



UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

**FAKULTÄT FÜR
ELEKTROTECHNIK,
INFORMATIK UND
MATHEMATIK**

Zur Berechnung elektromagnetischer Eigenwertprobleme in der numerischen Simulation von Nanostrukturen mit periodischen und transparenten Randbedingungen

Von der Fakultät für Elektrotechnik, Informatik und Mathematik
der Universität Paderborn

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften (Dr.-Ing.)

genehmigte Dissertation

von

Dipl.-Ing. Bastian Tobias Bandlow

Erster Gutachter:
Zweiter Gutachter:

Prof. Dr.-Ing. Rolf Schuhmann
Prof. Dr.-Ing. Thomas Weiland

Tag der mündlichen Prüfung: 31.03.2011

Paderborn 2011

Diss. EIM-E/274

Vorwort

Diese Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Fachgebiet Theoretische Elektrotechnik der Universität Paderborn. Bedanken möchte ich mich an dieser Stelle bei all denjenigen, die auf unterschiedliche Weise zum erfolgreichen Gelingen meiner Arbeit beigetragen haben. Insbesondere möchte ich folgenden Personen danken:

- Herrn Prof. Dr.-Ing. Rolf Schuhmann für die Ermöglichung dieser Arbeit, für die hervorragende wissenschaftliche Betreuung, für die Ermöglichung der Teilnahme an Konferenzen, Tagungen und Forschungsreisen sowie für das hohe Maß an Freiheit und Selbstbestimmung, das seine Mitarbeiter erfahren.
- Herrn Prof. Dr.-Ing. Thomas Weiland für die freundliche Übernahme des Korreferats und den damit verbundenen Mühen bei der Ausarbeitung des zweiten Gutachtens,
- Herrn Prof. Dr. Igor A. Tsukerman for having a good time at The University of Akron,
- Herrn Dr.-Ing. Matthias Stallein für mehr als fünf Jahre Kollegentum, Koautorentum und die Durchsicht des Manuskripts,
- Herrn Dr.-Ing. Grzegorz Lubkowski für das kollegiale Maß an Konkurrenz bei der Forschung zu Metamaterialien und die resultierenden gemeinsamen Veröffentlichungen,
- Herrn M. Sc. Christoph Claßen für seinen an den Tag gelegten „Biss“ bei der zeitnahen Fertigstellung gemeinsamer Veröffentlichungen sowie für seine Denkanstöße zu FLAME und das Gegenlesen des Manuskripts,
- Herrn Dipl.-Ing. Stefan Kirsch für das Gegenlesen des Manuskripts,
- Frau Nina Kuthan und Marcel für die gefundenen Rechtschreibfehler und monierten Ausdrucksformen („man“),
- Frau Gabriele Freitag für die langjährige Unterstützung bei vielen organisatorischen und verwaltungstechnischen Fragen,
- allen weiteren Kollegen aus Paderborn,
- allen Kollegen vom Institut für Theorie elektromagnetischer Felder der TU Darmstadt.

Mein spezieller Dank gilt meiner Freundin Bernadette Pulsmeier für ihre Geduld, die liebevolle Rücksichtnahme und entgegengebrachtes Verständnis während der Promotionsphase.

Paderborn, im April 2011

Bastian Bandlow

Inhaltsverzeichnis

1. Einleitung	1
1.1. Motivation und Zielsetzung	2
1.2. Vorgehensweise	3
1.3. Gliederung	4
2. Elektromagnetische Wellenausbreitung	7
2.1. Maxwell'sche Gleichungen	7
2.1.1. Stetigkeit an Materialgrenzen	8
2.1.2. Wellengleichung	8
2.1.3. Dispersion und Kausalität	10
2.1.4. Energieerhaltung	16
2.1.5. Orthogonalität und Streumatrizen	18
2.2. Methode der finiten Integration (FIT)	20
2.2.1. Räumlich diskrete Gitter-Maxwell-Gleichungen	20
2.2.2. Zeitbereich	24
2.2.3. Frequenzbereich	24
2.2.4. Randbedingungen	25
2.2.5. Fehler	29
2.3. Methode der flexiblen lokalen Approximation (FLAME)	30
2.3.1. Differenzenschemata in FLAME	30
2.3.2. Eigenschaften	31
2.3.3. Fehler und Konvergenz	32
2.3.4. Minimalbeispiel	33
3. Formulierungen für periodische und isolierte Nanostrukturen	37
3.1. Streumatrix-Ansatz zur Berechnung der Dispersion in periodischen Strukturen	37
3.1.1. Streumatrix für fundamentale Wellenleitermoden	38
3.1.2. Streumatrix für mehrere Wellenleitermoden	38
3.2. Eigenwertformulierungen zur Berechnung der Dispersion in periodischen Strukturen	39
3.2.1. Verlustfreie Elementarzellen	40
3.2.2. Verlustbehaftete Elementarzellen, zeitliches Abklingverhalten	40
3.2.3. Verlustbehaftete Elementarzellen, räumliches Abklingverhalten	41
3.3. Berechnung von Ausbreitungskonstanten in längshomogenen Wellenleitern	43
3.3.1. Wellenleiter mit transversal transparenten Rändern	44
3.4. Eigenwerte isolierter Strukturen im Freiraum	44
3.5. Eigenwerte skalarer 2D Problemstellungen mit FLAME	45
3.5.1. Struktur der Systemmatrix	46
3.5.2. Eigenwertformulierung in FLAME	46

