, & n^2\pi^2 > |\Re(\omega^2)| \end{cases} \quad (4.34)$$

In den numerischen Experimenten müssen wir noch die Frage klären welche Referenzlösung zur Messung des Fehlers $\|u - u_h\|_{L^2(\Omega_{\text{int}})}$ herangezogen wird. Dazu verwenden wir ebenfalls wie in Abbildung 4.8 dargestellt einen weiteren transparenten Rand Γ_{ref} , über den wir durch Setzen einer hohen Polynomordnung $p = 8$ und feinen mesh size $h = 0.01$ eine Referenzlösung auf einem Gebiet $\Omega_{\text{ref}} \supset \Omega_{\text{int}}$ erhalten. Für die Simulationen wählen wir

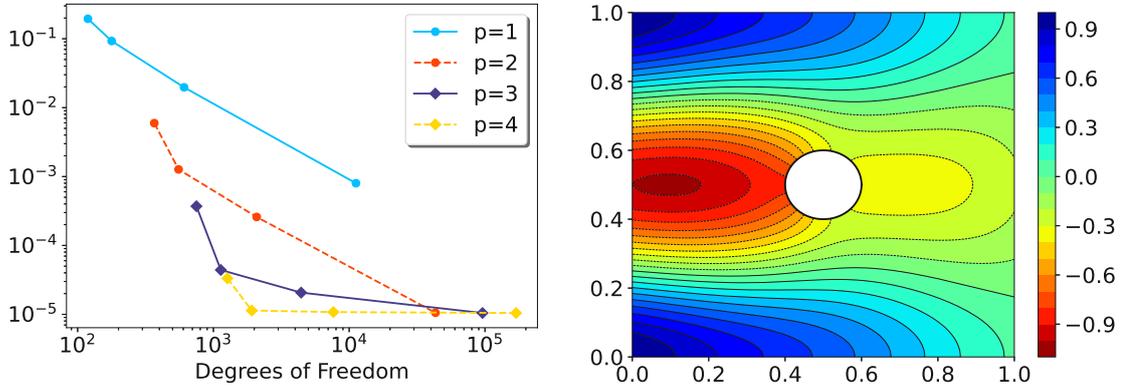


Abbildung 4.9: Links: $L^2(\Omega_{\text{int}})$ -Fehler für $N_{\text{ext}} = 7$, Rechts: Darstellung des Realteils der approximierten Lösung $\Re(u_h)$, Wellenzahl $\omega = 6.5$

nun als Streukörper K einen Kreis mit Mittelpunkt $(0.5, 0.5)$ und Radius 0.1 und erhalten die nachfolgenden Geometrien:

$$[(0, 2) \times (0, 1)] \setminus K =: \Omega_{\text{ref}} \supset \Omega_{\text{int}} := [(0, 1) \times (0, 1)] \setminus K. \quad (4.35)$$

Für die Gewichtsfunktion in (4.34) können wir somit $a = 1$ und $b = 0.6$ wählen. Weiters setzen wir die Dirichlet Daten auf Γ_D als $u_0(y) = \cos(2\pi y)$ voraus. In Abbildung 4.9 können wir für grobe mesh sizes h steigende Konvergenzgeschwindigkeit für steigende Polynomordnungen erkennen. Je feiner das mesh gewählt wurde, desto weniger unterschieden sich die Fehler in Abhängigkeit der Ordnung, was natürlich ebenfalls am Fehler der Referenzlösung selbst liegt. Durch Visualisieren der Referenzlösung kann man erkennen, dass die Lösung in der Nähe des Streukörpers ein wenig an Regularität verliert, wonach der Fehler der diskreten Lösung in jenem Bereich auch dominant ist. Um bessere Approximationen zu erhalten, könnte man das mesh in jenem Bereich zusätzlich verfeinern, was wir jedoch an dieser Stelle nicht weiter überprüfen wollen.

Stattdessen wollen wir noch eine physikalische Interpretation des Streukörpers anführen. Betrachten wir die sogenannte „Eingangsschwingung“ u_0 auf Γ_D , die sich entlang des Wellenleiters fortbewegt, so erkennen wir, dass diese Schwingung durch den Streukörper K gebrochen wird. Dieses Phänomen ist jedoch nur bedingt in Abbildung 4.9 zu sehen, da die Wellenzahl ω so klein gewählt ist, dass größere Frequenzen exponentiell gedämpft werden. Um dieses Verhalten besser zu illustrieren, wenden wir das Learned Infinite Elemente Ver-

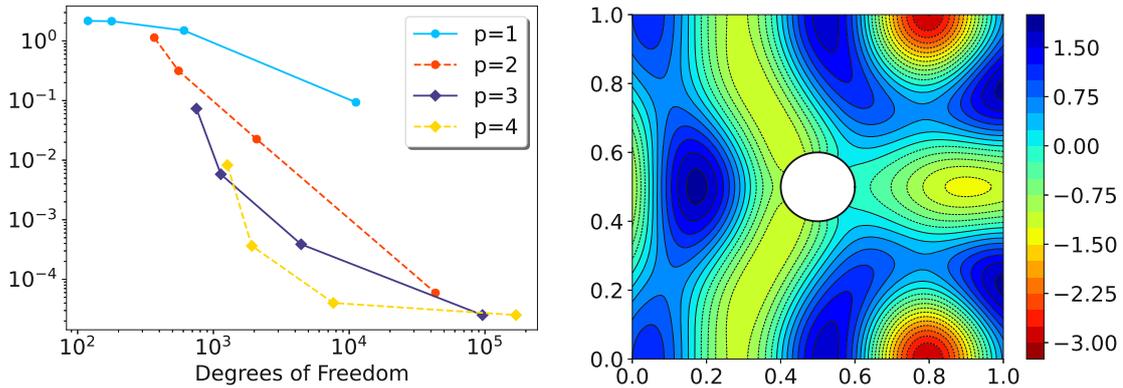


Abbildung 4.10: Links: $L^2(\Omega_{int})$ -Fehler für $N_{ext} = 7$, Rechts: Darstellung des Realteils der approximierten Lösung $\Re(u_h)$, Wellenzahl $\omega = 13.5$

fahren nochmal auf dieselbe Geometrie Ω_{int} an, indem wir eine größere Wellenzahl $\omega = 13.5$ und als Randdaten ebenfalls $u_0(y) = \cos(2\pi y)$ setzen. Die Ergebnisse der Simulationen befinden sich in Abbildung 4.10 und zeigen nun exakt jenes „Brechen“ der Wellen anhand des Streukörpers, das wir bereits erahnt haben. In Abbildung 4.9 können wir nun überdies ein ähnliches Verhalten wie in dem Experiment zuvor erkennen: Für steigende Polynomordnungen flacht der Fehler zunehmend bei 10^{-4} ab, was abermals durch den Fehler der Referenzlösung selbst begründbar ist. Schlussendlich wollen wir noch die Problemstellung aus Kapitel 4.2.2 im Kontext von Wellenleitern mit Streukörpern betrachten. Dazu betrachten wir die folgende stückweise konstante Wellenzahl

$$\omega(x) = \begin{cases} 15.2 & , \quad x \leq 3 \\ 10.5 & , \quad x > 3 \end{cases}$$

mit dem zu Grunde liegenden Gebiet $\Omega_{int} := [(0, 1) \times (0, 1)] \setminus K$, wobei

$$K := B_{0.075}(0.25, 0.25) \cup B_{0.075}(0.25, 0.5) \cup B_{0.075}(0.25, 0.75)$$

bezeichnet, siehe auch Abbildung 4.11. Zur Bestimmung einer geeigneten Gewichtsfunktion nützen wir nun wieder die Tatsache aus, dass sich für $x > 0.325$ die analytische Lösung wie in Formel (4.27) verhält und wir somit die Wahl

$$\omega_n = \begin{cases} 1, & n^2\pi^2 \leq (15.2)^2 \\ |a_n e^{i\sqrt{(15.2)^2 - n^2\pi^2} \cdot 0.675} + b_n e^{-i\sqrt{(15.2)^2 - n^2\pi^2} \cdot 0.675}|, & n^2\pi^2 > (15.2)^2 \end{cases}$$

rechtfertigen können. Dabei sind die Koeffizienten a_n, b_n aus dem Gleichungssystem (4.19) mit $a = 0.325$ und $x^* = 3$ zu berechnen. Die nachfolgende Abbildung zeigt, dass sich die Resultate in Bezug auf die Konvergenzgeschwindigkeit ähnlich wie in dem vorausgehenden Beispiel verhalten. Weiters ist gut zu sehen, wie „einfallende Schwingung“ $u_0 = \cos(\pi y)$ entlang des Streukörpers in kleinere Frequenzen gebrochen wird, welche sich dann unterschiedlich entlang des Wellenleiters fortbewegen. Zusammenfassend stellen wir also fest,

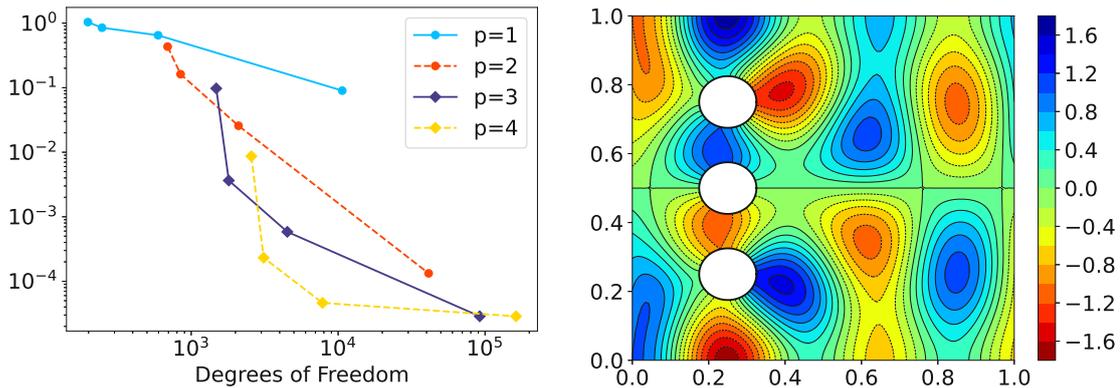


Abbildung 4.11: Links: $L^2(\Omega_{int})$ -Fehler für $N_{ext} = 7$, Rechts: Darstellung des Realteils der approximierten Lösung $\Re(u_h)$, Wellenzahlen $\omega_1 = 15.2$ und $\omega_2 = 10.5$

dass die in [HLP21] vorgestellte Learned Infinite Elemente Methode sehr gute Resultate hinsichtlich Konvergenz für die vorgestellten Problemstellungen lieferte. Es sei jedoch angemerkt, dass für Wellenleiter die Berechnung der exakten Eigenwerte von $-\Delta_\Gamma$ und in weiterer Hinsicht auch das Aufstellen des exakten DtN Operators vergleichsweise einfach war. Der wohl schwierigste und auch entscheidende Schritt in allen Beispielen war eine geeignete Wahl der Gewichtsfunktion ω_n und der Anzahl L_{max} der zu approximierenden DtN Nummern $\mathcal{A}tn(\lambda_n)$. Um diesen Schritt korrekt durchführen zu können, ist in jedem Fall eine Interpretation der Darstellung des Operators $\mathcal{D}t\mathcal{N}$ und des Verhaltens der Lösung des Außenraumproblems notwendig.

5 Anwendung auf Eigenwertprobleme

Während wir uns im letzten Kapitel im Wesentlichen mit dem Lösen von Streuproblemen beschäftigt haben, wollen wir abschließend die bisherigen Erkenntnisse nutzen, um Eigenwertprobleme auf unbeschränkten Gebieten zu lösen. Dazu werden wir zunächst in einem abstrakten Setting einen Algorithmus zur Berechnung von Eigenwerten präsentieren und dann daran anschließend diskutieren, wie wir diesen in Kombination mit der Learned Infinite Elemente Methode verwenden können. Abschließend werden wir diese neue Methode an verschiedenen Problemen testen und analysieren.

5.1 Integralmethode

In diesem Abschnitt werden wir die in [Bey12] vorgestellte, sogenannte Integralmethode zur Berechnung von Eigenwerten vorstellen. Diese vergleichsweise neue Methode hat sich als überaus nützlich bei nicht-linearen Eigenwertproblemen erwiesen, wodurch wir im Folgenden die Idee des oben genannten Verfahrens erläutern wollen. Für Details verweisen wir auf weiterführende Literatur wie [Nan22, Kapitel 8.3] bzw. [Bey12], auf deren Resultate dieses Kapitel auch aufbaut.

Betrachten wir für den Rest des Abschnitts das folgende Problem:

$$A(\lambda)u = 0, \quad u \in \mathbb{C}^n, u \neq 0, \lambda \in \mathbb{C}. \quad (5.1)$$

Dabei setzen wir den Operator $A : \Lambda \rightarrow \mathbb{C}^{n \times n}$ als holomorph auf einem Gebiet $\Lambda \subset \mathbb{C}$ voraus. Zunächst wollen wir den folgenden Begriff definieren:

Definition 5.1. Sei $A : \Lambda \rightarrow \mathbb{C}^{n \times n}$ wie oben definiert. Dann ist $\lambda \in \mathbb{C}$ ein *halb-einfacher Eigenwert*, wenn eine Orthonormalbasis $v_{1,1}, \dots, v_{1,L}$ von $\ker A(\lambda)$ aus Eigenvektoren existiert, sodass ebenfalls $A'(\lambda)v_{1,l} \notin A(\lambda)(\mathbb{C}^n)$ gilt. Als Eigenvektoren bezeichnen wir dabei $u \in \ker A(\lambda) \setminus \{0\}$.

Ausgangspunkt der folgenden Überlegungen ist der *Satz von Keldysh*, welcher einfach ausgedrückt die Resolvente des Operators in Form einer Laurent Reihe darstellt.

Satz 5.2 (Satz von Keldysh). *Sei $\mathbb{C} \supset \Lambda \ni \lambda \mapsto A(\lambda)$ holomorph und es existiert zumindest ein $\lambda \in \Lambda$ sodass die Inverse von $A(\lambda)$ existiert. Weiters sei λ_1 ein halb-einfacher Eigenwert mit Orthonormalbasis $v_{1,1}, \dots, v_{1,L} \in \ker A(\lambda_1) \setminus \{0\}$.*

Dann ist $\dim \ker(A^H(\lambda_1)) = L$ und es existiert eine Basis $w_{1,1}, \dots, w_{1,L}$ von $\ker(A^H(\lambda_1))$ sodass überdies

$$w_{1,l}^H A'(\lambda_1) v_{1,k} = \delta_{lk}$$

gilt. Weiters existiert eine Umgebung $\lambda_1 \in U \subset \mathbb{C}$, sodass für $\lambda \in U$

$$A(\lambda)^{-1} = (\lambda - \lambda_1)^{-1} \sum_{l=1}^L v_{1,l} w_{1,l}^H + R(\lambda),$$

wobei $\lambda \mapsto R(\lambda) \in \mathbb{C}^{n \times n}$ holomorph ist.

Für einen Beweis verweisen wir abermals auf [Bey12, Theorem 2.4]. Obwohl der obige Satz auf den ersten Blick sehr technisch aussieht, bildet er die Grundlage der Integralmethode. Sei nun nämlich Λ ein beschränktes Gebiet in \mathbb{C} und habe $\lambda \mapsto A(\lambda)$ in Λ nur endlich viele Eigenwerte $\lambda_1, \dots, \lambda_k \in \Lambda$, welche zusätzlich halb-einfach sind mit $\dim \ker A(\lambda_i) = L_i$ und Orthonormalbasis $\{v_{i1}, \dots, v_{iL_i}\}$, dann können wir $A(\lambda)^{-1}$ wie folgt schreiben:

$$A(\lambda)^{-1} = \sum_{n=1}^k (\lambda - \lambda_n)^{-1} \sum_{l=1}^{L_n} v_{n,l} w_{n,l}^H + R(\lambda), \quad \lambda \in \Lambda \setminus \{\lambda_1, \dots, \lambda_k\}.$$

Dabei ist nun $\lambda \mapsto R(\lambda)$ eine in ganz Λ holomorphe Funktion. Sei nun Γ eine positiv orientierte Jordan Kurve, die ganz in $\Lambda \setminus \{\lambda_1, \dots, \lambda_k\}$ liegt und in deren Inneren die Eigenwerte $\{\lambda_1, \dots, \lambda_k\}$ liegen, dann gilt mit Hilfe des Residuensatzes (2.1) für beliebige holomorphe Funktionen $\lambda \mapsto f(\lambda) \in \mathbb{C}$ die folgende Identität:

$$\frac{1}{2\pi i} \int_{\Gamma} f(\lambda) A(\lambda)^{-1} d\lambda = \sum_{n=1}^k f(\lambda_n) \sum_{l=1}^{L_n} v_{n,l} w_{n,l}^H. \quad (5.2)$$

Es bleibt an der Stelle die Frage offen, welche Wahl der Funktionen f zielführend ist. Bevor wir der Beantwortung jener Frage nachgehen, spezifizieren wir nun zuerst ein wenig die Voraussetzungen. Für den Rest dieses Abschnitts sei Γ wie oben definiert eine positiv orientierte Jordan Kurve, die ganz in der Resolventenmenge $\rho(A)$ liegt und in deren Inneren nur endlich viele halb-einfache Eigenwerte $\lambda_1, \dots, \lambda_k \in \sigma(A)$ mit Vielfachheiten L_k liegen. Mit der oben eingeführten Notation definieren wir nun mit den Eigenvektoren $v_{ij} \in \ker A(\lambda_i)$ bzw. Eigenvektoren des adjungierten Problems $w_{ij} \in \ker A^H(\lambda_i)$ die folgenden Matrizen

$$V := (v_{1,1}, \dots, v_{kL_k}) \in \mathbb{C}^{n \times L}, \quad W := (w_{1,1}, \dots, w_{kL_k}) \in \mathbb{C}^{n \times L},$$

wobei $L := \sum_{n=1}^k L_n$ bezeichnet. Als nächsten setzen wir $\hat{V} \in \mathbb{C}^{n \times J}$ als Zufallsmatrix voraus, wobei zumindest $L \leq J$ und typischerweise $J \ll n$ gelten muss. Mit diesen Vorbereitungen können wir nun die folgenden Matrizen definieren:

$$A_0 := \frac{1}{2\pi i} \int_{\Gamma} A(\lambda)^{-1} \hat{V} d\lambda, \quad A_1 := \frac{1}{2\pi i} \int_{\Gamma} \lambda A(\lambda)^{-1} \hat{V} d\lambda. \quad (5.3)$$

Man sieht unmittelbar, dass $A_0, A_1 \in \mathbb{C}^{n \times J}$ gilt. Nun gilt aber mit der Identität (5.2), indem wir $f(\lambda) = 1$ bzw. $f(\lambda) = \lambda$ setzen und der Tatsache, dass \hat{V} konstant ist:

$$A_0 = V W^H \hat{V}, \quad A_1 = V \text{diag}(\lambda_1, \dots, \lambda_k) W^H \hat{V}. \quad (5.4)$$

In diesen Matrizen kommen nun erstmals die gesuchten Eigenwerte des Operators A vor. Angenommen es gelingt uns die Matrizen A_0 und A_1 zu berechnen, so bleibt lediglich die Frage offen wie man schlussendlich an jene Eigenwerte kommt.