4. Numerische Methoden zur Berechnung algebraischer Eigenwertprobleme	49
4.1. Einführung	51
4.1.1. Klassifizierung von Eigenwertproblemen	51
4.1.2. Kenngrößen	54
4.1.3. Fehler-Definitionen	56
4.1.4. Kondition	57
4.2. Löser für das lineare Eigenwertproblem	59
4.2.1. Iterative Löser für einzelne Eigenwerte des Standard eigenwertproblems	60
4.2.2. Direkte Löser für das Standard eigenwertproblem	63
4.2.3. Direkte Löser für das allgemeine Eigenwertproblem	64
4.3. Grundlagen des Jacobi-Davidson-Verfahrens für Standard eigenwertprobleme	65
4.3.1. Einführung	66
4.3.2. Verwandtschaft mit dem Newton-Verfahren	67
4.3.3. Deflation	69
4.3.4. Neustarts	69
4.3.5. Berechnung äußerer Eigenwerte nicht-hermitescher Matrizen	70
4.3.6. Berechnung innerer Eigenwerte nicht-hermitescher Matrizen	72
4.3.7. Lösung der Korrekturgleichung	75
4.4. Löser für lineare Gleichungssysteme	76
4.4.1. Direkte Löser	76
4.4.2. Iterative Löser	77
4.4.3. Vorkonditionierung	77
4.4.4. Umindizierung der FIT-Matrizen für ILU-Zerlegungen	78
4.5. Erweiterung des Selektionsprozesses im Jacobi-Davidson-Verfahren	79
4.5.1. Selektionskriterien zur gezielten Auswahl gesuchter Eigenvektoren	79
4.5.2. Erweiterter Selektionsprozess	81
4.5.3. Selektion bei Berechnung äußerer Eigenwerte	82
4.5.4. Selektion bei Berechnung innerer Eigenwerte	83
4.6. Löser für polynomiale Eigenwertprobleme	84
4.6.1. Linearisierung polynomialer Eigenwertprobleme	84
4.6.2. Newton-Verfahren	85
4.6.3. Jacobi-Davidson-Verfahren für polynomiale Eigenwertprobleme	85
4.6.4. Implementierungsdetails des Jacobi-Davidson-Verfahren für polynomi- ale Eigenwertprobleme	87
4.7. Löser für nichtlineare Eigenwertprobleme	89
4.7.1. Fixpunkt-Iteration	90
4.7.2. Sekanten-Verfahren	91
5. Ergebnisse und Anwendungen	93
5.1. Dispersionseigenschaften von Elementarzellen im Zentimeterbereich	93
5.1.1. Simulationsansätze	95
5.1.2. Numerische Ergebnisse	96
5.1.3. Fazit	99
5.2. Dispersionseigenschaften von Elementarzellen im Nanometerbereich	100
5.2.1. Linkshändige Metamaterialien bei optischen Frequenzen	101
5.2.2. Lösung der Eigenwert-Formulierung	103
5.2.3. Numerische Ergebnisse	104
5.2.4. Fazit	106

5.3.	Berechnung von Wellenleitermoden in photonischen Kristallfasern	107
5.3.1.	Apriori Wissen und physikalisch motiviertes Selektionskriterium	108
5.3.2.	Numerische Konfiguration	110
5.3.3.	Performanzvergleich	111
5.3.4.	Fazit	112
5.4.	Eigenmoden isolierter Nanostrukturen im Freiraum	114
5.4.1.	Nanoresonatoren aus photonischen Kristallen	114
5.4.2.	Resonanzen in Microdisks	120
5.4.3.	Fehlereinfluss durch linearisierte PML	128
5.5.	Eigenwertberechnung zweidimensionaler Nanoresonatoren mit FLAME	131
5.5.1.	Nanoresonator in FLAME-Diskretisierung	131
5.5.2.	Konvergenzordnung für die Streuung am Zylinder	132
5.5.3.	Lösung des FLAME-Eigenwertproblems	133
5.5.4.	Fazit	135
6.	Zusammenfassung und Ausblick	137
A.	Anhang	143
A.1.	Berechnung der Dispersionskurve aus der Streumatrix	143
A.2.	Matrizendefinitionen	144
A.3.	Methoden zur Eigenpaarberechnung von Davidson und Jacobi	144
A.4.	Geometrie der Elementarzellen im Zentimeterbereich	148
	Abkürzungs- und Symbolverzeichnis	151
	Literaturverzeichnis	157

1. Einleitung

Während die Datenübertragung durch Lichtwellenleiter sowie die Datenspeicherung heutzutage optisch durchgeführt werden, stehen die verfügbaren programmierbaren Recheneinheiten zur Informationsverarbeitung weiterhin auf einer elektronischen Basis. Dies bedeutet unweigerlich, dass eine Schnittstelle zwischen Optik und Elektronik notwendig ist. Das Gebiet der Optoelektronik und Photonik befasst sich mit genau dieser Schnittstelle, deren Absicht die gezielte Beeinflussung von Licht auf Basis der Elektrodynamik ist. Laser-Pointer und Leuchtdioden sind zwei Beispiele für aktive Bauelemente der Photonik, die inzwischen im alltäglichen Leben verwendet werden. Darüber hinaus gibt es eine Reihe passiver Bauelemente, die im Folgenden näher betrachtet werden.

Eine Form passiver Bauelemente stellen Materialien mit geometrisch komplexem Aufbau dar. Die Forschungen auf diesem Gebiet sind seit der Jahrtausendwende eng mit dem Begriff des *Metamaterials* verwoben. *Metamaterialien* sind keine Materialien im herkömmlichen Sinn, sondern künstlich hergestellte, meist periodische Strukturen aus Elementarzellen, die sich in einem umgebenden Medium befinden. Das Interesse von Forschergruppen unterschiedlichster Disziplinen an *Metamaterialien* liegt darin begründet, die elektromagnetischen Eigenschaften eines Materials maßgeschneidert entwerfen und herstellen zu können. Die elektromagnetischen Eigenschaften des *Metamaterials* werden weniger durch die Mittelwerte der elementaren chemischen Bestandteile der Elementarzellen bestimmt, sondern sind stark von der Geometrie der in der Elementarzelle enthaltenen Substruktur abhängig. Die Substruktur kann dabei sowohl aus Dielektrika als auch aus Metallen bestehen. Experimentell konnte im Mikrowellenbereich Brechung in negative Richtung an der Grenzfläche zu *Metamaterialien* beobachtet werden [114], die bei homogenen natürlichen Materialien nicht auftritt. Eine solche *Metamaterial*-Struktur ist in Abbildung 1.1 dargestellt. Das *Metamaterial* besteht aus einer Kombination von metallischen *Split-Ring-Resonatoren* [109] und Dipolen, die jeweils als magnetische bzw. elektrische Resonatoren oberhalb ihrer Resonanzfrequenz betrieben werden. Seit Anbeginn der Forschung auf diesem Gebiet wird versucht, *Metamaterialien* effektive Materialparameter zuzuschreiben. Insbesondere für den Fall der Brechung in negative Richtung hat dies zu recht exotischen, negativen Werten der effektiven Materialparameter geführt¹, die in der Natur bei homogenen Materialien nicht beobachtet werden.

Ein weiterer Typ von periodischen Strukturen stellen die *photonischen Kristalle* dar [4, 103, 122]. Die periodische Anordnung von (meist) dielektrischen Materialien führt zu einer elektromagnetischen Bandlücke. Bei Frequenzen innerhalb der Bandlücke ist keine elektromagnetische Wellenausbreitung möglich, so dass *photonische Kristalle* auch als optische Halbleiter bezeichnet werden. Die Elementarzelle *photonischer Kristallgitter* hat typischerweise die Abmessung einer halben Wellenlänge. Aufgrund der Skalierbarkeit der MAXWELL'schen Gleichungen lassen sich *photonische Kristalle* für beliebige Frequenzen entwerfen, wobei der Fokus der Forschung auf Frequenzen im optischen Bereich liegt. Die Eigenschaft der elektromagnetischen Bandlücke, keine Wellenausbreitung zuzulassen, wird verwendet, um Licht

¹Die negativen Materialparameter approximieren das elektromagnetische Verhalten der Struktur jeweils in einem sehr kleinen Frequenzband.