Um diese Frage beantworten zu können, müssen wir zunächst eine zusätzliche Forderung an die Eigenvektoren $v_{ij} \in \ker(A(\lambda_i) \setminus \{0\})$ stellen. Diese besteht darin, dass wir voraussetzen, dass V bzw. W (also die Matrizen bestehend aus den Eigenvektoren) jeweils vollen Spaltenrang L haben. Diese Einschränkung mag vielleicht plausibel erscheinen, jedoch müssen bei allgemeinen nicht-linearen Eigenwertproblemen die Eigenvektoren verschiedener Eigenwerte nicht unbedingt linear unabhängig sein. Bei genauerem Nachdenken ist dies jedoch trotzdem keine wesentliche Einschränkung, da man notfalls immer zu einer (kleineren) Kurve $\tilde{\Gamma}$ übergehen kann, in deren Inneren weniger oder auch gar nur ein einziger Eigenwert liegt. Mit jener zusätzlichen Voraussetzung ist nun sichergestellt, dass $V, W \in \mathbb{C}^{n \times L}$ sowie ebenfalls $\hat{V} \in \mathbb{C}^{n \times J}$ als Zufallsmatrix maximalen Rang haben, womit sich diese Eigenschaft unmittelbar auf die Matrix $A_0 \in \mathbb{C}^{n \times J}$ überträgt. Bilden wir nun eine Singulärwertzerlegung der Matrix A_0 , so erhalten wir die folgende Darstellung

$$A_0 = \tilde{V} \Sigma \tilde{W}^H. \quad (5.5)$$

Dabei sind die Matrizen $\tilde{V} \in \mathbb{C}^{n \times J}$, $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_J)$ und $\tilde{W} \in \mathbb{C}^{J \times J}$. Aufgrund der zusätzlichen Voraussetzung an die Matrizen V und W , welche maximalen Spaltenrang gewährleisten, sehen wir unmittelbar, dass für die berechneten Singulärwerte (wir sortieren dabei die Singulärwerte absteigend) gilt:

$$\sigma_{L+1} = \dots = \sigma_J = 0. \quad (5.6)$$

Somit können wir anstatt der vollen Singulärwertzerlegung ebenso die „reduzierte Zerlegung“ betrachten

$$A_0 = \tilde{V}_0 \Sigma_0 \tilde{W}_0^H, \quad (5.7)$$

wobei nun mehr $\tilde{V}_0 \in \mathbb{C}^{n \times L}$, $\tilde{W}_0 \in \mathbb{C}^{J \times L}$, sowie $\Sigma_0 = \text{diag}(\sigma_1, \dots, \sigma_L)$ gilt. Überdies gilt die folgende Identität:

$$\tilde{V}_0^H \tilde{V}_0 = \tilde{W}_0^H \tilde{W}_0 = \text{Id}_{L \times L}.$$

Dies bedeutet insbesondere, dass $S := \tilde{V}_0^H V$ regulär ist, womit wir ausgehend von der Zerlegung in (5.7) die folgende Identität erhalten:

$$\Sigma_0 \tilde{W}_0^H = \tilde{V}_0^H A_0 \stackrel{(5.4)}{=} \underbrace{\tilde{V}_0^H V}_{=S} W^H \hat{V} \Leftrightarrow S^{-1} \Sigma_0 \tilde{W}_0^H = W^H \hat{V}. \quad (5.8)$$

Diese brauchen wir nun nur mehr in die Darstellung von A_1 aus (5.4) einzusetzen und erhalten:

$$\begin{aligned} A_1 &= V \underbrace{\text{diag}(\lambda_1, \dots, \lambda_k)}_{=:D} W^H \hat{V} \Leftrightarrow A_1 = V D S^{-1} \Sigma_0 \tilde{W}_0^H \\ &\Leftrightarrow \tilde{V}_0^H A_1 = \underbrace{\tilde{V}_0^H V}_{=S} D S^{-1} \Sigma_0 \tilde{W}_0^H \\ &\Leftrightarrow \tilde{V}_0^H A_1 \tilde{W}_0 \Sigma_0^{-1} = S D S^{-1}. \end{aligned}$$

Somit reicht es also, um an die Eigenwerte λ_i zu kommen, die Eigenwerte der Matrix

$$B := \tilde{V}_0^H A_1 \tilde{W}_0 \Sigma_0^{-1} \in \mathbb{C}^{L \times L} \quad (5.9)$$

zu berechnen. Wir heben an dieser Stelle hervor, dass L genau die Summe der geometrischen Vielfachheiten aller Eigenwerte λ_i innerhalb von Γ bezeichnet, also $L \ll n$ gilt. Somit haben wir zumindest theoretisch das Eigenwertproblem einer $n \times n$ Matrix auf ein vergleichsweise einfach zu lösendes Problem zurückgeführt. Der Preis, den wir dafür zahlen müssen, liegt jedoch in der Berechnung der Matrizen A_0 sowie A_1 . Diese beinhalten, wie in der Definition (5.4) bereits ersichtlich ist, Integrale die ohne weiteres nicht explizit berechnet werden können. Einen Ausweg bietet die Berechnung dessen mittels Quadraturformeln, welche bei einer cleveren Wahl der Stützpunkte exponentiell konvergieren. Wir fassen dazu kurz die Resultate aus [Nan22, Lemma 8.12, Satz 8.13] zusammen:

Satz 5.3. Sei $R > r > 0$ und $D := \{z \in \mathbb{C} \mid r < \frac{|z|}{r_0} < R\}$ sowie $\lambda \mapsto f(\lambda) \in \mathbb{C}^n$ eine auf D holomorphe Funktion. Weiters bezeichne

$$\mathcal{Q}(f) := \frac{1}{2\pi i} \int_{|\lambda|=r_0} f(\lambda) d\lambda,$$

dann gilt für die folgende Quadraturformel

$$\mathcal{Q}_N(f) := \frac{R}{N} \sum_{n=0}^{N-1} \omega_N^n f(r_0 \omega_N^n), \quad \omega_N^n = \exp\left(\frac{2\pi i n}{N}\right), \quad (5.10)$$

wobei N die Anzahl der Quadraturpunkte bezeichnet, die folgende Fehlerabschätzung

$$|\mathcal{Q}(f) - \mathcal{Q}_N(f)| \leq \max_{|\lambda|=\sigma r_0} \|f(\lambda)\| \frac{\sigma^{-N}}{1 - \sigma^{-N}} + \max_{|\lambda|=\rho r_0} \|f(\lambda)\| \frac{\rho^{-N}}{1 - \rho^{-N}}$$

für alle $\rho \in (1, r), \sigma \in (1, R)$.

Für einen Beweis verweisen wir abermals auf [Nan22, Lemma 8.12]. Verwendet man nun als Integrationsweg einen Kreis $\Gamma(t) := z_0 + R \exp(2\pi i t)$ mit $t \in (0, 1)$, so gilt für die in (5.3) definierten Matrizen und den durch die oben definierten Quadraturformel approximierten Matrizen $A_{0,1}^{(N)}$ die folgende Fehlerabschätzung

$$\|A_{0,1}^{(N)} - A_{0,1}\| \leq C(d_-^{N-r+1} + d_+^{N-r+1}),$$

wobei r die maximale Ordnung der Pole von $A(\lambda)^{-1}$ bezeichnet und

$$d_- := \max_{\lambda \in \sigma(A), |\lambda - z_0| < r_0} \frac{|\lambda - z_0|}{R}, \quad d_+ := \max_{\lambda \in \sigma(A), |\lambda - z_0| > r_0} \frac{R}{|\lambda - z_0|}.$$

Somit haben wir alle wesentlichen Bestandteile diskutiert, die für die Integralmethode notwendig sind. Zusammenfassend können wir also den Algorithmus in drei Schritten zusammenfassen:

1. Berechne die Matrizen $A_0, A_1 \in \mathbb{C}^{n \times J}$ mittels Quadraturformel

2. Berechne die reduzierte Singulärwertzerlegung $A_0 = \tilde{V}_0 \Sigma_0 \tilde{W}_0^H$ wie in (5.7)
3. Berechne Eigenwerte der Matrix $B := \tilde{V}_0^H A_1 \tilde{W}_0 \Sigma_0^{-1} \in \mathbb{C}^{L \times L}$

Die beiden Letzteren sind vergleichsweise einfach durchzuführen. Den meisten Aufwand benötigt die Berechnung der Matrizen $A_0^{(N)}$ bzw. $A_1^{(N)}$, da für jeden Quadraturpunkt ein lineares Gleichungssystem gelöst werden muss.

5.2 Einführendes numerisches Beispiel

Um die bereits vorgestellte Theorie zu veranschaulichen, wollen wir in diesem Abschnitt mit einem einfachen Beispiel fortfahren. Für den Rest des Abschnitts wollen wir die Eigenwerte λ des negativen Laplace Operators $-\Delta$ auf dem Einheitsquadrat $\Omega := (0, 1) \times (0, 1)$, wobei wir zusätzlich Neumann Randbedingungen voraussetzen, numerisch berechnen. Dies entspricht der folgenden schwachen Formulierung:

Suche $(\lambda, u) \in \mathbb{C} \times H^1(\Omega)$, $u \neq 0$, sodass

$$\int_{\Omega} \nabla u \overline{\nabla v} \, dx = \lambda \int_{\Omega} u \overline{v} \, dx, \quad \forall v \in H^1(\Omega). \quad (5.11)$$

Wir bemerken an dieser Stelle, dass dies zwar einem linearen Eigenwertproblem entspricht, jedoch soll dieses Modellproblem ohnehin nur zur Illustration der recht technischen Schritte des vergangenen Abschnitts dienen. Zunächst bringen wir die Formulierung (5.11) in den Kontext von Abschnitt 5.1. Dabei gehen wir wie üblich in ein diskretes Setting über, indem wir die Lösung $u_h \in X_h$ mit $n := \dim X_h < \infty$ ansetzen. Weiters sei $\{\phi_1, \dots, \phi_n\}$ eine dazugehörige Basis, womit wir die folgenden Matrizen aufstellen können:

$$K \in \mathbb{C}^{n \times n}, \quad K_{ij} := \int_{\Omega} \nabla \phi_i \overline{\nabla \phi_j} \, dx$$

$$M \in \mathbb{C}^{n \times n}, \quad M_{ij} := \int_{\Omega} \phi_i \overline{\phi_j} \, dx.$$

Damit können wir nun unmittelbar den folgenden Operator für beliebiges $\Lambda \subset \mathbb{C}$ definieren:

$$A : \Lambda \rightarrow \mathbb{C}^{n \times n}, \lambda \mapsto A(\lambda) := K - \lambda M. \quad (5.12)$$

Dieser hängt offensichtlich holomorph von λ ab. Um die Eigenwerte dieses Operators zu berechnen, stellen wir nun eine Implementierung der Integralmethode in Python bzw. Netgen/NGSolve vor. Dabei geben wir auch Teile des Codes an, welche wir ebenfalls im nächsten Abschnitt in einer komplizierteren Problemstellung weiterverwenden werden. Ausgangspunkt ist dabei ein mesh bzw. Triangulierung des Einheitsquadrats Ω , auf dem das Modellproblem (5.11) mittels Finiten Elemente Diskretisierung aufgesetzt wird:

```
# FEM space
fes = H1(mesh, complex=True, order=p)
u, v = fes.TnT()
```

```

K = BilinearForm(fes)
K += grad(u)*grad(v)*dx
K.Assemble()
M = BilinearForm(fes)
M += u*v*dx
M.Assemble()

```

Sind die Matrizen K, M einmal initialisiert (dies geschieht über die `Assemble()` Funktion), können wir den Operator $\lambda \mapsto A(\lambda)$ aus (5.12) implementieren, dessen Eigenwerte wir mit der Integralmethode berechnen wollen.

```

def operatorA(lam,K,M):
    A = K.mat.CreateMatrix()
    A.AsVector().data = K.mat.AsVector() - lam*M.mat.AsVector()
    return A

```

Um die Matrizen A_0 bzw A_1 aus (5.3) aufstellen zu können, müssen wir zunächst eine „Solve Routine“ implementieren, welche für gegebenes $\lambda \in \Lambda \subset \mathbb{C}$ und „rechte Seite“ $\hat{V} \in \mathbb{C}^{n \times J}$ das folgende Quellproblem löst:

$$A(\lambda)X = \hat{V}, \quad \bar{Y} \in \mathbb{C}^{n \times J}.$$

Im nachfolgenden Code wird dabei die Inverse von $A(\lambda)$ direkt mittels einem in `NGSolve` vorimplementierten Solver berechnet. Dies ist jedoch nicht zwingend notwendig, da man alternativ auch eine *LU* Zerlegung der Matrix $A(\lambda)$ aufstellen und anschließend das auftretende Gleichungssystem mittels Vorwärts- bzw. Rückwärts-Substitution lösen könnte.

```

def Solve_Problem(lam,Vhat):
    #calculate A(lambda) and its inverse
    A = operatorA(lam,K,M)
    Ainv = A.Inverse(freedofs=fes.FreeDofs())
    #initialize result vector
    tmp = GridFunction(fes)
    #for each column of Vhat solve linear system
    for j in range(J):
        for i in range(X.vec.size):
            tmp.vec[i] = Vhat[i,j]
        tmp.vec.data = Ainv*tmp.vec
        #store result in matrix for output
        if j == 0:
            X = np.array(tmp.vec.FV().NumPy())
        else:
            X = np.c_[ X, tmp.vec.FV() ]
    return X

```

Wir sind nun in der Lage die Matrizen A_0, A_1 numerisch mittels der Quadraturformel aus (5.10) berechnen zu können. Wir beschränken uns dabei auf kreisförmige Kurven, nehmen

also an, dass Γ der folgenden Gestalt ist:

$$\Gamma : [0, 1] \rightarrow \mathbb{C}, t \mapsto z_0 + R \exp(2\pi ti).$$

Dabei ist weiters $z_0 \in \mathbb{C}$ und $R > 0$ vorausgesetzt. Verwenden wir eine Zerlegung des Intervalls $[0, 1]$ mit äquidistanten Stützstellen $t_j := j/N$ für $j = 0, \dots, N$, so erhalten wir mit $\lambda_j := \Gamma(t_j)$ die Quadraturformel für A_0

$$A_0^{(N)} := \frac{R}{N} \sum_{j=0}^{N-1} A(\lambda_j)^{-1} \hat{V} \exp\left(\frac{2\pi i j}{N}\right) \quad (5.13)$$

und nach geeignetem Zusammenfassen ebenfalls für A_1 :

$$A_1^{(N)} := z_0 A_0^{(N)} + \frac{R^2}{N} \sum_{j=0}^{N-1} A(\lambda_j)^{-1} \hat{V} \exp\left(\frac{4\pi i j}{N}\right). \quad (5.14)$$

Wir bezeichnen ab sofort mit $N \in \mathbb{N}$ die Anzahl der Quadraturpunkte. Eine Implementierung der Quadraturformeln in Python sieht somit wie folgt aus:

```
def calculate_A0A1(N):
    for n in range(N):
        lam = gamma(z0,R,n,N)
        W = Solve_Problem(lam,What)
        if n == 0:
            A0 = W*np.exp(1j*2*math.pi*n/N)
            A1 = W*np.exp(1j*4*math.pi*n/N)
        else:
            A0 = A0 + W*np.exp(1j*2*math.pi*n/N)
            A1 = A1 + W*np.exp(1j*4*math.pi*n/N)
    A0 = A0*R/N
    A1 = z0*A0 + R**2/N*A1
    return A0, A1
```

Dabei wurden die einzelnen Quadraturpunkte λ_j über die nachfolgende Funktion berechnet:

```
def gamma(z0,R,j,N):
    return z0+R*np.exp(1j*2*math.pi*j/N)
```

Als nächstes muss die nachfolgende Singulärwertzerlegung wie in (5.5) beschrieben für $A_0^{(N)}$ durchgeführt werden:

$$A_0^{(N)} = \tilde{V} \Sigma \tilde{W}^H.$$

Dies kann in Python durch die folgenden Zeilen

```
V, sigma, WH = np.linalg.svd(A0, full_matrices=False)
```

mittels dem `numpy.linalg.svd` Paket von [Har+20] durchgeführt werden. Um aus dieser Darstellung dann schlussendlich die reduzierte Singulärwertzerlegung (5.7) zu erhalten, geben wir uns eine Toleranz τ vor und suchen $k \in \mathbb{N}$, sodass

$$\sigma_1 \geq \dots \geq \sigma_k \geq \tau > \sigma_{k+1} \approx \dots \approx \sigma_J \approx 0 \quad (5.15)$$

gilt. Kann diese Wahl erfolgreich durchgeführt werden, müssen lediglich die ersten k Spalten bzw. Reihen von V_0 bzw. W_0 behalten werden. Wir bemerken an dieser Stelle, dass die Wahl der Toleranz τ ein kritischer Punkt sein kann, da sich die „künstlichen“ Singulärwerte in der Praxis nicht unbedingt bei 0 anhäufen müssen. In der nachfolgenden Implementierung fordern wir deshalb, dass der Algorithmus abbrechen soll, sofern der Rang Test (5.15) nicht erfolgreich durchgeführt werden kann.

```
def check_sv(V,sigma,WH,tol):
    k = 0
    for l in range(len(sigma)):
        if sigma[l] > tol:
            k += 1
    if k == 0:
        raise ValueError("Rank Test not successful - no SVs found")
    elif k == J:
        raise ValueError("Rank Test not successful - increase J")
    else:
        sigmainv = inv(np.diag(sigma[:k]))
        VOH = V[:, :k].conjugate().transpose()
        W0 = WH[:k, :].conjugate().transpose()
        return VOH, sigmainv, W0
```