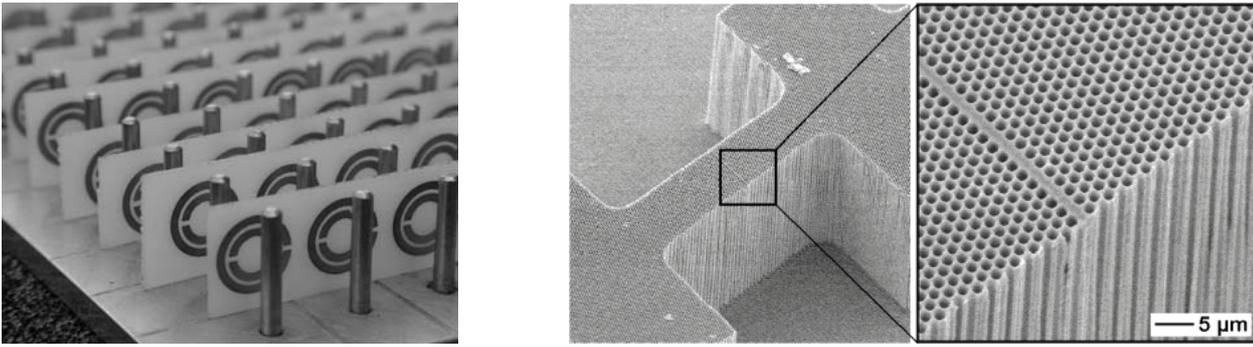


Abbildung 1.1.: Links: Strukturgitter eines *Metamaterials* bestehend aus metallischen *Split-Ring-Resonatoren* und Dipolen (Quelle: [119]). Mitte und rechts: Mikroskopisch kleine *photonische Kristallstruktur* aus einem perforierten Dielektrikum mit linienförmigem Defekt zur Wellenleitung (Quelle: [92]).

kontrolliert zu führen. Dazu werden in das periodische *photonische Kristallgitter* linienförmige Defekte eingebracht, die einen Wellenleiter bilden. In Abbildung 1.1 ist ein *photonischer Kristall* dargestellt, bei dem der Wellenleiter sichtbar wird. Punktförmige Defekte stellen Resonatoren dar, mit denen Licht räumlich „gefangen“ werden kann. Die typische Abmessung der Elementarzellen im Bereich einer halben Wellenlänge verursacht, dass keine Modellierung in Form von effektiven Materialparametern möglich ist.

1.1. Motivation und Zielsetzung

Strukturgitter aus *Metamaterialien* mit negativem Brechungsindex sind zuerst im Mikrowellenbereich hergestellt und untersucht worden. In der einschlägigen Literatur manifestiert sich ein großes Interesse an der Beschreibung von *Metamaterialien* durch effektive Materialparameter. Das Konzept der effektiven Materialparameter bietet die Möglichkeit, die Modellierung drastisch zu vereinfachen. Anstelle der numerisch aufwendigen Diskretisierung der mikroskopischen Substruktur einer Elementarzelle können für die Simulation homogene effektive Materialparameter angesetzt werden, wie z.B. in [22] gezeigt wurde. Die Elementarzellen von *Metamaterialien* sind kleiner als die Wellenlänge der elektromagnetischen Strahlung für die sie entworfen werden. Jedoch liegt der Größenunterschied mit etwa 1:10 im Bereich einer einzigen Größenordnung – oder sogar darunter. Daher wird in dieser Arbeit untersucht, inwieweit sich das Konzept der effektiven Materialparameter als tragfähig erweist und an welcher Stelle Grenzen auftreten.

Mit der seit Anbeginn der Forschung zu *Metamaterialien* fortschreitenden Miniaturisierung der technologischen Fertigungsprozesse konnten Strukturen für immer kleinere Wellenlängen hergestellt werden. Dabei sind die geometrischen Strukturen innerhalb der Elementarzellen beim Übergang vom Spektralbereich der Mikrowellen in den infraroten Bereich² des Lichts verändert worden, um auch dort effektiv negative Brechungsindizes demonstrieren zu können. Die im Mikrowellenbereich funktionierende Anordnung von *Split-Ring-Resonator* und Dipol ist zu einer Struktur modifiziert worden, die als *Fishnet* bezeichnet wird [12]. Die in den Elementarzellen verwendeten Bestandteile weisen im infraroten Spektralbereich andere Materialeigenschaften als im Mikrowellenbereich auf. Daher wird untersucht, inwieweit

²Wellenlängen von etwa 780 nm bis 1 mm bzw. Frequenzen von 300 GHz bis rund 400 THz.

Algorithmen zur Simulation beim Übergang aus dem Mikrowellenbereich in den infraroten Spektralbereich angepasst oder erweitert werden müssen.

Neben den *Metamaterialien*, die oft als räumlich unendlich periodisches Strukturgitter betrachtet werden können, spielen Nanostrukturen mit begrenzten periodischen Strukturgittern für *photonische Kristalle* eine wichtige Rolle. Ein *photonisches Kristallgitter* in zwei Dimensionen bildet die Querschnittsfläche einer *photonischen Kristallfaser*, die als Lichtwellenleiter eingesetzt wird und kommerziell erhältlich ist. Zur Berechnung der Ausbreitungseigenschaften von Licht in *photonischen Kristallfasern* stoßen bekannte Algorithmen aus dem Mikrowellenbereich an ihre Grenzen. Einerseits muss die Querschnittsfläche aufgrund der komplexen Geometrie vom Rechengitter fein aufgelöst werden und andererseits erfolgt nicht in jedem Fall eine starke Führung der ausbreitungsfähigen Wellenleitermoden, so dass das Auftreten von Strahlungsmoden in der Modellierung berücksichtigt werden muss. Darüber hinaus können die gesuchten Wellenleitermoden von den bekannten Lösungsalgorithmen nicht leicht identifiziert werden. Daher werden Erweiterungen vorgestellt, mit denen die gezielte Berechnung der gesuchten Wellenleitermoden möglich ist.

Eine weitere Klasse von Problemstellungen, für die erweiterte Lösungsalgorithmen benötigt werden, sind isolierte Nanostrukturen. Isolierte Nanostrukturen sind anders als die Elementarzelle eines Strukturgitters nicht Bestandteil einer periodischen Anordnung. Sie sind mit einzelnen Antennen im Freiraum vergleichbar, die bei ihrer Resonanzfrequenz Wellen abstrahlen können und daher beispielsweise für den Bau von Lasern geeignet sind. Zur Analyse der Resonanzen muss die zugrunde liegende numerische Modellierung durch eine sogenannte transparente Randbedingung sicherstellen, dass elektromagnetische Wellen durch die Grenzen des Rechengebiets hindurch abgestrahlt werden können. Dadurch wird wiederum die Lösung der resultierenden algebraischen Systeme aufwendiger. In dieser Arbeit werden Lösungsalgorithmen vorgestellt, mit denen gezielt gesuchte Resonanzen von Strukturen berechnet werden können, deren Modellierung die Verwendung transparenter Randbedingungen im numerischen Modell erfordert.

Das übergeordnete Ziel dieser Arbeit ist die Weiterentwicklung von Algorithmen aus dem Mikrowellenbereich für die Analyse und für den Entwurf linearer passiver Nanostrukturen für Optoelektronik und Photonik. Der Fokus liegt dabei auf den Resonanzen und Eigenwellen von Nanostrukturen.

1.2. Vorgehensweise

Die numerische Modellbildung erfolgt auf Basis der Methode der finiten Integration [38], die seit Jahrzehnten eine etablierte Theorie zur Berechnung elektromagnetischer Felder darstellt. Hauptvorteil ist dabei, dass sich neue Formulierungen, die auf Herleitungen in der kontinuierlichen Vektoranalysis basieren, leicht mit der Methode der finiten Integration in die diskrete algebraische Form bringen lassen. Als zweites Verfahren kommt die Methode der flexiblen lokalen Approximation exemplarisch für ebene Problemstellungen zum Einsatz. Gegenüber der Methode der finiten Integration bietet die Wahl von problem-angepassten Basisfunktionen eine schnellere Approximation der Feldlösungen in der Nähe runder Konturen, die insbesondere in *photonischen Kristallen* vielfach auftreten.

Zur Analyse von Elementarzellen periodischer Strukturgitter wird das Dispersionsdiagramm verwendet, das die Relation von Wellenzahl und Wellenlänge bzw. Frequenz beschreibt. Das

Dispersionsdiagramm kann dabei sowohl aus einem Ansatz gewonnen werden, der auf der Streumatrix basiert, als auch aus einem Ansatz bei dem eine Folge von dreidimensionalen Eigenwertproblemen mit periodischer Randbedingung gelöst wird.

Für die Lösung der algebraischen Eigenwertprobleme wird das JACOBI–DAVIDSON–Verfahren gelöst. Das JACOBI–DAVIDSON–Verfahren ist ein auf Unterraumprojektion basierendes iteratives Verfahren. Vorteilhaft an diesem Verfahren ist, dass Erweiterungen leicht vorgenommen werden können. Berechnete Näherungen für Eigenwerte und Eigenvektoren werden sukzessive verbessert. Der Selektionsprozess, der aus den vorhandenen approximierten Eigenwerten die gesuchten auswählt, ist herkömmlicherweise darauf beschränkt, Eigenwerte nahe bei einem Zielwert zu selektieren. Informationen, die in den Eigenvektorapproximationen zur Verfügung stehen, werden dabei nicht berücksichtigt. Daher wird das JACOBI–DAVIDSON–Verfahren mit einem erweiterten Selektionsprozess ausgestattet, das ermöglicht, gezielt Eigenvektorapproximationen für die weitere Iteration zu selektieren, die bereits a priori bekannte Eigenschaften erfüllen. Die a priori verfügbaren Informationen variieren dabei je nach Anwendungsfall. In komplexen Wellenleiterstrukturen wie der *photonischen Kristallfaser* ist beispielsweise von vornherein bekannt, dass geführte Wellenleitermoden ihr Maximum der Leistung im Kern der Faser besitzen, woraus sich ein Selektionskriterium formen lässt.