Konnten alle bisherigen Schritte erfolgreich durchgeführt werden, wird nun abschließend die Matrix $B \in \mathbb{C}^{k \times k}$ wie in (5.9) aufgestellt und mittels `np.linalg.eig` Paket die Eigenwerte und dazugehörigen Eigenvektoren berechnet. Die Matrix B selbst ist typischerweise von einer kleinen Größe, weshalb die Eigenwerte vergleichsweise einfach zu ermitteln sind.

```
def calc_ev(VOH, sigmainv, W0, A1):
    mat = VOH @ A1 @ W0 @ sigmainv
    eigval, eigvec = np.linalg.eig(mat)
    V = VOH.transpose().conjugate() @ eigvec
    return eigval, V
```

Die oben vorgestellten Code-Teile können dann nach Belieben in einer einzigen Funktion bzw. Routine zusammengefasst werden. Kommen wir nun zurück zu der konkreten Problemstellung aus (5.11), so befinden wir uns in der komfortablen Lage, die exakten Eigenwerte analytisch darzustellen bzw. berechnen zu können. Diese lassen sich mittels eines Separationsansatzes berechnen. Dazu setzen wir Eigenfunktionen u der folgenden Bauart an:

$$u(x, y) = \cos(m\pi x) \cos(n\pi y), \quad m, n \in \mathbb{R}.$$

So erhalten wir mit Hilfe der Neumann Randbedingungen

$$\begin{aligned} \frac{\partial}{\partial x} u(1, y) &= -m\pi \sin(m\pi) \cos(n\pi y) \stackrel{!}{=} 0 \Rightarrow m \in \mathbb{N}_0 \\ \frac{\partial}{\partial y} u(x, 1) &= -n\pi \cos(m\pi x) \sin(n\pi) \stackrel{!}{=} 0 \Rightarrow n \in \mathbb{N}_0. \end{aligned}$$

Daraus können wir die Eigenwerte

$$\lambda_{nm} = (n^2 + m^2)\pi^2, \quad n, m \in \mathbb{N}_0 \quad (5.16)$$

mit dazugehörigen Eigenfunktionen

$$u_{nm} = \cos(m\pi x) \cos(n\pi y), \quad n, m \in \mathbb{N}_0 \quad (5.17)$$

ablesen, womit wir auch über die geometrischen Vielfachheiten der einzelnen Eigenwerte Bescheid wissen. Wir wollen nun die ersten von 0 verschiedenen Eigenwerten mit der

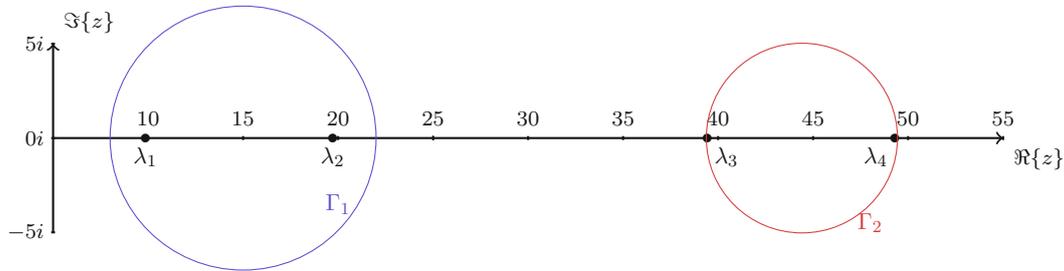


Abbildung 5.1: Integrationswege Γ_i zur Berechnung der Eigenwerte

Integralmethode berechnen. Mit der Notation von (5.16) entspricht dies gerade den Werten $\lambda_1 := \lambda_{01} = \pi^2$, $\lambda_2 := \lambda_{11} = 2\pi^2$, $\lambda_3 := \lambda_{02} = 4\pi^2$ und $\lambda_4 := \lambda_{12} = 5\pi^2$. In den nachfolgenden numerischen Ergebnissen haben wir dazu die Kurven

$$\begin{aligned} \Gamma_1 &: t \mapsto 15 + 7 \exp(2\pi it) \\ \Gamma_2 &: t \mapsto 4.5\pi^2 + \left(\frac{\pi^2}{2} + 0.1\right) \exp(2\pi it) \end{aligned}$$

gewählt, siehe auch Abbildung 5.1. Bevor wir die numerischen Ergebnisse präsentieren, wollen wir noch kurz diskutieren, welche (maximalen) Konvergenzordnungen wir überhaupt für Problem (5.12) erwarten können. Dabei führen wir im Wesentlichen die Resultate aus [NFP21] und [Nan22, Satz 7.6, Bemerkung 7.7] an.

Setzt man allgemein eine Lösung $u_h \in X_h \subset H^1(\Omega)$ in (5.11) mit $\dim X_h < \infty$ an, so ist von entscheidender Bedeutung in der Konvergenztheorie von Eigenwerten, wie gut die kontinuierlichen Eigenfunktionen $u \in E(\lambda) := \ker A(\lambda) \setminus \{0\}$ durch X_h approximiert werden können. Um etwaige Unklarheiten zu beseitigen, sagen wir, dass λ_h ein diskreter Eigenwert mit Eigenfunktion $u_h \neq 0$ ist, wenn

$$\int_{\Omega} \nabla u_h \overline{\nabla v_h} dx = \lambda_h \int_{\Omega} u_h \overline{v_h} dx, \quad \forall v_h \in X_h \quad (5.18)$$

gilt. Können wir überdies voraussetzen, dass

$$\sup_{\substack{u \in E(\lambda), \\ \|u\|=1}} \inf_{v \in X_h} \|u - v\|_{H^1(\Omega)} = \mathcal{O}(h^p)$$

für $p \in \mathbb{N}$ gilt, dann kann man wie in [Nan22, Abschnitt 7, Bemerkung 7.7] nachweisen, dass für das Modellproblem (5.12) die diskreten Eigenwerte λ_h mit Ordnung $2p$ gegen die kontinuierlichen Eigenwerte konvergieren, also die folgende Abschätzung gilt:

$$|\lambda - \lambda_h| = \mathcal{O}(h^{2p}). \quad (5.19)$$

In NGSolve verwenden wir wie üblich Spline Räume der Ordnung p , setzen also $X_h := S^p(\mathcal{T})$ für eine reguläre Triangulierung \mathcal{T} von Ω , wodurch diese Voraussetzung erfüllt ist. Man kann nämlich ähnlich wie in [NFP21, Theorem 3.5 bzw. Korollar 3.6] nachweisen, dass für $u \in H^{p+1}(\Omega)$ die folgende Abschätzung gilt:

$$\inf_{v_h \in S^p(\mathcal{T})} \|u - v_h\|_{H^1(\Omega)} \leq Ch^p.$$

Mit zusätzlichen Techniken kann man überdies zeigen, dass die Forderung $u \in H^{p+1}(\Omega)$ für Eigenfunktionen von (5.11) keine wesentliche Einschränkung ist, da dies für konvexe Ω ohnehin gilt. In den numerischen Experimenten konnten wir nun auch jene theoretischen Konvergenzraten beobachten. Während wir bei den ersten beiden Versuchen die Kurve Γ_1 so wählten, dass sich die Eigenwerte λ_1 und λ_2 „mittig“ innerhalb der Kurve befanden, wählten wir im zweiten Experiment die Kurve so, dass die exakten Eigenwerte sehr nahe an deren Rand liegen. Auf die Konvergenzgeschwindigkeit scheint dies jedoch keine Auswirkungen zu haben, wie in Abbildung 5.2 ersichtlich ist.

Zusammenfassend können wir in allen vier Fällen sehen, dass die erhaltenen numerischen Konvergenzraten auch theoretisch begründbar sind bzw. ohnehin aus (5.19) zu erwarten waren. Für die Approximationen der Matrizen A_0 bzw. A_1 haben wir überdies die Quadraturformel aus (5.13) bzw. (5.14) mit $N = 24$ für die Berechnung von λ_1 und λ_2 bzw. $N = 32$ Stützpunkten für die von λ_3 und λ_4 gewählt. Da die zu erwartende Summe der geometrischen Vielfachheit der innerhalb der Kurven liegenden Eigenwerte verglichen mit (5.17) maximal 4 war, haben wir überdies $J = 6$ für die Anzahl der Spalten der Zufallsmatrix \hat{V} gewählt.

Zuletzt wollen wir noch herausheben, dass bei diesem Modellproblem der Rang Test (5.15) ohne Probleme durchführbar war, da die künstlichen Singulärwerte sich sehr stark von den „echten“ Singulärwerten unterscheiden haben. Tatsächlich lagen die gesuchten Singulärwerte in der Größenordnung 10^3 bis 10^5 , während die künstlichen einer Größenordnung von 10^{-8} entsprachen und somit problemlos aussortiert werden konnten.

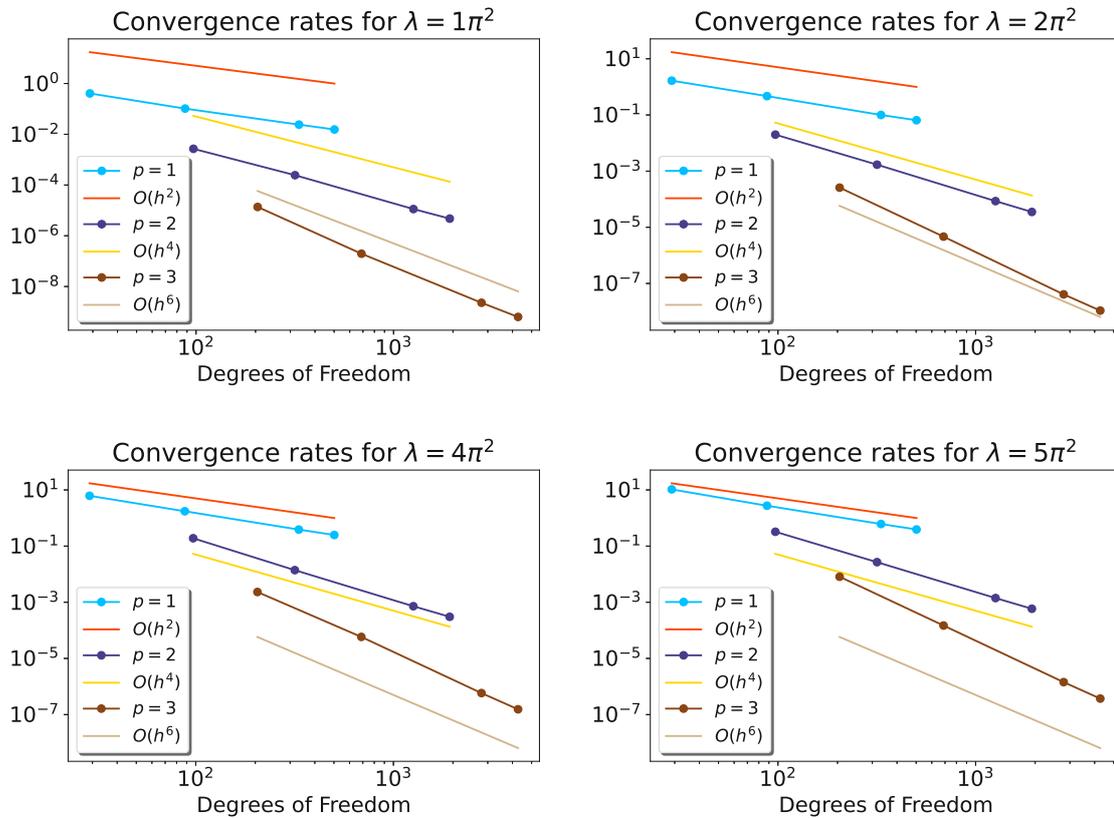


Abbildung 5.2: Darstellung der Fehler $|\lambda - \lambda_h|$ für verschiedene Eigenwerte

5.3 Kombination mit Learned Infinite Elemente Verfahren

Dieser Abschnitt stellt nun eine Kombination der bisher in dieser Arbeit präsentierten Methoden dar. Dies bedeutet, dass wir die Integralmethode gemeinsam mit dem Learned Infinite Elemente Verfahren zur Berechnung von Eigenwerten auf unbeschränkten Gebieten verwenden werden. Dabei suchen wir wie schon zuvor die Eigenwerte des negativen Laplace Operators $-\Delta$ mit Neumann Randbedingungen auf einem diesmal unbeschränkten Gebiet Ω . Als Geometrie verwenden wir in den nachfolgenden Experimenten dazu einen einseitig offenen und unbeschränkten Wellenleiter, der zusätzlich einen Streukörper K in deren Inneren beinhaltet. Dieser Sachverhalt wird in der nachfolgenden Abbildung illustriert. Nach



Abbildung 5.3: Unbeschränkte Geometrie zur Berechnung von Eigenwerten

Einführung eines transparenten Randes Γ_a führt dies auf das folgende Eigenwertproblem:

Suche $(\lambda, u) \in \mathbb{C} \times H^1(\Omega_{\text{int}}) \setminus \{0\}$, sodass

$$\int_{\Omega_{\text{int}}} \nabla u \overline{\nabla v} dx - \lambda \int_{\Omega_{\text{int}}} u \overline{v} dx - \int_{\Gamma_a} \mathcal{DtN}(u|_{\Gamma_a}) \overline{v|_{\Gamma_a}} ds = 0 \quad (5.20)$$

für alle $v \in H^1(\Omega_{\text{int}})$ gilt. Um das Problem (5.20) diskretisieren zu können, ersetzen wir abermals u, v durch $u_h, v_h \in X_h \subset H^1(\Omega_{\text{int}})$ mit $\dim X_h =: n < \infty$. Weiters bezeichnen wir mit $n|_{\Gamma_a}$ die Anzahl der Freiheitsgrade des FEM Raumes, welche zu Γ_a gehören. Ersetzen wir nun ebenfalls wie in Kapitel 4.1 beschrieben, den exakten Dirichlet-to-Neumann Operator mit Hilfe des Learned Infinite Elemente Ansatzes, erhält man ein endlichdimensionales Eigenwertproblem der folgenden Form:

Suche $(\lambda, u) \in \mathbb{C} \times \mathbb{C}^m \setminus \{0\}$, sodass

$$A(\lambda)u = 0, \quad A(\lambda) \in \mathbb{C}^{m \times m}, \quad m := n + (N_{\text{ext}} - 1) \cdot n|_{\Gamma_a}. \quad (5.21)$$

Dabei bezeichnet zur Erinnerung N_{ext} die Anzahl der externen Freiheitsgrade im Sinne von Kapitel 4.1. Wir halten an der Stelle fest, dass die Matrix A aus zwei verschiedenen Blöcken besteht: einerseits aus der FEM-Matrix

$$S(\lambda) := (S_{ij})_{i,j=1}^n = \int_{\Omega_{\text{int}}} \nabla \varphi_i \overline{\nabla \varphi_j} dx - \lambda \int_{\Omega_{\text{int}}} \varphi_i \overline{\varphi_j} dx$$

und andererseits aus dem approximierten DtN Operator

$$L(\lambda) := \begin{pmatrix} L_{\Gamma\Gamma}(\lambda) & L_{\Gamma E}(\lambda) \\ L_{E\Gamma}(\lambda) & L_{EE}(\lambda) \end{pmatrix}, \quad L_{\Gamma\Gamma}(\lambda) \in \mathbb{C}^{n|_{\Gamma_a} \times n|_{\Gamma_a}}, \quad L_{EE}(\lambda) \in \mathbb{C}^{(N_{\text{ext}}-1) \cdot n|_{\Gamma_a} \times (N_{\text{ext}}-1) \cdot n|_{\Gamma_a}}.$$

Um die weitere Präsentation zu vereinfachen, sei nun Γ die Menge der Indexe der Freiheitsgraden von X_h zugehörig zu Γ_a und I die Menge alle anderen (also „inneren“) Freiheitsgrade von X_h . Mit jener Notation lässt sich die Matrix A wie folgt umordnen:

$$A(\lambda) = \begin{pmatrix} S_{II}(\lambda) & S_{I\Gamma}(\lambda) & 0 \\ S_{\Gamma I}(\lambda) & S_{\Gamma\Gamma}(\lambda) + L_{\Gamma\Gamma}(\lambda) & L_{\Gamma E}(\lambda) \\ 0 & L_{E\Gamma}(\lambda) & L_{EE}(\lambda) \end{pmatrix}. \quad (5.22)$$

Daraus wird nun ebenfalls klar, dass die Matrizen $S(\lambda)$ und $L(\lambda)$ über die Freiheitsgrade auf Γ_a miteinander gekoppelt werden. Somit sind wir wieder in der gleichen Ausgangslage wie in Kapitel 5.1, sofern wir nachweisen können, dass $\lambda \mapsto A(\lambda)$ holomorph auf einer Teilmenge $\Lambda \subset \mathbb{C}$ ist. Als problematisch für die Definition einer solchen Teilmenge Λ erweist sich tatsächlich die Matrix $L(\lambda)$ in der Formulierung (5.22). Dies wird insofern ersichtlich, indem man sich die Definition des exakten DtN Operators genauer ansieht. Bezeichnen wir wie üblich mit (λ_n, v_n) die Eigenpaare des negativen Laplace Beltrami Operators $-\Delta_{\Gamma_a}$, wobei die Orthonormalbasis bezüglich des L^2 Skalarproduktes gewählt wird, dann erhalten wir nach (3.43) die folgende Darstellung für $u_0 \in L^2(\Gamma_a)$:

$$\mathcal{DtN}u_0 = \sum_{n=0}^{\infty} \underbrace{i\sqrt{\lambda - \lambda_n}}_{=: dtn(\lambda_n)} (u_0, v_n)_{L^2(\Gamma_a)} v_n. \quad (5.23)$$

Erst aus jener Darstellung ist zu erkennen, dass dann ebenfalls die Matrix A aus (5.22) nicht linear von λ abhängt, womit wir im Vergleich zu dem vorherigen Abschnitt tatsächlich ein nichtlineares Eigenwertproblem vorliegen haben. Des Weiteren können wir $\Lambda \subset \mathbb{C}$ nicht mehr beliebig wählen, da die komplexe Wurzel, wie aus der Komplexen Analysis bekannt, nicht auf ganz \mathbb{C} holomorph ist. Somit müssen wir, wie bereits in einem früheren Kapitel erwähnt, die sogenannten „Branch-Cuts“, welche bei der Definition einer jeden komplexen Wurzel entstehen, aus dem Definitionsbereich von A ausnehmen. Ferner, müssen wir bei einer Implementierung der Integralmethode sicherstellen, dass der Integrationsweg nicht über einen solchen Branch-Cut verläuft. Um etwaige Ungereimtheiten zu vermeiden, geben wir nun die Definition der komplexen Wurzel an, die wir in den nachfolgenden Experimenten verwenden:

$$\sqrt{\cdot} : \mathbb{C} \setminus i\mathbb{R}_{<0} \rightarrow \mathbb{C}, z \mapsto \sqrt{z} := \sqrt{|z|} \exp\left(\frac{i \arg z}{2}\right). \quad (5.24)$$