1.3. Gliederung

Nach der Einführung der Grundlagen in Kapitel 2 werden in Kapitel 3 die Formulierungen für die Berechnung spezifischer Problemstellungen eingeführt. Kapitel 4 enthält die Methoden zur numerischen Lösung, wobei ausführlich auf die Berechnung der Eigenwertprobleme eingegangen wird. Kapitel 5 enthält die Ergebnisse zu jeweils konkreten Anwendungen. Diese Arbeit gliedert sich in die fortfolgenden Kapitel:

Kapitel 2: Elektromagnetische Wellenausbreitung

Die Grundlagen zur Modellierung elektromagnetischer Wellenausbreitung werden in drei Abschnitten beschrieben. Im ersten Abschnitt liegt der Fokus auf den analytischen Methoden, die aus den MAXWELL'schen Gleichungen hergeleitet werden. Dabei wird insbesondere auf die Dispersion eingegangen. In den weiteren Abschnitten werden die Grundlagen der Methode der finiten Integration sowie die Grundzüge der Methode der flexiblen lokalen Approximation eingeführt.

Kapitel 3: Formulierungen für periodische und isolierte Nanostrukturen

Die Formulierungen für die eingangs motivierten Problemstellungen werden in diesem Kapitel hergeleitet. Im Einzelnen sind dies:

1. Der Streumatrix–Ansatz dient zur Berechnung der Dispersionseigenschaften periodischer Strukturen.
2. Die Dispersionseigenschaften periodischer Strukturen lassen sich ebenfalls aus periodisch berandeten Eigenwertprobleme berechnen, denen eine dreidimensionale Diskretisierung des Rechengebiets zugrunde liegt.
3. Die Formulierung für die Berechnung von Ausbreitungskonstanten in längshomogenen Wellenleitern mit komplexen Querschnittsgeometrien.

4. Besonderheiten in der Berechnung von Resonanzen isolierter dreidimensionaler Strukturen im Freiraum.
5. Eigenwertformulierung zur Berechnung ebener Problemstellungen mit der Methode der flexiblen lokalen Approximation.

Kapitel 4: Numerische Methoden für die Berechnung algebraischer Eigenwertprobleme

Der Lösung der zuvor motivierten algebraischen Eigenwert- und Gleichungssysteme ist dieses Kapitel gewidmet. Nach der Vorstellung bekannter Verfahren wird auf die Erweiterungen und Anpassungen der Algorithmen eingegangen.

Kapitel 5: Ergebnisse und Anwendungen

Die im vorherigen Kapitel eingeführten Algorithmen werden auf ausgewählte Nanostrukturen angewendet. Dazu gehören:

1. Die Berechnung der Dispersionseigenschaften dreidimensionaler Elementarzellen periodischer Strukturgitter mit Abmessungen im Zentimeterbereich für den Einsatz im Spektralbereich der Mikrowellen.
2. Die Berechnung der Dispersionseigenschaften dreidimensionaler Elementarzellen periodischer Strukturgitter mit Abmessungen im Nanometerbereich für den Einsatz im infraroten Spektralbereich.
3. Die gezielte Berechnung einiger ausgewählter Wellenleitermoden in geometrisch komplexen Leiterquerschnitten. Als Beispiel dient eine *photonische Kristallfaser*.
4. Die Berechnung der Resonanzfrequenzen von dielektrischen Nanoresonatoren und *Microdisks*, die sich im Freiraum befinden.
5. Berechnung der Resonanzfrequenzen eines planaren Nanoresonators mit der Methode der flexiblen lokalen Approximation.

Kapitel 6: Zusammenfassung und Ausblick

Das zusammenfassende Kapitel endet mit einem Ausblick.

Die einheitliche Notation der verwendeten Symbole und Konventionen ist ab Seite 151 aufgeführt. Die für diese Arbeit entwickelten Algorithmen sind in der Programmiersprache *Matlab* [126] und C implementiert. Für Operationen der linearen Algebra gibt es von allen großen Prozessorherstellern optimierte Bibliotheken, hier wird vor allem die *Math Kernel Library* [124] von *Intel* verwendet.

2. Elektromagnetische Wellenausbreitung

Der erste Abschnitt dieses Kapitels beschreibt die Ausbreitung elektromagnetischer Wellen mit Hilfe der kontinuierlichen MAXWELL'schen¹ Gleichungen. In den nachfolgenden Abschnitten werden mit der Methode der finiten Integration und der Methode der flexiblen lokalen Approximation diskrete Verfahren zur Berechnung elektromagnetischer Wellenausbreitung vorgestellt.