Dabei bezeichnet $\arg z$ den Winkel (Polarkoordinaten!), für den wir den folgenden Definitionsbereich wählen:

$$-\frac{\pi}{2} < \arg z \leq \frac{3\pi}{2}.$$

Die nachfolgende Abbildung stellt eine Visualisierung der oben definierten komplexen Wurzel dar. Auf der linken Seite befindet sich eine Darstellung des Betrags der Wurzel in der komplexen Ebene sowie ebenfalls ein Weg, der über eine Unstetigkeitsstelle verläuft. Die Auswirkung dessen wird auch rechts in Abbildung 5.4 durch den Sprung an $\frac{3\pi}{2}$ ersichtlich. Mit jenen Vorbereitungen sind wir nun in der Lage, die in Kapitel 5.1 vorgestellte Integral-

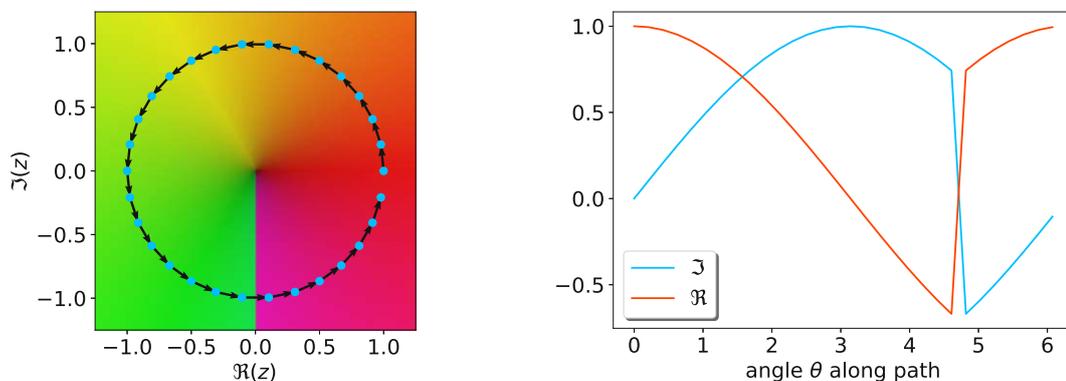


Abbildung 5.4: Links: Betrag der komplexen Wurzel mit Integrationsweg über eine Unstetigkeitsstelle, Rechts: Real- und Imaginärteil der komplexen Wurzel entlang des Weges

methode auf Problem (5.20) bzw. (5.22) anzuwenden. Um sicher zu stellen, dass $\lambda \mapsto A(\lambda)$ holomorph ist, setzen wir den Definitionsbereich wie folgt voraus:

$$\Lambda \subset \mathbb{C} \setminus \bigcup_{n \in \mathbb{N}_0} \{\lambda_n + i\mathbb{R}_{<0}\}. \quad (5.25)$$

Weiters sei Γ eine geschlossene Kurve, die ganz in Λ verläuft. Bis auf wenige Ausnahmen können wir nun die in Kapitel 5.2 definierten Funktionen wiederverwenden. Im Kern des Algorithmus steckt, wie bereits diskutiert, die Lösung des folgenden (Streu-) Problems

$$A(\lambda)\hat{U} = \hat{V}, \quad \hat{V} \in \mathbb{C}^{m \times J}, \quad (5.26)$$

wobei $A(\lambda) \in \mathbb{C}^{m \times m}$ wie in (5.22) definiert, \hat{V} eine Zufallsmatrix und $J \in \mathbb{N}$ größer als die Summe aller geometrischen Vielfachheiten der innerhalb von Γ liegenden Eigenwerte λ_i ist. Wir wollen nun noch genauer auf die spezielle Struktur der Matrix $A(\lambda)$ eingehen. Wie in Kapitel 4.1 bereits geschildert, wird der exakte DtN Operator über eine Erweiterung der „FEM“-Matrix $S(\lambda)$, bzw. in weiterer Form durch das Schur Komplement der Matrix $L(\lambda)$ approximiert. Dies bedeutet, dass wir die künstlichen, externen Freiheitsgrade in der Lösung des Streuproblems (5.26) und analog in der Zufallsmatrix \hat{V} nicht berücksichtigen müssen. Analog zu den in Kapitel 4.2 präsentierten Dirichlet-Streuproblemen ist also zusammengefasst das folgende Problem zu lösen:

$$\underbrace{\begin{pmatrix} A_{II}(\lambda) & A_{I\Gamma}(\lambda) & 0 \\ A_{\Gamma I}(\lambda) & A_{\Gamma\Gamma}(\lambda) & A_{\Gamma E}(\lambda) \\ 0 & A_{E\Gamma}(\lambda) & A_{EE}(\lambda) \end{pmatrix}}_{=A(\lambda)} \begin{pmatrix} U_I \\ U_\Gamma \\ U_E \end{pmatrix} = \begin{pmatrix} \hat{V}_I \\ \hat{V}_\Gamma \\ 0 \end{pmatrix}.$$

Somit macht es nun ebenfalls Sinn, anstatt der „vollen Matrix“ $A(\lambda)^{-1}\hat{V} \in \mathbb{C}^{m \times J}$ aus Definition (5.4) die Matrix

$$\hat{U}(\lambda) := \begin{pmatrix} U_I \\ U_\Gamma \end{pmatrix} \in \mathbb{C}^{n \times J}$$

zu wählen, um A_0 und A_1 aufzustellen, da ja die externen Freiheitsgrade nur in Form des Schur Komplements zur Lösung beitragen und deshalb U_E selbst nicht relevant ist. Die einzige Änderung im Vergleich zu der im vorherigen Kapitel präsentierten Vorgehensweise stellt also die Berechnung der Matrizen A_0 und A_1 dar, die nun wie folgt zu Berechnen sind:

$$A_0 := \frac{1}{2\pi i} \int_{\Gamma} \hat{U}(\lambda) d\lambda, \quad A_1 := \frac{1}{2\pi i} \int_{\Gamma} \lambda \hat{U}(\lambda) d\lambda. \quad (5.27)$$

In Hinblick auf eine numerische Implementierung können diese wieder mit der Quadraturformel aus (5.13) bzw. (5.14) approximiert werden, womit der Rest des Algorithmus analog wie im vorherigen Kapitel, anhand eines einfachen Modellproblems erklärt, durchgeführt werden kann.

Bemerkung 5.4. An dieser Stelle wollen wir noch hervorheben, dass bei der Kombination der Integralmethode mit dem Learned Infinite Elemente Verfahren nicht nur für jeden Quadraturpunkt bzw. jede Frequenz λ ein lineares Gleichungssystem gelöst werden muss, sondern ebenfalls die gelernten Matrizen $L(\lambda)$ neu aufgestellt werden müssen. Dies bedeutet, dass zusätzlich für jeden Schritt ein Optimierungsproblem gelöst werden muss, da ja $L(\lambda)$ aus jenen Teilmatrizen besteht, deren Schur Komplement den exakten Dirichlet-to-Neumann Operator am „besten“ approximiert. Dies macht die ohnehin schon aufwendige Integralmethode in gewisser Weise noch „teurer“.

5.3.1 Aufsetzen der Problemstellung

In diesem Abschnitt werden wir nun die bisherigen Überlegungen auf eine konkrete Problemstellung anwenden. Um die Präsentation möglichst einfach zu gestalten, verwenden wir als Geometrie einen Wellenleiter der Breite π mit einem Quadrat der Seitenlänge $\pi/4$ innerhalb des Wellenleiters als Streukörper. Mathematisch fassen wir jenes Setting durch Setzen von $K := (\frac{3}{8}\pi, \frac{5}{8}\pi) \times (\frac{3}{8}\pi, \frac{5}{8}\pi)$ und dem daraus resultierenden (unbeschränkten) Gebiet

$$\Omega := ((0, \infty) \times (0, \pi)) \setminus K.$$

Nach Wählen des transparenten Randes $\Gamma_a := \{\pi\} \times (0, \pi)$ erhalten wir als Innenraumgeometrie $\Omega_{\text{int}} := (0, \pi)^2 \setminus K$, siehe auch Abbildung 5.5. Um einen geeigneten Definitionsbereich

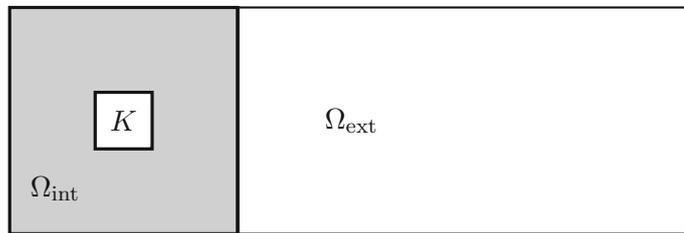


Abbildung 5.5: Geometrie zur Berechnung von Eigenwerten mittels Integralmethode

für den in (5.22) definierten Operator $\lambda \mapsto A(\lambda)$ zu finden, benötigen wir zuerst die Eigenwerte λ_n des negativen Laplace-Beltrami Operators mit Neumann Randbedingungen auf Γ_a bzw. auf dem Intervall $(0, \pi)$. Wir haben bereits gezeigt, dass diese durch $\lambda_n = n^2$ für $n \in \mathbb{N}_0$ bestimmt sind. Somit können wir den Definitionsbereich mittels der Formel (5.25) wie folgt bestimmen

$$\Lambda \subset \mathbb{C} \setminus \bigcup_{n \in \mathbb{N}_0} \{n^2 \times i\mathbb{R}_{<0}\}$$

und schlussfolgern, dass $\lambda \mapsto A(\lambda)$ mit $\lambda \in \Lambda$ eine holomorphe Funktion darstellt. Während wir im letzten Kapitel Referenzlösungen als Ausgangspunkt zur Wahl einer geeigneten Kurve Γ wählen konnten, ist dies mit der diesmal zu Grunde liegenden Geometrie klarerweise nicht möglich. Im Allgemeinen ist es jedoch möglich nachzuweisen, dass Eigenwerte λ von A einen negativen Imaginärteil haben, also die folgende Implikation gilt:

$$\forall (\lambda, u) \in \mathbb{C} \times \mathbb{C}^m, u \neq 0 : A(\lambda)u = 0 \Rightarrow \Im(\lambda) < 0.$$

Dies schränkt zumindest unsere Suche auf die untere Halbebene in der komplexen Ebene ein. Das bedeutet jedoch auch, dass wir dort auf Branch-Cuts der komplexen Wurzel treffen, über die nicht integriert werden darf, siehe auch Abbildung 5.6. Dies müssen wir bei der Wahl einer geeigneten Kurve natürlich zusätzlich berücksichtigen. Man könnte nun versuchen bzw. hoffen, durch simples Ausprobieren mehrerer Kurven auf Eigenwerte zu treffen, wir bedienen uns in dieser Arbeit jedoch einer anderen Vorgehensweise. Im Detail verwenden wir einen zusätzlichen Eigenwertrechner, basierend auf der PmL (*Perfectly*

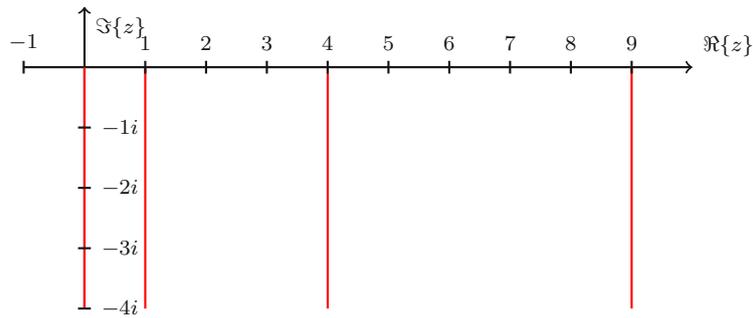


Abbildung 5.6: Unstetigkeitsstellen im Definitionsbereich $\Lambda \subset \mathbb{C}$ von A

Matched Layer) Methode, welcher uns eine gewisse Vorahnung gibt, wo sich Eigenwerte befinden werden. Da die PmL Methode selbst nicht Gegenstand dieser Arbeit ist, werden wir nicht näher auf den Algorithmus und dessen Funktionsweise eingehen, lediglich die Ergebnisse interpretieren, die in Abbildung 5.7 zusammengefasst sind. Grundsätzlich kann man zunächst die Unstetigkeitsstellen entlang der Branch-Cuts sehen (vergleiche auch mit Abbildung 5.6). Diese werden über die PmL Methode als verzerrte, punktierte Geraden sichtbar gemacht. Nimmt man jene sogenannten „künstlichen Eigenwerte“ aus, so bleiben die farbigen

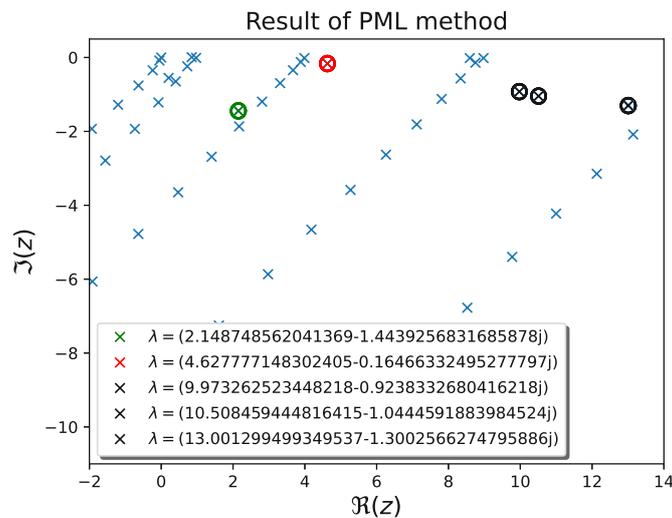


Abbildung 5.7: Ergebnisse der Eigenwertberechnung mittels PmL Methode

gekennzeichneten Punkte als Kandidaten für Eigenwerte von A übrig. Diese bilden den Ausgangspunkt der weiteren Untersuchungen mit Hilfe der Integralmethode. Dabei wollen wir insbesondere die Wechselwirkung zwischen dem approximierten Dirichlet-to-Neumann Operator und der Integralmethode selbst untersuchen. Ein kritischer Punkt der Integralmethode ist, wie bereits diskutiert, das Aufstellen einer reduzierten Singulärwertzerlegung wie in (5.7) beschrieben, bei der die künstlichen Singulärwerte aussortiert werden müssen. Interessant hinsichtlich der Qualität des Verfahrens ist nun sicherlich, ob sich dies immer

bewerkstelligen lässt. Als Motivation der nachfolgenden Experimente formulieren wir die folgenden drei Fragen:

1. Welche Integrationswege Γ sind sinnvoll und was ändert sich an der Konvergenz der Eigenwerte, wenn sich die Kurve einer Unstetigkeitsstelle nähert?
2. Können die künstlichen Singulärwerte immer eindeutig aussortiert werden und wie hängt der entstehende „Gap“ von den Modell-Parametern ab?
3. Wie verhält sich die Konvergenzgeschwindigkeit des Fehlers in Abhängigkeit der Qualität der Approximation des DtN Operators, insbesondere der Anzahl der externen Freiheitsgrade N_{ext} ?

Anhand von ausgewählten Beispielen werden wir nun versuchen diese Fragen zu beantworten. Dabei versuchen wir zunächst mit Hilfe der Integralmethode die in Abbildung 5.7 ermittelten Kandidaten zu bestätigen.

5.3.2 Einzelne Eigenwerte innerhalb einer Kontur

In einem ersten Experiment beschränken wir uns auf Integrationswege Γ , in deren Inneren lediglich ein einziger Eigenwert liegt. Dabei nehmen wir zunächst den rot markierten Referenzwert aus Abbildung 5.7 als Vorbild, nämlich

$$\lambda_{\text{ref}} = 4.627777148302405 - 0.16466332495277797i. \quad (5.28)$$

Diesen (potenziellen) Eigenwert versuchen wir nun mit der Integralmethode zu bestätigen. Bevor wir mit der Interpretation der Ergebnisse beginnen, sei noch bemerkt, dass λ_{ref} durch die PmL Methode mit einer Polynomordnung $p = 8$ und mesh size $h = 0.05$ berechnet wurde.