2.1. Maxwell'sche Gleichungen

Der Ausgangspunkt für die klassische Beschreibung makroskopischer Effekte elektromagnetischer Felder ist mit den Gleichungen von MAXWELL [7, 8] vollständig gegeben. Eine vollständige Darstellung der MAXWELL'schen Theorie kann den einschlägigen Lehrbüchern, wie z.B. [3] entnommen werden. Für ruhende Medien und beliebige Flächen $A \subset \mathbb{R}^3$ und Volumina $V \subseteq \mathbb{R}^3$ bei Zeitpunkten $t \in \mathbb{R}$ lauten sie in Integral- und Differenzialschreibweise [3]

$$\oint_{\partial A} \vec{E}(\vec{r}, t) \cdot d\vec{s} = -\frac{d}{dt} \int_A \vec{B}(\vec{r}, t) \cdot d\vec{A} \quad \Leftrightarrow \operatorname{rot} \vec{E}(\vec{r}, t) = -\frac{\partial \vec{B}(\vec{r}, t)}{\partial t} \quad (2.1)$$

$$\oint_{\partial A} \vec{H}(\vec{r}, t) \cdot d\vec{s} = \int_A \left(\frac{\partial \vec{D}(\vec{r}, t)}{\partial t} + \vec{J}(\vec{r}, t) \right) \cdot d\vec{A} \quad \Leftrightarrow \operatorname{rot} \vec{H}(\vec{r}, t) = \frac{\partial \vec{D}(\vec{r}, t)}{\partial t} + \vec{J}(\vec{r}, t) \quad (2.2)$$

$$\oint_{\partial V} \vec{B}(\vec{r}, t) \cdot d\vec{A} = 0 \quad \Leftrightarrow \operatorname{div} \vec{B}(\vec{r}, t) = 0 \quad (2.3)$$

$$\oint_{\partial V} \vec{D}(\vec{r}, t) \cdot d\vec{A} = \int_V \rho_L(\vec{r}, t) dV \quad \Leftrightarrow \operatorname{div} \vec{D}(\vec{r}, t) = \rho_L(\vec{r}, t) \quad (2.4)$$

mit den ortsabhängigen Größen der *elektrischen Feldstärke* $\vec{E}(\vec{r}, t)$, der *Verschiebungsstromdichte* $\vec{D}(\vec{r}, t)$, der *magnetischen Feldstärke* $\vec{H}(\vec{r}, t)$, der *magnetischen Flussdichte* $\vec{B}(\vec{r}, t)$, der *Raumladungsdichte* $\rho_L(\vec{r}, t)$ und der *Stromdichte* $\vec{J}(\vec{r}, t)$. Die MAXWELL'schen Gleichungen bestehen aus dem Induktionsgesetz von FARADAY² (2.1), dem Gesetz von AMPÈRE³ (2.2) und dem Gesetz von GAUSS⁴ (2.4). MAXWELL's Beitrag zur Vervollständigung der geschlossenen Theorie elektromagnetischer Phänomene war die Einführung des MAXWELL'schen *Verschiebungsstroms* $\vec{D}(\vec{r}, t)$. Die differenzielle Form wird jeweils durch Anwendung der Integralsätze von GAUSS und STOKES⁵ erhalten. Die Verknüpfung der Feldstärken mit den Flussdichten

¹James Clerk Maxwell (1831-1879), schottischer Physiker.

²Michael Faraday (1791-1867), englischer Naturforscher.

³André-Marie Ampère (1775-1836), französischer Physiker und Mathematiker.

⁴Johann Carl Friedrich Gauß (1777-1855), deutscher Mathematiker, Astronom, Geodät und Physiker.

⁵George Gabriel Stokes (1819-1903), irischer Mathematiker und Physiker.

erfolgt mit den ortsabhängigen tensoriellen Größen der *Permeabilität* $\bar{\mu}(\vec{r})$, der *Permittivität* $\bar{\epsilon}(\vec{r})$ und der *Konduktivität* $\bar{\kappa}(\vec{r})$

$$\vec{B}(\vec{r},t) = \bar{\mu}(\vec{r})\vec{H}(\vec{r},t) = \mu_0\vec{H}(\vec{r},t) + \vec{M}(\vec{r},t), \quad (2.5)$$

$$\vec{D}(\vec{r},t) = \bar{\epsilon}(\vec{r})\vec{E}(\vec{r},t) = \epsilon_0\vec{E}(\vec{r},t) + \vec{P}(\vec{r},t), \quad (2.6)$$