Um die Güte der Konvergenz festzustellen, verwenden wir die folgende Konfiguration der Modell-Parameter:

- Quadraturpunkte $N = 16$, $t_j := j/N$ für $j = 0, \dots, N - 1$
- Anzahl äußerer Freiheitsgrade $N_{\text{ext}} = 5$ für DtN Operator
- Anzahl zu berücksichtigten DtN Nummern $L_{\text{max}} = 10$ für das Optimierungsproblem
- Anzahl der Spalten $J = 3$ von \hat{V}
- Integrationsweg $\Gamma(t) := (4.6 - 0.15i) + 0.5 \exp(2\pi it)$
- Gewichtsfunktion für $\kappa_j := \Gamma(t_j)$:

$$\omega_{jn} = \begin{cases} 1, & n^2 \leq |\Re(\kappa_j)| \\ |\exp(i\sqrt{\kappa_j - n^2} \cdot (\pi - 5/8\pi))|, & n^2 > |\Re(\kappa_j)| \end{cases}$$

Untersucht wird zunächst die Konvergenzgeschwindigkeit in Bezug auf die mesh size h für unterschiedliche Polynomordnungen p . Neben dem Integrationsweg Γ sind die Ergebnisse

der Simulationen in Abbildung 5.8 zusammengefasst. Erstaunlicherweise sind keine signifikanten Unterschiede der Geschwindigkeit in Abhängigkeit der Ordnung p des Spline Raumes $S^p(\mathcal{T})$ zu erkennen. Dies kann verschiedene Gründe haben, die wir zu einem späteren Zeitpunkt noch genauer durchleuchten wollen. Zunächst sei aber gesagt, dass anders als im Modellproblem des letzten Kapitels, die zu Grunde liegende Geometrie Ω_{int} aufgrund des Streukörpers K nicht konvex ist. Dies hat zur Folge, dass ohne weiteres keine Aussagen über die Regularität der Eigenfunktionen $u \in \ker A(\lambda) \setminus \{0\}$ gemacht werden können, welche in weiterer Konsequenz maßgeblich für die Konvergenzgeschwindigkeit der Eigenwerte verantwortlich ist. Werfen wir nun einen Blick auf das Resultat der Singulärwertzerlegungen

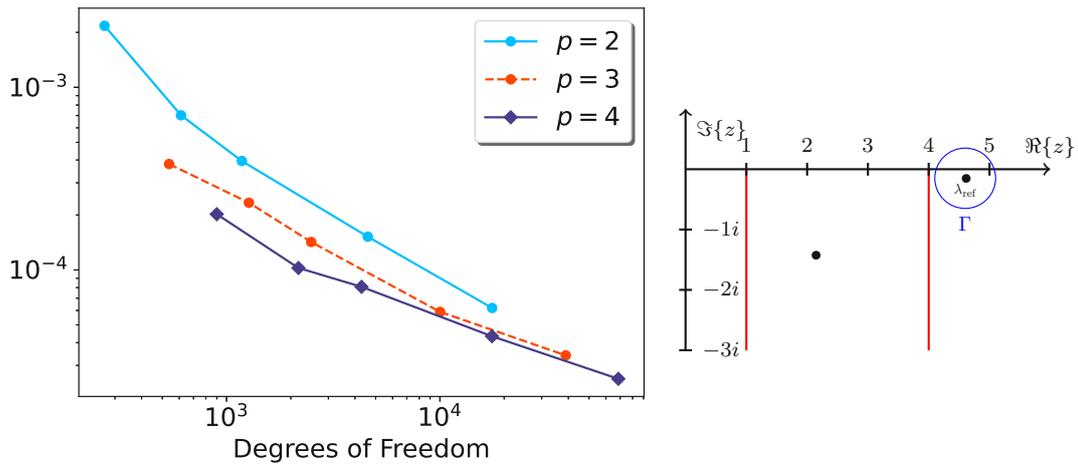


Abbildung 5.8: Links: Fehler $|\lambda_{\text{ref}} - \lambda^{(h)}|$ gemessen an dem Referenzwert aus (5.28), Rechts: Integrationsweg Γ für die Integralmethode

der approximierten Matrix $A_0^{(N)}$ wie in (5.7). Dabei bezeichnen wir mit dem sogenannten „Gap“ die Differenz zwischen σ_L und σ_{L+1} im Sinne von Definition (5.6). Die nachfolgende Tabelle fasst sowohl den Gap als auch die Größenordnungen der ersten beiden Singulärwerte zusammen:

mesh	$p = 2$		$p = 3$		$p = 4$	
	Gap	σ_1/σ_2	Gap	σ_1/σ_2	Gap	σ_1/σ_2
0.5	30.85	$10^1/10^{-3}$	26.66	$10^1/10^{-3}$	31.46	$10^1/10^{-2}$
0.3	22.67	$10^1/10^{-2}$	38.13	$10^1/10^{-2}$	33.28	$10^1/10^{-2}$
0.2	35.55	$10^1/10^{-2}$	42.32	$10^1/10^{-2}$	51.58	$10^1/10^{-2}$
0.1	446.80	$10^2/10^{-2}$	486.09	$10^2/10^{-2}$	410.77	$10^2/10^{-2}$
0.05	2996.64	$10^3/10^{-1}$	3167.88	$10^3/10^{-1}$	3262.73	$10^3/10^{-1}$

Abbildung 5.9: Lage der Singulärwerte für Simulationen aus Abbildung 5.8

Zunächst ist in Tabelle 5.9 erkennbar, dass sich der Gap zwischen σ_1 und σ_2 (bei absteigender Sortierung der Singulärwerten) vergrößert, je feiner das mesh gewählt wird. Tatsächlich

war mit der oben gewählten Konfiguration der Rang-Test durch Wählen von $\tau := 10^{-1}$ problemlos durchführbar. Um ausschließen zu können, dass die Approximation des DtN Operators den dominanten Teil des Fehlers darstellt, werden wir den Fehler in Abhängigkeit der Anzahl an äußeren Freiheitsgrade N_{ext} als nächstes analysieren. Gleichzeitig gehen wir ebenfalls der Frage nach, wie der Algorithmus auf Kurven Γ in der Nähe von Unstetigkeitsstellen reagiert. Dazu passen wir die zuvor gesetzte Konfiguration insofern an, dass wir eine fixe mesh-size von $h = 0.1$ und Polynomordnung $p = 3$ verwenden. Die Ergebnisse

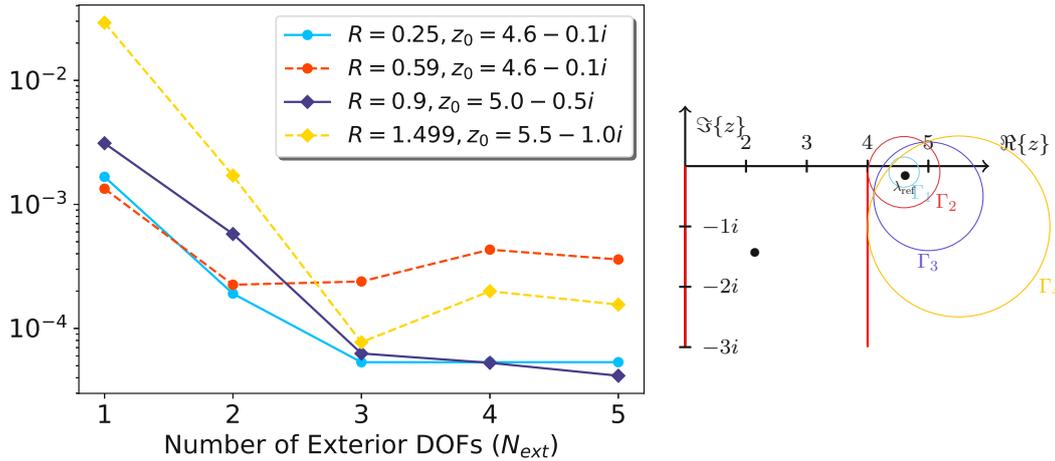


Abbildung 5.10: Links: Fehler $|\lambda_{\text{ref}} - \lambda^{(h)}|$ gemessen an dem Referenzwert aus (5.28), Rechts: Integrationswege Γ_i für die Integralmethode

befinden sich in Abbildung 5.10 und lassen erahnen, dass die Wahl der Kurve Γ bzw. deren Lage in Bezug auf Unstetigkeitsstellen im Definitionsbereich von A nur bedingt Auswirkung auf die Konvergenzgeschwindigkeit hat. Zwar liefern die Kurven Γ_1 und Γ_3 bessere Approximationen, jedoch sind die Unterschiede nicht signifikant genug, um an diesem Beispiel als möglichen Grund dafür die Nähe des Weges zur Unstetigkeitsstelle ausmachen zu können. Vielmehr muss man ebenfalls für Kurve Γ_4 den größeren Radius berücksichtigen, der wohlmöglich zusätzliche Quadraturpunkte für eine bessere Approximation der Integrale verlangt. Interessant ist überdies auch, dass bei allen Kurven die „maximale Konvergenz“ bereits bei einer sehr geringen Anzahl an externen Freiheitsgraden (in diesem Fall $N_{\text{ext}} = 3$) erreicht wurde. Dies ist ebenfalls ein Anzeichen dafür, dass die Approximation des DtN Operators nicht den dominanten Fehler darstellt. Überprüfen wir nun noch die Lage der Singulärwerte σ_i im Hinblick auf die Größe des Gaps. Diese sind in der nachfolgenden Tabelle (Abbildung 5.11) zusammengefasst. In einem ersten Experiment konnten wir somit keine signifikanten Unterschiede der Lage der Singulärwerte in Abhängigkeit der Nähe von Kurven Γ zu Unstetigkeitsstellen nachweisen. In der Tat unterscheiden sich die künstlichen Singulärwerte in sämtlichen Beispielen um eine Größenordnung von mindestens 10^2 von den gesuchten Werten.

Wie bereits erwähnt, wollen wir nun noch die Konvergenzgeschwindigkeiten aus Abbildung

N_{ext}	Γ_1		Γ_2		Γ_3		Γ_4	
	Gap	σ_1/σ_2	Gap	σ_1/σ_2	Gap	σ_1/σ_2	Gap	σ_1/σ_2
1	455.53	$10^2/10^{-2}$	429.23	$10^2/10^0$	420.09	$10^2/10^0$	447.57	$10^2/10^1$
2	375.48	$10^2/10^{-1}$	369.21	$10^2/10^0$	437.54	$10^2/10^0$	479.87	$10^2/10^0$
3	463.89	$10^2/10^{-2}$	461.85	$10^2/10^0$	316.67	$10^2/10^{-1}$	463.61	$10^2/10^0$
4	334.26	$10^2/10^{-2}$	327.59	$10^2/10^0$	356.78	$10^2/10^{-1}$	350.80	$10^2/10^0$
5	404.80	$10^2/10^{-6}$	350.30	$10^2/10^{-1}$	426.23	$10^2/10^{-1}$	450.89	$10^2/10^{-1}$

Abbildung 5.11: Lage der Singulärwerte für Simulationen aus Abbildung 5.10

5.8 besser verstehen. Dazu erinnern wir abermals daran, dass die Matrizen A_0 und A_1 der Integralmethode, wie in (5.27) beschrieben, im Wesentlichen aus den Lösungen verschiedener Streuproblemen bestehen. Wir versuchen im Folgenden die Konvergenzgeschwindigkeiten aus Abbildung 5.8 zu plausibilisieren, indem wir die resultierenden Streuprobleme hinsichtlich der Konvergenz analysieren. Dazu wollen wir die in Abschnitt 4.2.3 bereits präsentierten Funktionen und Resultate (wieder-)verwenden und skalieren die Geometrie Ω_{int} , bzw. Ω auf eine Breite von 1. Dies hat, wie wir bereits gesehen haben, zur Folge, dass sich die Eigenwerte von $-\Delta_\Gamma$ sowie in weiterer Konsequenz ebenfalls die gesuchten Eigenwerte um den Faktor π^2 ändern. Zusammenfassend resultiert dies in der Analyse des folgenden Problems:

$$\begin{aligned}
 -\Delta u - \omega u &= 0, & \text{in } \Omega_{\text{int}} \\
 u &= u_0, & \text{auf } \Gamma_D := \{0\} \times (0, 1) \\
 \frac{\partial}{\partial \eta} u &= 0, & \text{auf } \Gamma_N := \partial\Omega_{\text{int}} \setminus \underbrace{(\Gamma_D \cup \{1\} \times (0, 1))}_{=: \Gamma_a}.
 \end{aligned} \tag{5.29}$$

Auf Grund der skalierten Geometrie werden wir Frequenzen ω in der Nähe von $\lambda_{\text{ref}} \cdot \pi^2 \approx 46 - 2i$ untersuchen. Dies entspricht nämlich in etwa den Frequenzen, die man bei Anwendung der Integralmethode zur Berechnung des durch π^2 skalierten Eigenwertes (5.28) auf dem entsprechenden Gebiet Ω_{int} verwenden müsste. Überdies setzen wir für die Analyse als Dirichlet Daten $u_0 = \sum_{l=1}^3 \cos(l\pi \cdot)$ voraus. Für die Approximation des DtN Operators verwenden wir die gleichen Parameter N_{ext} und L_{max} mit dem einzigen Unterschied, dass wir die Gewichtsfunktion ω_n an die Geometrie wie folgt anpassen müssen:

$$\omega_n = \begin{cases} 1, & n^2 \pi^2 \leq |\Re(\omega)| \\ |\exp(i\sqrt{\omega - \pi^2 n^2} \cdot (3/8))|, & n^2 \pi^2 > |\Re(\omega)|. \end{cases} \tag{5.30}$$

Um die Konvergenzgeschwindigkeit schließlich zu analysieren, verwenden wir als Referenzlösung, wie in Abschnitt 4.2.3 erläutert, die Lösung auf einem Gebiet $\tilde{\Omega} \supset \Omega_{\text{int}}$ mit Polynomordnung $p = 6$ und einer feinen mesh size ($h = 0.01$). In Abbildung 5.12 fassen wir die numerischen Ergebnisse für $\omega := 46 - 2i$ zusammen. Tatsächlich konnten wir analog zu Abbildung 5.8 keine Unterschiede der Konvergenzgeschwindigkeit des L^2 Fehlers der Lösung des Streuproblems in Bezug auf die Polynomordnung des Spline Raumes erkennen. Ausgehend von diesen Ergebnissen wirken die Geschwindigkeiten der approximierten Eigenwerten in Abbildung 5.8 ebenfalls wenig überraschend, da ja die Matrizen

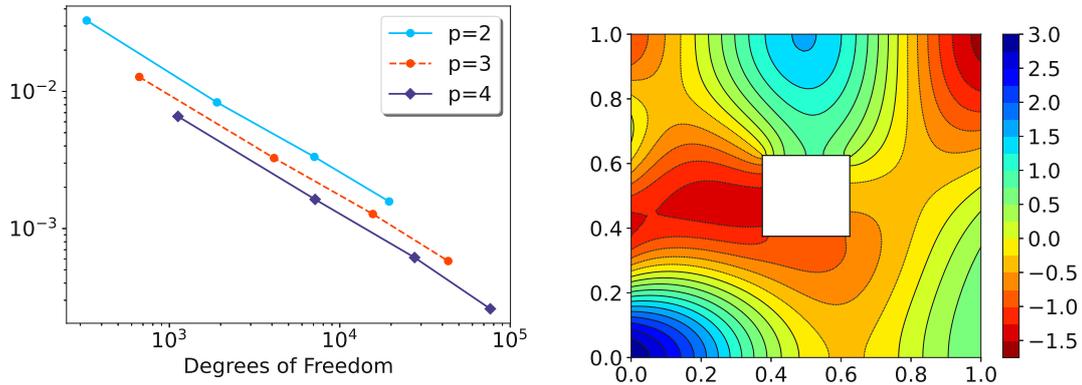


Abbildung 5.12: Links: $L^2(\Omega_{int})$ Fehler des Streuproblems (5.29) mit $\omega = 46 - 2i$ in Abhängigkeit der mesh size h , Rechts: Visualisierung der Referenzlösung $\Re(u_h)$

der Integralmethode gerade aus jenen Lösungen bestehen (die Dirichlet Daten u_0 werden durch zufällige Inhomogenitäten, also den Spalten von \hat{V} ersetzt). Dies ist zwar in keinem Fall der Beweis, dass durch Setzen von anderen Modellparametern keine besseren Konvergenzgeschwindigkeiten erreicht werden können, jedoch hilft das vorausgehende Experiment, um die zumindest auf den ersten Blick überraschenden Konvergenzgeschwindigkeiten aus Abbildung 5.8 besser zu verstehen.

Ausgehend von den bisherigen Experimenten können wir bereits erste Schlüsse der Integralmethode in Kombination mit der Learned Infinite Elemente Methode ziehen. Zunächst konnten wir feststellen, dass die Approximationen gegen den zuvor mit der PmL Methode ermittelten Eigenwert konvergieren. Die Geschwindigkeit der Konvergenz hängt dabei mit der Konvergenz der resultierenden Streuproblemen zusammen, wie das vorausgehende Experiment ebenfalls bestätigt hat. Der entscheidende Vorteil der Integralmethode ist natürlich, dass die Lage eines möglichen Kandidaten zuverlässig überprüft bzw. bestätigt werden kann. Wir möchten nun die bisherigen Experimente anhand eines weiteren Beispiels vertiefen, genauer versuchen wir einen weiteren Eigenwert aus Abbildung 5.7 mittels Integralmethode zu „verifizieren“. Wir wählen dafür als Referenzlösung den in Abbildung 5.7 grün markierten Eigenwert, nämlich

$$\lambda_{\text{ref}} = 2.148748562041369 - 1.4439256831685878i. \quad (5.31)$$

Abermals wurde dieser Referenzwert durch das PmL Verfahren auf einem „feinen mesh“ mit mesh size $h = 0.05$ und hoher Polynomordnung ($p = 8$) berechnet. Das Ziel wird zunächst wieder sein, die Lage des Kandidatens mit der Integralmethode zu bestätigen. Dabei verwenden wir ähnlich wie zuvor die folgende Konfiguration der Input Parameter:

- Quadraturpunkte $N = 16$, $t_j := j/N$ für $j = 0, \dots, N - 1$
- Anzahl äußerer Freiheitsgrade $N_{\text{ext}} = 5$ für DtN Operator
- Anzahl zu berücksichtigter DtN Nummern $L_{\text{max}} = 8$ für das Optimierungsproblem

- Anzahl der Spalten $J = 3$ von $\hat{V} \in \mathbb{C}^{m \times J}$
- Integrationsweg $\Gamma(t) := (2.5 - 1.5i) + 0.5 \exp(2\pi it)$
- Gewichtsfunktion für $\kappa_j := \Gamma(t_j)$:

$$\omega_{jn} = \begin{cases} 1, & n^2 \leq |\Re(\kappa_j)| \\ |\exp(i\sqrt{\kappa_j - n^2} \cdot (\pi - 5/8\pi))|, & n^2 > |\Re(\kappa_j)|. \end{cases}$$

Wir wiederholen nun die zuvor präsentierten Experimente am Beispiel des Referenzwertes aus (5.31) und beginnen mit der Analyse des Fehlers in Abhängigkeit der Freiheitsgraden bzw. mesh size h für verschiedene Polynomordnungen p . Die in Abbildung 5.13 zusammen-

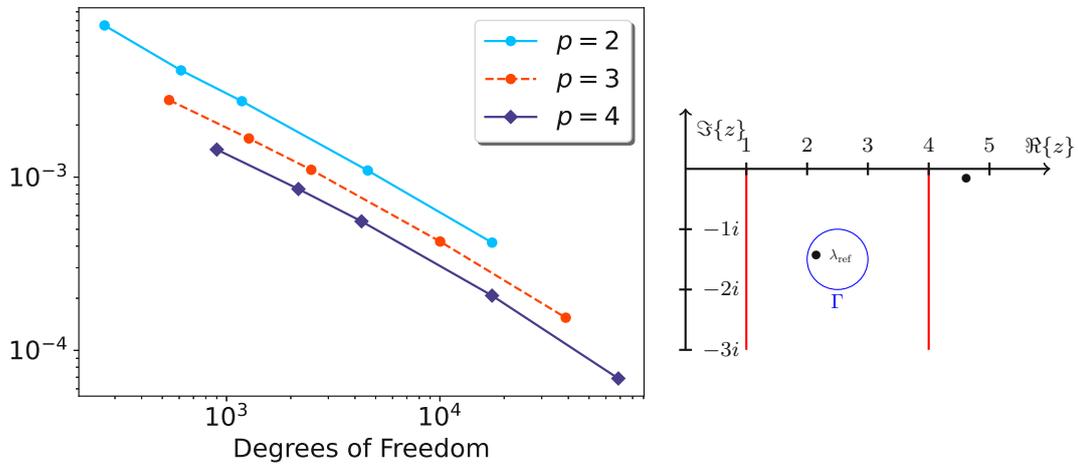


Abbildung 5.13: Links: Fehler $|\lambda_{ref} - \lambda^{(h)}|$ gemessen an dem Referenzwert aus (5.31), Rechts: Integrationsweg Γ für die Kontur-Integralmethode

mengefassten Ergebnisse zeigen ein sehr ähnliches Verhalten des Fehlers wie zuvor. Insbesondere konnten wir erneut nachweisen, dass die Geschwindigkeit der Konvergenz des Fehlers nicht von der Ordnung des Spline Raumes abzuhängen scheint. Um den Grund für die Geschwindigkeiten zu erläutern, wollen wir auch diesmal die Lösungen der resultierenden Streuprobleme (5.29) untersuchen. Erneut gehen wir zu einer auf Breite 1 skalierten Geometrie über, womit wir Frequenzen ω in der Nähe von $\lambda_{ref} \cdot \pi^2$ untersuchen werden. Wir vervollständigen die Konfiguration mit derselben Gewichtsfunktion wie in (5.30), sowie Dirichlet Daten $u_0 := \sum_{l=1}^2 \cos(l\pi \cdot)$ und einer zu untersuchenden Frequenz $\omega := 21 - 15j$. Analog zu den bisher vorgestellten Experimenten erhalten wir eine Referenzlösung durch Lösen des Streuproblems auf einem größeren Gebiet $\tilde{\Omega} \supset \Omega_{int}$ mit kleiner mesh-size $h = 0.01$ und hoher Polynomordnung ($p = 6$). Die Resultate der Simulationen befinden sich in Abbildung 5.14 und zeigen wie schon in der Berechnung der Eigenwerte keine signifikanten Unterschiede der Konvergenz des L^2 Fehlers der Lösung in Bezug auf die Polynomordnung p . Mit dieser Kenntnis wirken nun ebenfalls die Resultate aus Abbildung 5.13 wenig erstaunlich.

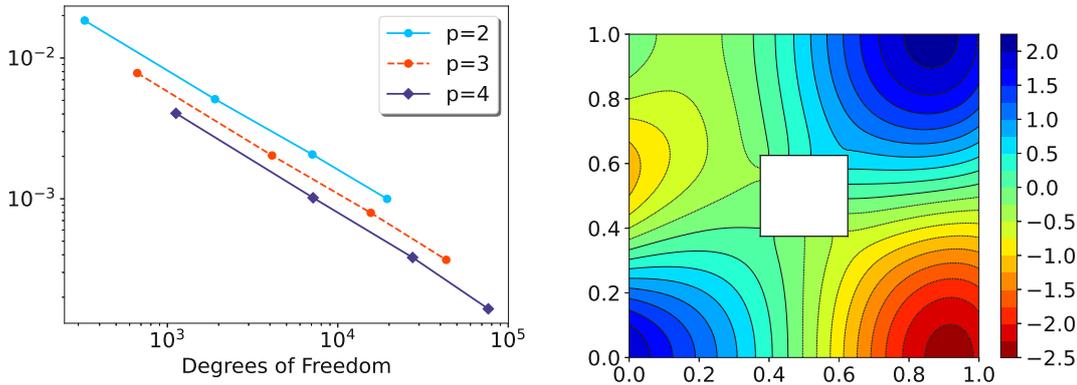


Abbildung 5.14: Links: $L^2(\Omega_{int})$ Fehler des Streuproblems (5.29) mit $\omega = 21 - 15i$ in Abhängigkeit der mesh size h , Rechts: Visualisierung der Referenzlösung $\Re(u_h)$

Wir kehren nun zurück zur Analyse der Integralmethode und betrachten den Gap zwischen den Singulärwerten σ_1 und σ_2 . Wir konnten erneut beobachten, dass sich auch bei diesem Beispiel die künstlichen Singulärwerte leicht aussortieren lassen, wie in der nachfolgenden Tabelle ersichtlich ist:

mesh	$p = 2$		$p = 3$		$p = 4$	
h	Gap	σ_1/σ_2	Gap	σ_1/σ_2	Gap	σ_1/σ_2
0.5	43.13	$10^1/10^{-4}$	82.72	$10^1/10^{-4}$	83.58	$10^1/10^{-3}$
0.3	109.70	$10^2/10^{-4}$	117.40	$10^2/10^{-3}$	219.50	$10^2/10^{-4}$
0.2	183.766	$10^2/10^{-4}$	137.58	$10^2/10^{-3}$	253.63	$10^2/10^{-3}$
0.1	908.66	$10^2/10^{-4}$	1193.58	$10^3/10^{-3}$	751.45	$10^2/10^{-3}$
0.05	3208.01	$10^3/10^{-4}$	2958.17	$10^3/10^{-4}$	2046.01	$10^3/10^{-3}$

Abbildung 5.15: Lage der Singulärwerte für Simulationen aus Abbildung 5.13

Auffallend in Tabelle 5.15 ist überdies, dass sich die künstlichen Singulärwerte sogar noch stärker unterscheiden bzw. bei Null häufen als es im vorherigen Beispiel in Tabelle 5.9 ersichtlich war. Wir versuchen als nächstes wieder die Parameter bzw. Kurven so zu verändern, dass Probleme auftreten könnten. Dazu verwenden wir unterschiedliche Konturen Γ_i , die unterschiedlich nahe eines Branch-Cuts liegen. Des Weiteren testen wir den Algorithmus mit einem Integrationsweg Γ_2 , der über eine Unstetigkeitsstelle im Definitionsbereich Λ verläuft, was laut der in Kapitel 5.1 vorgestellten Theorie ausgeschlossen war. In den nachfolgenden Experimenten verwenden wir überdies die gleiche Konfiguration wie zuvor, ergänzen diese lediglich mit einer fixen mesh-size $h = 0.1$ und Polynomordnung $p = 3$. Die Resultate sind in Abbildung 5.16 zusammengefasst.

Wie aus der Theorie zu erwarten war, liefert die in Abbildung 5.16 rot markierte Kurve Γ_2 keine (zufriedenstellende) Konvergenz des Fehlers. Wir erkennen also, dass die Holomorphie (insbesondere Stetigkeit) des Operators $\lambda \mapsto A(\lambda)$ unabdingbar ist. Was die Konvergenz bezüglich der anderen Szenarien anbelangt, so erkennen wir wie schon zuvor, dass allein die

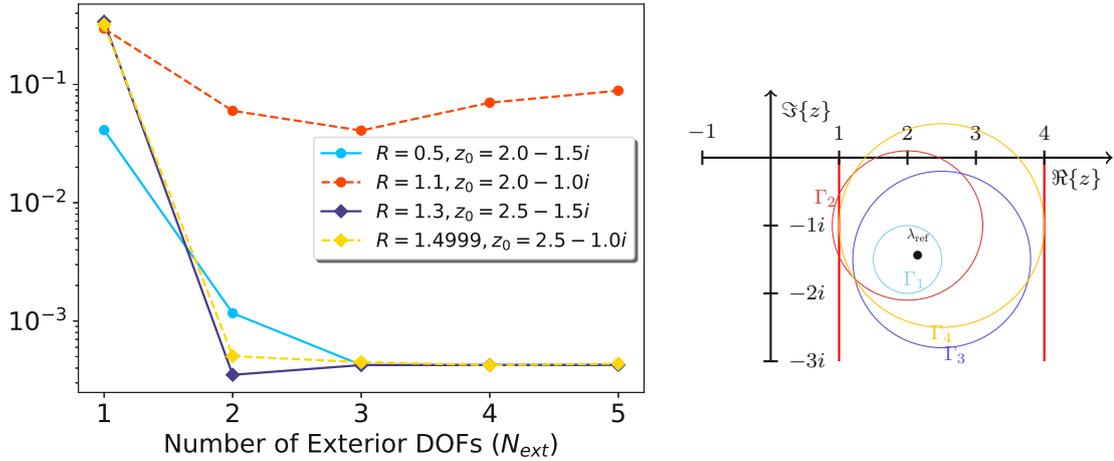


Abbildung 5.16: Links: Fehler $|\lambda_{ref} - \lambda^{(h)}|$ gemessen an dem Referenzwert aus (5.31), Rechts: Integrationswege Γ_i für Integralmethode

Nähe einer Kurve Γ_i zu Unstetigkeitsstellen im Definitionsbereich des Operators A keine unmittelbaren Auswirkungen auf die Konvergenzgeschwindigkeit der Eigenwerte zu haben scheint. Überdies wird die „maximale Konvergenz“ erneut bei einer geringer Anzahl an externen Freiheitsgraden N_{ext} erreicht. Wir führen zuletzt noch die Lage der Singulärwerte σ_i in der nachfolgenden Tabelle an: Bei Betrachten von Tabelle 5.17 wird nun ersichtlich,

N_{ext}	Γ_1		Γ_2		Γ_3		Γ_4	
	Gap	σ_1/σ_2	Gap	σ_1/σ_2	Gap	σ_1/σ_2	Gap	σ_1/σ_2
1	121.94	$10^2/10^1$	796.99	$10^2/10^1$	7864.27	$10^3/10^2$	3502.54	$10^3/10^2$
2	899.21	$10^2/10^{-1}$	579.42	$10^2/10^1$	605.84	$10^2/10^0$	852.30	$10^2/10^1$
3	652.09	$10^2/10^{-1}$	229.56	$10^2/10^1$	656.96	$10^2/10^0$	679.34	$10^2/10^1$
4	852.72	$10^2/10^{-6}$	690.09	$10^2/10^1$	739.24	$10^2/10^{-1}$	786.94	$10^2/10^1$
5	609.76	$10^2/10^{-5}$	541.01	$10^2/10^1$	685.52	$10^2/10^{-1}$	919.23	$10^2/10^1$

Abbildung 5.17: Lage der Singulärwerte für Simulationen aus Abbildung 5.16

dass der Gap für Γ_1 (hinsichtlich Größenordnung!) und $N_{ext} > 3$ am deutlichsten war. Erstaunlicherweise konnten auch die zu Γ_2 gehörigen Singulärwertzerlegungen durchgeführt werden. Zwar näherten sich die künstlichen Singulärwerten nicht so stark bei 0 wie bei den anderen Kurven, jedoch konnte trotzdem eine Aussortierung nach Anpassen der Toleranz τ durchgeführt werden.

5.3.3 Mehrere Eigenwerte innerhalb einer Kontur

Bislang wurden die Kurven Γ immer so gewählt, dass nur ein einzelner Eigenwert λ in deren Inneren lag. Der Vorteil der Integralmethode besteht jedoch unter anderem darin,

alle Eigenwerte innerhalb einer gegebenen Kontur berechnen zu können. Sei L die Summe aller geometrischen Vielfachheiten solcher Eigenwerte, dann resultiert die Integralmethode in einem (klassischen) Eigenwertproblem der Matrix $B \in \mathbb{C}^{L \times L}$, wie in (5.9) beschrieben wurde. Dies bedeutete in den bisherigen Experimenten, dass B als Konstante bereits den gesuchten Eigenwert beinhaltet. Somit war in jenem Fall keine klassische Eigenwertberechnung mehr nötig. Dies ändert sich, sobald mehrere Eigenwerte innerhalb einer Kurve liegen.

In diesem Abschnitt wollen wir verschiedene Szenarien testen, bei denen mehrere Eigenwerte innerhalb einer Kurve liegen. Als Referenzwerte eignen sich dafür die in schwarz gekennzeichneten Kandidaten aus Abbildung 5.7. Die nachfolgenden Experimente verwenden die folgende Konfiguration:

- Quadraturpunkte $N = 32$, $t_j := j/N$ für $j = 0, \dots, N - 1$
- Äußere Freiheitsgrade $N_{\text{ext}} = 4$ für DtN Operator
- Anzahl zu berücksichtigter DtN Nummern $L_{\text{max}} = 10$ für das Optimierungsproblem
- Anzahl der Spalten $J = 4$ von $\hat{V} \in \mathbb{C}^{m \times J}$
- Integrationsweg $\Gamma(t) := (11.5 - i) + 2 \exp(2\pi it)$
- Gewichtsfunktion für $\kappa_j := \Gamma(t_j)$:

$$\omega_{jn} = \begin{cases} 1, & n^2 \leq |\Re(\kappa_j)| \\ |\exp(i\sqrt{\kappa_j - n^2} \cdot (\pi - 5/8\pi))|, & n^2 > |\Re(\kappa_j)| \end{cases}$$

Die Kurve Γ ist diesmal so gewählt, dass drei (mögliche) Eigenwerte in deren Inneren liegen, siehe auch Abbildung 5.18. Dies hat nun ebenfalls zur Folge, dass $J \geq 3$ gelten muss.

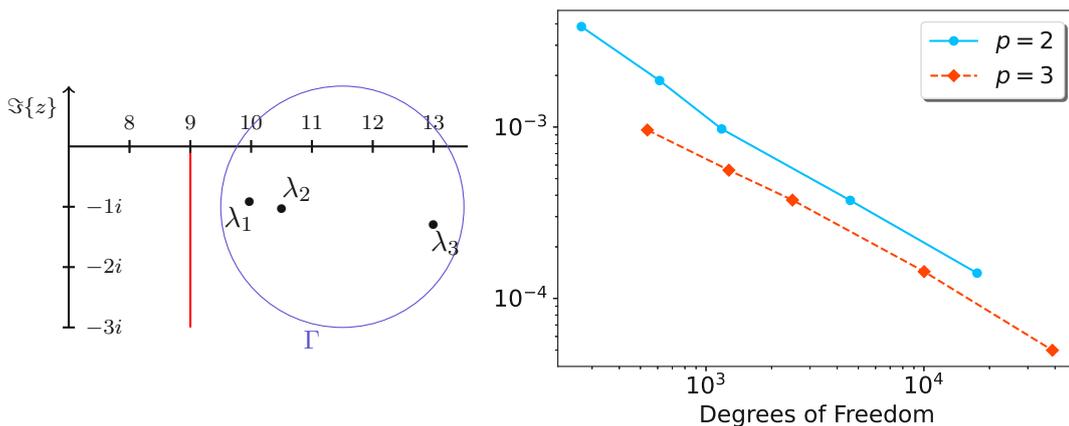


Abbildung 5.18: Links: Integrationsweg Γ für Integralmethode, Rechts: Fehler $|\lambda_1 - \lambda_1^{(h)}|$

Ebenfalls ist eine Erhöhung der Anzahl der Quadraturpunkte N nötig, um das Integral

über die Kontur ausreichend gut approximieren zu können. Um weitere Missverständnisse zu vermeiden, bemerken wir an dieser Stelle, dass die Referenzwerte λ_1 , λ_2 und λ_3 abermals mittels PmL Methode berechnet wurden. Die genauen Werte befinden sich in Abbildung 5.7, weshalb wir diese nicht erneut anführen.

Zu erwarten ist nach den bisherigen Erkenntnissen, dass alle drei Eigenwerte gegen die dazugehörigen Referenzwerte konvergieren. Allerdings sind geringe Unterschiede in der Konvergenz zwischen den einzelnen Eigenwerten möglich, da bei der Berechnung der Eigenwerte der Matrix B ebenfalls nicht alle Eigenwerte gleich gut approximiert werden können. In den Abbildungen 5.18 und 5.19 befinden sich die numerischen Ergebnisse, welchen den Fehler der Approximationen der Eigenwerte in Bezug auf die Anzahl der Freiheitsgraden veranschaulichen. Die Resultate können analog zu denen aus dem vorangehenden Kapitel interpretiert werden, zeigen demnach keine signifikanten Unterschiede in Bezug auf die Polynomordnung. Die Konvergenzraten zwischen den einzelnen Eigenwerten unterscheiden sich

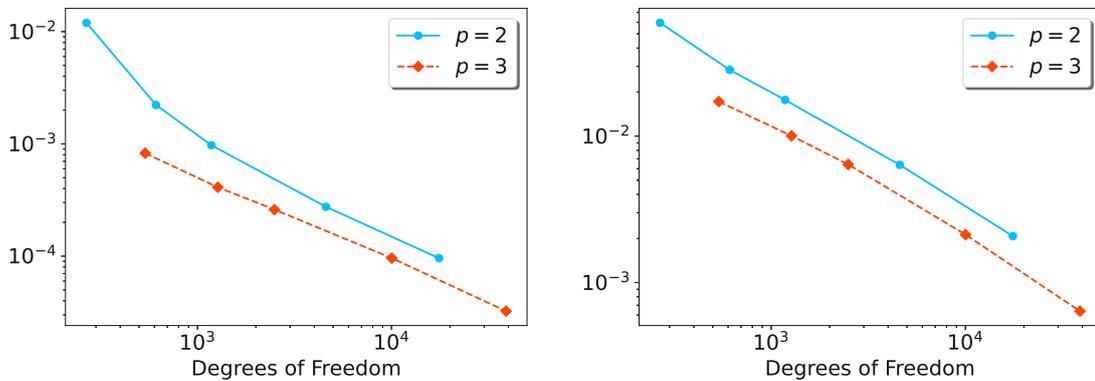


Abbildung 5.19: Links: Fehler $|\lambda_2 - \lambda_2^{(h)}|$, Rechts: Fehler $|\lambda_3 - \lambda_3^{(h)}|$

ebenfalls nur minimal. Im Detail wird der betragsmäßig kleinste Eigenwert λ_1 am besten approximiert (Größenordnung 10^{-4}), wohingegen die Approximation an den betragsmäßig größten Eigenwert λ_3 einen Fehler von $\mathcal{O}(10^{-3})$ aufweist. Diese Unterschiede können unter anderem, wie schon erklärt, auf die eigentliche Eigenwertberechnung der Matrix B aus (5.9) mittels des in Python zur Verfügung stehenden Pakets `np.linalg.eig` zurückgeführt werden. Es bleibt noch die Frage offen, wie sich die Lage der Singulärwerte σ_i bei mehrfachen Eigenwerten innerhalb einer Kurve verhält. Dies wird in Abbildung 5.20 sichtbar gemacht.

Wie man unmittelbar sieht, war auch bei diesem Experiment mit mehreren Eigenwerten innerhalb einer Kontur der Rang-Test problemlos durchführbar bzw. die künstlichen Singulärwerte eindeutig aussortierbar. Zur Erinnerung, ein erfolgreicher Rang-Test bestand darin, für eine gegebene Toleranz τ ein $k \in \mathbb{N}$ zu finden, sodass

$$\sigma_1 \geq \dots \geq \sigma_k > \tau > \sigma_{k+1} \approx \dots \approx \sigma_J \approx 0 \quad (5.32)$$

gilt. Bei genauerem Betrachten ist hier ebenfalls eine Korrelation zwischen der mesh-size h und der Größe des Gaps zu erkennen: Je feiner das mesh, desto deutlicher ist die Trennung

mesh	$p = 2$				$p = 3$			
h	σ_1	σ_2	σ_3	σ_4	σ_1	σ_2	σ_3	σ_4
0.5	$\mathcal{O}(10^2)$	$\mathcal{O}(10^2)$	$\mathcal{O}(10^1)$	$\mathcal{O}(10^{-5})$	$\mathcal{O}(10^2)$	$\mathcal{O}(10^2)$	$\mathcal{O}(10^1)$	$\mathcal{O}(10^{-6})$
0.3	$\mathcal{O}(10^2)$	$\mathcal{O}(10^1)$	$\mathcal{O}(10^1)$	$\mathcal{O}(10^{-6})$	$\mathcal{O}(10^2)$	$\mathcal{O}(10^2)$	$\mathcal{O}(10^2)$	$\mathcal{O}(10^{-5})$
0.2	$\mathcal{O}(10^2)$	$\mathcal{O}(10^1)$	$\mathcal{O}(10^1)$	$\mathcal{O}(10^{-5})$	$\mathcal{O}(10^2)$	$\mathcal{O}(10^2)$	$\mathcal{O}(10^1)$	$\mathcal{O}(10^{-6})$
0.1	$\mathcal{O}(10^3)$	$\mathcal{O}(10^2)$	$\mathcal{O}(10^1)$	$\mathcal{O}(10^{-5})$	$\mathcal{O}(10^3)$	$\mathcal{O}(10^2)$	$\mathcal{O}(10^2)$	$\mathcal{O}(10^{-5})$
0.05	$\mathcal{O}(10^4)$	$\mathcal{O}(10^3)$	$\mathcal{O}(10^2)$	$\mathcal{O}(10^{-4})$	$\mathcal{O}(10^4)$	$\mathcal{O}(10^3)$	$\mathcal{O}(10^2)$	$\mathcal{O}(10^{-4})$

Abbildung 5.20: Lage der Singulärwerte für Simulationen aus Abbildung 5.18 und 5.19

zwischen den künstlichen und den echten Singulärwerten.

Wir schließen diesen Abschnitt mit einem weiteren Beispiel. Da wir bereits die Kandidaten aus Abbildung 5.7 erfolgreich validiert haben, verwenden wir abermals das PmL Verfahren, um weitere Kandidaten zu erhalten. Wir befinden uns also in der Ausgangssituation von Abbildung 5.21. Analog zu den bisherigen Experimenten legen wir nun eine Kontur um die

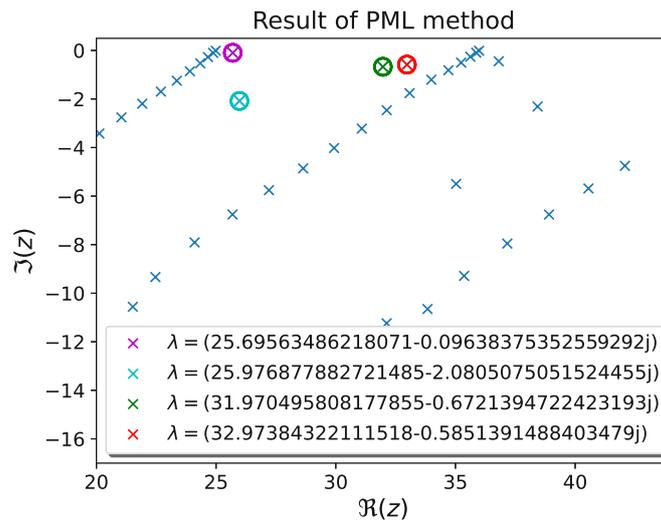


Abbildung 5.21: Ergebnisse der Eigenwertberechnung mittels PmL Methode

farbig markierten Kandidaten (vergleiche auch mit Abbildung 5.22). Aufgrund der Lage der Eigenwerte sind wir außerdem gezwungen eine Kurve zu wählen, die nahe an einem Branch-Cut liegt. Wie wir aus den bisherigen Experimenten bereits wissen, hat dies jedoch ad-hoc keinen signifikanten Einfluss auf die Konvergenz. Für die weiteren Experimente wurde die nachfolgende Konfiguration verwendet:

- Quadraturpunkte $N = 40$, $t_j := j/N$ für $j = 0, \dots, N - 1$
- Äußere Freiheitsgrade $N_{\text{ext}} = 7$ für DtN Operator
- Anzahl zu berücksichtigter DtN Nummern $L_{\text{max}} = 12$ für Optimierungsproblem

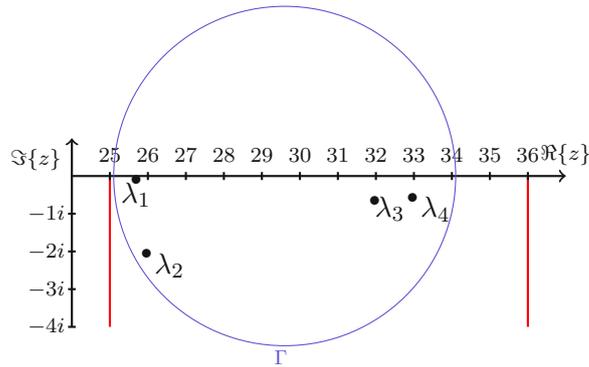


Abbildung 5.22: Integrationsweg Γ für die Berechnung der Eigenwerte aus Abb. 5.21

- Anzahl der Spalten $J = 6$ von $\hat{V} \in \mathbb{C}^{m \times J}$
- Integrationsweg $\Gamma(t) := 29.6 + 4.5 \exp(2\pi it)$
- Gewichtsfunktion für $\kappa_j := \Gamma(t_j)$:

$$\omega_{jn} = \begin{cases} 1, & n^2 \leq |\Re(\kappa_j)| \\ |\exp(i\sqrt{\kappa_j - n^2} \cdot (\pi - 5/8\pi))|, & n^2 > |\Re(\kappa_j)| \end{cases}$$

Bevor wir die Resultate präsentieren, diskutieren wir noch die Wahl der Parameter etwas genauer. Zunächst wurde die Anzahl der Quadraturpunkte im Vergleich zu den bisherigen Simulationen wieder erhöht, was durch die größere Länge der Kontur Γ gerechtfertigt werden kann. Die Anzahl der zu erwartenden geometrischen Vielfachheiten der Eigenwerten innerhalb der Kurve $L = 4$ erfordert zumindest $J \geq 4$. Für die nachfolgenden Untersuchungen, auch in Hinblick der Analyse der Lage der Singulärwerte, macht jedoch natürlich ein größerer Wert Sinn. An dieser Stelle sei jedoch ebenso bemerkt, dass eine größere Anzahl der Spalten J von \hat{V} auch mit zusätzlichem Aufwand verbunden ist. Die nachfolgenden

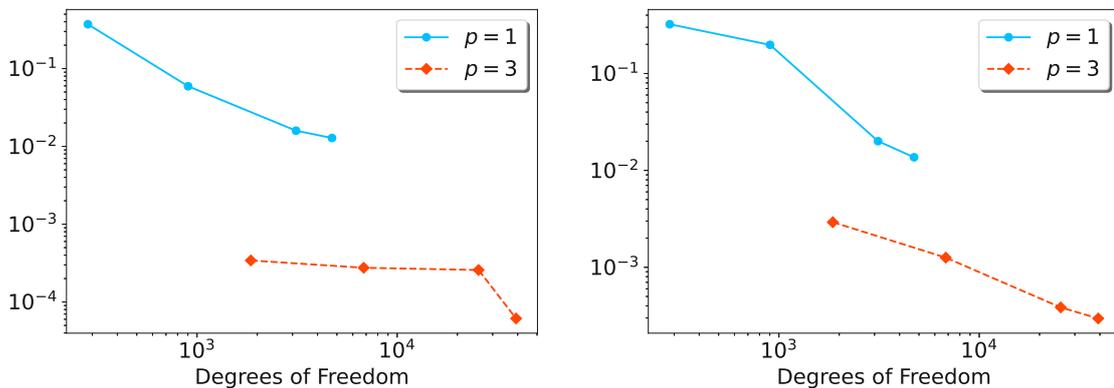


Abbildung 5.23: Links: Fehler $|\lambda_1 - \lambda_1^{(h)}|$, Rechts: Fehler $|\lambda_2 - \lambda_2^{(h)}|$

Abbildungen beinhalten den Fehler in Abhängigkeit der Freiheitsgraden, wobei erneut die Fehler bezüglich der durch das PmL Verfahren generierten Referenzwerte (mit $h = 0.05$ und Polynomordnung $p = 8$) gemessen werden. Ähnlich wie schon zuvor beobachten wir Unterschiede der Fehler zwischen den einzelnen Eigenwerten. Während für $p = 1$ der Fehler in Abhängigkeit der Freiheitsgrade zunehmend abnimmt, scheint dies für $p = 3$ nicht unbedingt der Fall zu sein. Dies kann zum einen an der Wahl der Parameter für die Approximation des DtN Operators liegen, zum anderen auch direkt an dem Fehler der Eigenwertberechnung der Matrix $B \in \mathbb{C}^{4 \times 4}$. Wir werfen nun noch einen Blick auf die resultierenden

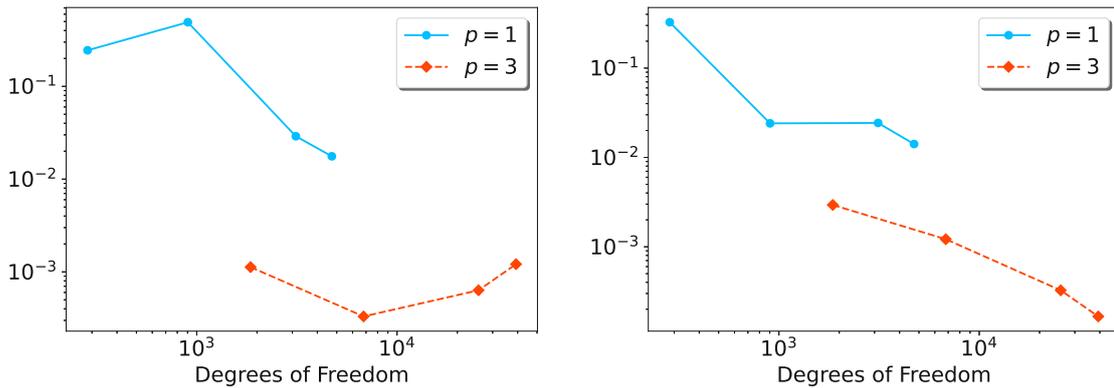


Abbildung 5.24: Links: Fehler $|\lambda_3 - \lambda_3^{(h)}|$, Rechts: Fehler $|\lambda_4 - \lambda_4^{(h)}|$

Singulärwerte. Wie gewohnt, sind die diese in absteigender Reihenfolge sortiert, womit wir in der Tabelle 5.25 nur den größten Singulärwert σ_1 , den Gap zwischen σ_4 und σ_5 sowie σ_6 darstellen werden. Die Größenordnungen der Singulärwerte ermöglichen auch in diesem

mesh	$p = 1$				$p = 3$			
h	σ_1	σ_4	σ_5	σ_6	σ_1	σ_4	σ_5	σ_6
0.25	$\mathcal{O}(10^2)$	$\mathcal{O}(10^1)$	$\mathcal{O}(1)$	$\mathcal{O}(10^{-5})$	$\mathcal{O}(10^3)$	$\mathcal{O}(10^1)$	$\mathcal{O}(10^{-1})$	$\mathcal{O}(10^{-2})$
0.125	$\mathcal{O}(10^3)$	$\mathcal{O}(10^1)$	$\mathcal{O}(1)$	$\mathcal{O}(10^{-2})$	$\mathcal{O}(10^3)$	$\mathcal{O}(10^1)$	$\mathcal{O}(10^{-1})$	$\mathcal{O}(10^{-1})$
0.0625	$\mathcal{O}(10^3)$	$\mathcal{O}(10^2)$	$\mathcal{O}(10^1)$	$\mathcal{O}(10^{-3})$	$\mathcal{O}(10^3)$	$\mathcal{O}(10^2)$	$\mathcal{O}(1)$	$\mathcal{O}(10^{-1})$
0.05	$\mathcal{O}(10^4)$	$\mathcal{O}(10^2)$	$\mathcal{O}(10^1)$	$\mathcal{O}(10^{-3})$	$\mathcal{O}(10^4)$	$\mathcal{O}(10^2)$	$\mathcal{O}(1)$	$\mathcal{O}(10^{-1})$

Abbildung 5.25: Lage der Singulärwerte für Simulationen aus Abbildung 5.23 bzw. 5.24

Beispiel eine deutliche Trennung der künstlichen Werte, insbesondere liegen all diese in einer Größenordnung von maximal 10. Abschließend versuchen wir die in Abbildung 5.23 und 5.24 dargestellten Konvergenzen zu verbessern, indem wir zusätzliche Konturen Γ_1 und Γ_2 einführen, die näher an den gesuchten Eigenwerten liegen. Diese Vorgehensweise kann auch bei einer unbekanntem Lage von Eigenwerten verwendet werden: man versucht mit einer großen Kurve zu starten und die aus jener Simulation stammenden Eigenwerte durch gezieltes Setzen einer kleineren Kontur sukzessive zu bestätigen. In unserem Fall also

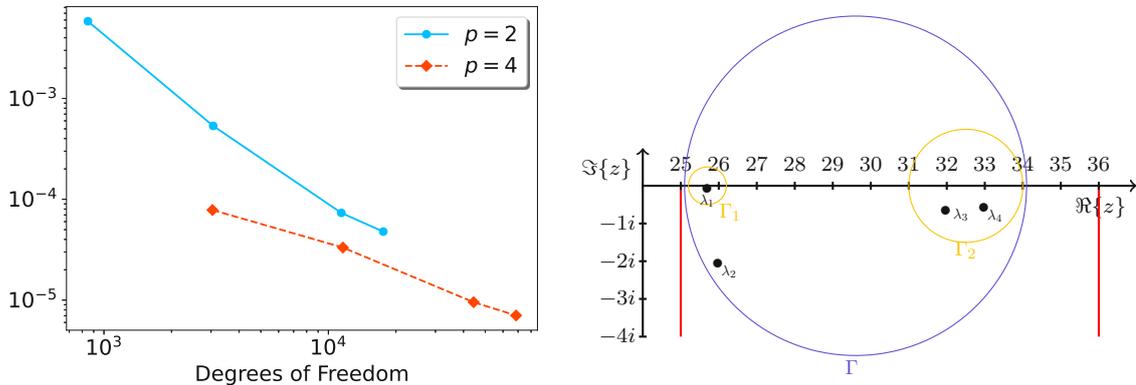


Abbildung 5.26: Links: Fehler $|\lambda_1 - \lambda_1^{(h)}|$, Rechts: Integrationswege Γ_1 und Γ_2 für die Integrationsmethode, Modellparameter: $N = 16, J = 3, N_{ext} = 5, L_{max} = 10$

führen wir die folgenden Kurven ein:

$$\Gamma_1 := 25.7 + 0.5 \exp(2\pi i \cdot), \quad \Gamma_2 := 32.5 + 1.5 \exp(2\pi i \cdot).$$

Diese Wahl wird in Abbildung 5.26 visualisiert. Natürlich können wir auf Grund der kleineren Konturen die Anzahl der Quadraturpunkte reduzieren, gleichzeitig erhöhen wir auch die Polynomordnung p . Die restlichen Modellparameter bleiben nahezu unverändert. Wie in Abbildung 5.26 und 5.27 unmittelbar zu beobachten ist, resultiert diese Wahl in einem kleineren Fehler als in den vorausgehenden Experimenten.

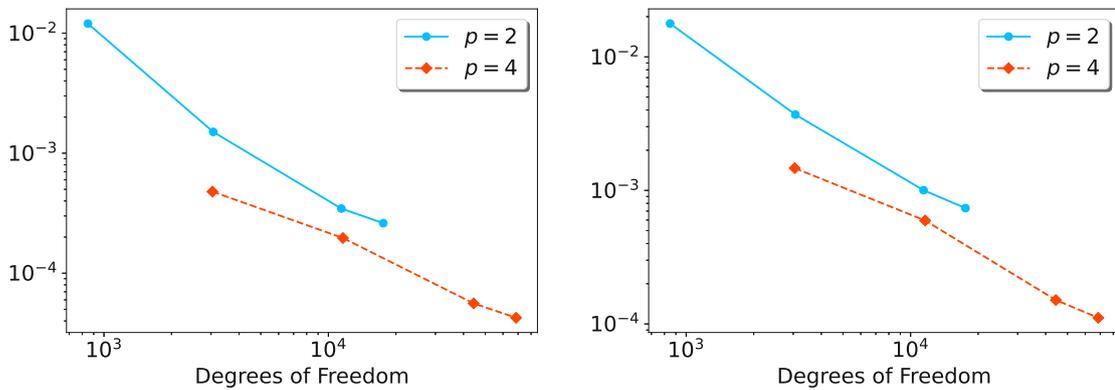


Abbildung 5.27: Links: Fehler $|\lambda_3 - \lambda_3^{(h)}|$, Rechts: Fehler $|\lambda_4 - \lambda_4^{(h)}|$, Modellparameter: $N = 24, J = 4, N_{ext} = 5, L_{max} = 10$

5.3.4 Weitere Experimente

In den bisher präsentierten Experimenten konnten wir zwar die Konvergenz der approximierten Eigenwerte nachweisen, jedoch schien die Geschwindigkeit dieser nur bedingt von der Polynomordnung der verwendeten Spline Räume $\mathcal{S}^p(\mathcal{T})$ abzuhängen. Der Grund dafür scheint an der speziellen Gestalt der gewählten Geometrie Ω_{int} zu liegen, da selbst der L^2 Fehler von Streuproblemen das gleiche Verhalten aufweist. In Kapitel 4.2.3 haben wir (zwar

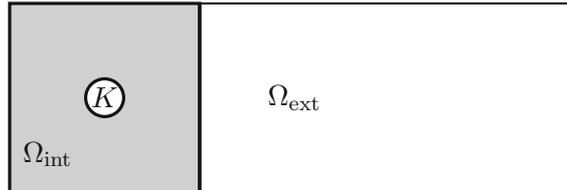


Abbildung 5.28: Unbeschränktes Gebiet zur Berechnung von Eigenwerten

nur für reelle Frequenzen ω) in den dort definierten Geometrien für Streuprobleme in der Tat unterschiedliche Konvergenzgeschwindigkeiten bezüglich der Polynomordnung nachweisen können. Dies lässt nun ebenfalls ein ähnliches Verhalten für die Berechnung der Eigenwerte mittels der Integralmethode erhoffen. Tatsächlich konnten wir mit einer ähnlichen Geometrie, wie in (4.35) verwendet, in einem weiteren Experiment teilweise unterschiedliche Geschwindigkeiten nachweisen. Wie in Abbildung 5.28 ersichtlich ist, wurde für die nachfolgenden Experimente der Streukörper aus Abschnitt 5.3.1 durch $K := B_{\frac{\pi}{10}}(\frac{\pi}{2}, \frac{\pi}{2})$ ersetzt.

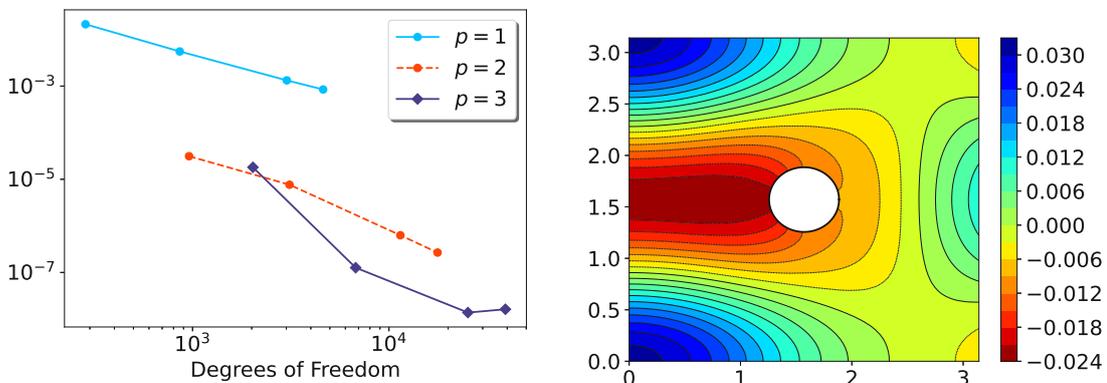


Abbildung 5.29: Links: Fehler $|\lambda_{\text{ref}} - \lambda^{(h)}|$ in Abhängigkeit der Freiheitsgrade, Rechts: Visualisierung der dazugehörigen Eigenfunktion u_h , $\lambda_{\text{ref}} \approx 4.5 - 0.25i$

Bemerkung 5.5. Da wir in diesem Abschnitt die Geometrie nicht nur um einen Faktor skalierten, verändert sich natürlich ad-hoc auch die Lage der zu ermittelnden Eigenwerte. Tatsächlich haben wir durch erneute Verwendung der PmL Methode eine sehr ähnliche Ausgangslage wie jener in Abbildung 5.7 vorfinden können. Dies hängt im Wesentlichen damit zusammen, dass die Eigenwerte stetig von der Konfiguration des zu untersuchenden

Modells abhängen. Insbesondere kann man also annehmen, dass eine „kleine“ Änderung der Geometrie in einer „kleinen Änderung“ der Lage der Eigenwerte resultiert. Diese Annahmen werden in den folgenden Experimenten ebenfalls bestätigt.

Somit ist klar, in welchen Bereichen wir nach möglichen Eigenwerten suchen müssen, wenn wir die in Kapitel 5.3.2 ermittelten Werte als Ausgangspunkt nehmen. Wir präsentieren an dieser Stelle nur die Resultate, da wir die Konfiguration der Modell-Parameter in diesem Abschnitt nur leicht verändert haben im Vergleich zu den Beispielen aus Kapitel 5.3.2. Der Referenzwert wurde diesmal im Vergleich zu den bisherigen Abschnitten ebenfalls mit der Integralmethode mit Polynomordnung $p = 6$ berechnet. Wie in Abbildungen 5.29 sowie 5.30 ersichtlich wird, können wir nun erstmals signifikante Unterschiede in der Geschwindigkeit der Konvergenz erkennen. Dies hängt nun ebenfalls mit den in Kapitel 4.2.3 vorgestellten Resultaten hinsichtlich der Konvergenz von Lösungen von Streuproblemen zusammen und bestätigt unsere zuvor genannte Vermutung.

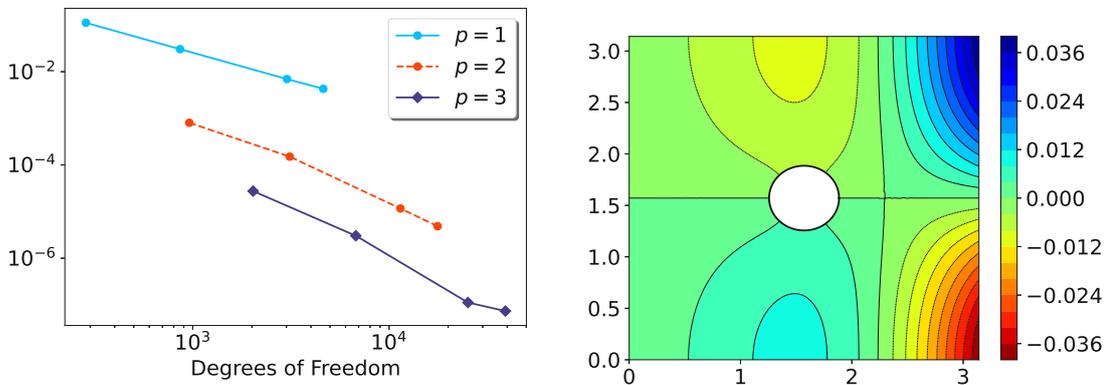


Abbildung 5.30: Links: Fehler $|\lambda_{ref} - \lambda^{(h)}|$ in Abhängigkeit der Freiheitsgrade, Rechts: Visualisierung der dazugehörigen Eigenfunktion u_h , $\lambda_{ref} \approx 1.98 - 1.84j$

Zusammenfassend konnten wir also mit Hilfe der Integralmethode in Kombination mit der Learned Infinite Elemente Diskretisierung sämtliche durch die PmL Methode generierten Kandidaten bestätigen. Der Gebrauch der PmL Methode diente dabei lediglich zur Kontrolle der ermittelnden Werte und spielte klarerweise in der eigentlichen Berechnung keine Rolle.

Was die Approximation des DtN Operators anbelangt, konnten wir in fast allen Bereichen feststellen, dass erstaunlicherweise bereits eine geringe Anzahl N_{ext} an externen Freiheitsgraden zu brauchbaren Konvergenzen der Eigenwerte führte. Besonders wichtig bei der Berechnung der Eigenwerte ist natürlich eine sinnvolle Wahl der Parameter L_{max} sowie einer entsprechenden Gewichtsfunktion ω_n , kurz gesagt also die Wahl der Learned Infinite Elemente Methode relevanten Parameter. Um diese geeignet zu wählen, muss man berücksichtigen, dass die Matrizen A_0 sowie A_1 , wie in (5.3) beschrieben, genau aus den Lösungen bestimmter Streuprobleme, siehe auch (5.26), mit zufälliger rechter Seite bestehen. Somit ist es naheliegend, dass man zuerst jene Lösungen untersuchen und verstehen muss. Dies entspricht genau jener Vorgehensweise, die wir in dieser Arbeit präsentiert ha-

ben. Abschließend wollen wir noch zur Veranschaulichung ein paar Eigenfunktionen zu den bereits berechneten Eigenwerten aus den Abschnitten 5.3.2 und 5.3.3 visualisieren.

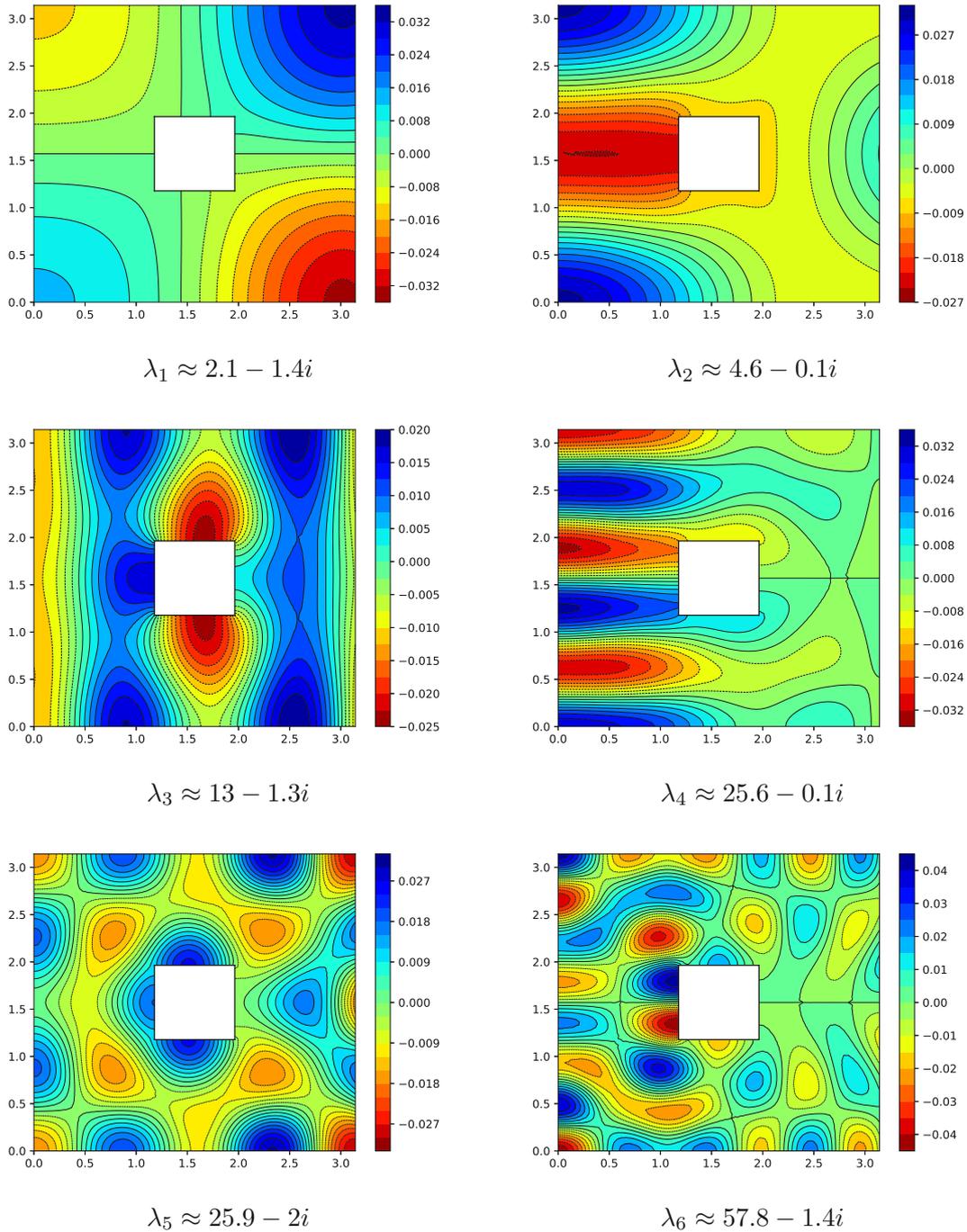


Abbildung 5.31: Visualisierung der Eigenfunktionen $u \in \ker A(\lambda_i)$ für verschiedene λ_i

Literatur

- [Gol82] Charles I. Goldstein. „A finite element method for solving Helmholtz type equations in waveguides and other unbounded domains“. In: *Math. Comp.* 39.160 (1982), S. 309–324. ISSN: 0025-5718. DOI: [10.2307/2007317](https://doi.org/10.2307/2007317). URL: <https://doi.org/10.2307/2007317>.
- [Lei86] Rolf Leis. *Initial-boundary value problems in mathematical physics*. B. G. Teubner, Stuttgart; John Wiley & Sons, Ltd., Chichester, 1986, S. viii+266. ISBN: 3-519-02102-1. DOI: [10.1007/978-3-663-10649-4](https://doi.org/10.1007/978-3-663-10649-4). URL: <https://doi.org/10.1007/978-3-663-10649-4>.
- [Hac92] W. Hackbusch. *Elliptic differential equations*. Bd. 18. Springer Series in Computational Mathematics. Theory and numerical treatment, Translated from the author’s revision of the 1986 German original by Regine Fadiman and Patrick D. F. Ion. Springer-Verlag, Berlin, 1992, S. xiv+311. ISBN: 3-540-54822-X. DOI: [10.1007/978-3-642-11490-8](https://doi.org/10.1007/978-3-642-11490-8). URL: <https://doi.org/10.1007/978-3-642-11490-8>.
- [Jän93] Klaus Jänich. *Funktionentheorie*. Third. Springer-Lehrbuch. [Springer Textbook]. Eine Einführung. [An introduction]. Springer-Verlag, Berlin, 1993. ISBN: 3-540-56337-7. DOI: [10.1007/978-3-662-11803-0](https://doi.org/10.1007/978-3-662-11803-0). URL: <https://doi.org/10.1007/978-3-662-11803-0>.
- [CK98] David Colton und Rainer Kress. *Inverse acoustic and electromagnetic scattering theory*. Second. Bd. 93. Applied Mathematical Sciences. Springer-Verlag, Berlin, 1998, S. xii+334. ISBN: 3-540-62838-X. DOI: [10.1007/978-3-662-03537-5](https://doi.org/10.1007/978-3-662-03537-5). URL: <https://doi.org/10.1007/978-3-662-03537-5>.
- [Ihl98] Frank Ihlenburg. *Finite element analysis of acoustic scattering*. Bd. 132. Applied Mathematical Sciences. Springer-Verlag, New York, 1998, S. xiv+224. ISBN: 0-387-98319-8. DOI: [10.1007/b98828](https://doi.org/10.1007/b98828). URL: <https://doi.org/10.1007/b98828>.
- [BG99] A. Bendali und Ph. Guillaume. „Non-reflecting boundary conditions for waveguides“. In: *Math. Comp.* 68.225 (1999), S. 123–144. ISSN: 0025-5718. DOI: [10.1090/S0025-5718-99-01016-9](https://doi.org/10.1090/S0025-5718-99-01016-9). URL: <https://doi.org/10.1090/S0025-5718-99-01016-9>.
- [DI01] Leszek Demkowicz und Frank Ihlenburg. „Analysis of a coupled finite-infinite element method for exterior Helmholtz problems“. In: *Numer. Math.* 88.1 (2001), S. 43–73. ISSN: 0029-599X. DOI: [10.1007/PL00005440](https://doi.org/10.1007/PL00005440). URL: <https://doi.org/10.1007/PL00005440>.
- [Néd01] Jean-Claude Nédélec. *Acoustic and electromagnetic equations*. Bd. 144. Applied Mathematical Sciences. Integral representations for harmonic problems. Springer-Verlag, New York, 2001, S. x+316. ISBN: 0-387-95155-5. DOI: [10.1007/978-1-4757-4393-7](https://doi.org/10.1007/978-1-4757-4393-7). URL: <https://doi.org/10.1007/978-1-4757-4393-7>.

- [HN09] Thorsten Hohage und Lothar Nannen. „Hardy space infinite elements for scattering and resonance problems“. In: *SIAM J. Numer. Anal.* 47.2 (2009), S. 972–996. ISSN: 0036-1429. DOI: [10.1137/070708044](https://doi.org/10.1137/070708044). URL: <https://doi.org/10.1137/070708044>.
- [Hoa11a] Vu Hoang. „A RADIATION CONDITION FOR A PERIODIC SEMI-INFINITE WAVEGUIDE“. In: (2011). URL: <https://www.math.kit.edu/iana1/~hoang/media/radwaveguidehoang.pdf>.
- [Hoa11b] Vu Hoang. „The limiting absorption principle for a periodic semi-infinite waveguide“. In: *SIAM J. Appl. Math.* 71.3 (2011), S. 791–810. ISSN: 0036-1399. DOI: [10.1137/100791798](https://doi.org/10.1137/100791798). URL: <https://doi.org/10.1137/100791798>.
- [Bey12] Wolf-Jürgen Beyn. „An integral method for solving nonlinear eigenvalue problems“. In: *Linear Algebra Appl.* 436.10 (2012), S. 3839–3863. ISSN: 0024-3795. DOI: [10.1016/j.laa.2011.03.030](https://doi.org/10.1016/j.laa.2011.03.030). URL: <https://doi.org/10.1016/j.laa.2011.03.030>.
- [Kre14] Rainer Kress. *Linear integral equations*. Third. Bd. 82. Applied Mathematical Sciences. Springer, New York, 2014, S. xvi+412. DOI: [10.1007/978-1-4614-9593-2](https://doi.org/10.1007/978-1-4614-9593-2). URL: <https://doi.org/10.1007/978-1-4614-9593-2>.
- [Jün19] Ansgar Jüngel. *Partielle Differentialgleichungen*. Institut für Analysis und Scientific Computing, TU Wien, 2019.
- [Har+20] Charles R. Harris u. a. „Array programming with NumPy“. In: *Nature* 585.7825 (Sep. 2020), S. 357–362. DOI: [10.1038/s41586-020-2649-2](https://doi.org/10.1038/s41586-020-2649-2). URL: <https://doi.org/10.1038/s41586-020-2649-2>.
- [Jün20] Ansgar Jüngel. *Nichtlineare Partielle Differentialgleichungen*. Institut für Analysis und Scientific Computing, TU Wien, 2020.
- [HLP21] Thorsten Hohage, Christoph Lehrenfeld und Janosch Preuß. „Learned infinite elements“. In: *SIAM J. Sci. Comput.* 43.5 (2021), A3552–A3579. ISSN: 1064-8275. DOI: [10.1137/20M1381757](https://doi.org/10.1137/20M1381757). URL: <https://doi.org/10.1137/20M1381757>.
- [Jan21] Preuß Janosch. *Learned IE*. https://gitlab.gwdg.de/learned_infinite_elements/learned_ie. 2021.
- [NFP21] Lothar Nannen, Michael Feischl und Dirk Praetorius. *Numerics of Partial Differential Equations: stationary Problems*. Institut für Analysis und Scientific Computing, TU Wien, 2021.
- [AMT22] Sameer Agarwal, Keir Mierle und The Ceres Solver Team. *Ceres Solver*. Version 2.1. März 2022. URL: <https://github.com/ceres-solver/ceres-solver>.
- [Nan22] Lothar Nannen. *Eigenwertprobleme*. Institut für Analysis und Scientific Computing, TU Wien, 2022.
- [SCH22] BEN SCHWEIZER. „Inhomogeneous Helmholtz equations in wave guides – existence and uniqueness results with energy methods“. In: *European Journal of Applied Mathematics* (2022), S. 1–27. DOI: [10.1017/S0956792522000080](https://doi.org/10.1017/S0956792522000080).