$$\vec{J}(\vec{r},t) = \vec{J}_e(\vec{r},t) + \underbrace{\bar{\kappa}(\vec{r})\vec{E}(\vec{r},t)}_{\vec{J}_i(\vec{r},t)} + \underbrace{\rho_L(\vec{r},t)\vec{v}(\vec{r},t)}_{\vec{J}_k(\vec{r},t)}, \quad (2.7)$$

die in dieser Arbeit durchgängig linear von den entsprechenden Feldstärken abhängen und als zeitlich konstant angesetzt werden. Desweiteren sind $\vec{M}(\vec{r},t)$ und $\vec{P}(\vec{r},t)$ ortsabhängige *magnetische Polarisation* und *elektrische Polarisation*. Die Stromdichte $\vec{J}(\vec{r},t)$ besteht aus den Summanden der extern eingepprägten Stromdichte $\vec{J}_e(\vec{r},t)$, der Leitungsstromdichte $\vec{J}_l(\vec{r},t)$ und der Konvektionsstromdichte $\vec{J}_k(\vec{r},t)$, die durch Ladungen hervorgerufen wird, die durch andere physikalische Kräfte bewegt werden. Die Raumladungsdichte $\rho_L(\vec{r},t)$ und die Stromdichte $\vec{J}(\vec{r},t)$ erfüllen das Gesetz der Ladungserhaltung

$$\operatorname{div} \vec{J}(\vec{r},t) + \frac{\partial \rho_L(\vec{r},t)}{\partial t} = 0, \quad (2.8)$$

das auch als Kontinuitätsgleichung bekannt ist.

2.1.1. Stetigkeit an Materialgrenzen

Der Übergang zwischen zwei unterschiedlichen benachbarten Materialien wird makroskopisch üblicherweise als sprunghafter Übergang entsprechend Abbildung 2.1 modelliert. Die Stetigkeitsbedingungen, denen elektromagnetische Felder an solchen Materialgrenzen genügen müssen, sind [3]

$$\vec{n}_{12}(\vec{r}_g) \cdot (\vec{D}_2(\vec{r}_g) - \vec{D}_1(\vec{r}_g)) = \sigma_F(\vec{r}_g) \quad (2.9)$$

$$\vec{n}_{12}(\vec{r}_g) \cdot (\vec{B}_2(\vec{r}_g) - \vec{B}_1(\vec{r}_g)) = 0 \quad (2.10)$$

$$\vec{n}_{12}(\vec{r}_g) \times (\vec{E}_2(\vec{r}_g) - \vec{E}_1(\vec{r}_g)) = \vec{0} \quad (2.11)$$

$$\vec{n}_{12}(\vec{r}_g) \times (\vec{H}_2(\vec{r}_g) - \vec{H}_1(\vec{r}_g)) = \vec{J}_F(\vec{r}_g). \quad (2.12)$$

Hierbei ist $\vec{n}_{12}(\vec{r}_g)$ der lokale Normalenvektor auf der Materialgrenzfläche zwischen Gebiet 1 und Gebiet 2 am Ort auf der Grenzfläche \vec{r}_g . Die Feldwerte sind als Grenzwerte zu interpretieren, je nachdem aus welchem Gebiet die Annäherung an einen Punkt auf der Grenzfläche erfolgt. Die Flächenstromdichte $\vec{J}_F(\vec{r}_g)$ und die Flächenladungsdichte $\sigma_F(\vec{r}_g)$ sind Größen, die nur in der Grenzfläche auftreten können.

2.1.2. Wellengleichung

Aus den MAXWELL'schen Gleichungen lässt sich die vektorielle homogene zeitabhängige Wellengleichung im Freiraum für das elektrische Feld herleiten⁶

$$\Delta \vec{E}(\vec{r},t) - \mu_0 \epsilon_0 \frac{\partial^2 \vec{E}(\vec{r},t)}{\partial t^2} = 0, \quad (2.13)$$

⁶Ein analoges Ergebnis lässt sich für das magnetische Feld formulieren.

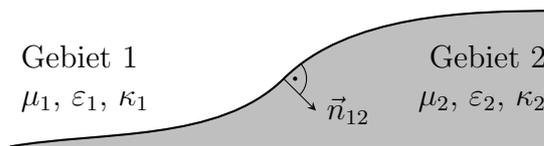


Abbildung 2.1.: Der Übergang zwischen zwei unterschiedlichen Materialien wird als Sprung der Materialparameter idealisiert.

mit der die elektromagnetische Wellenausbreitung beschrieben werden kann. Gleichung (2.13) ist eine lineare partielle Differentialgleichung zweiter Ordnung und zählt zur Klasse der hyperbolischen Differentialgleichungen. Hierbei ist der LAPLACE⁷-Operator gegeben durch

$$\Delta = \nabla^2 = \text{grad div} - \text{rot rot}. \quad (2.14)$$

Reelle Feldfunktionen mit sinusförmiger Zeitabhängigkeit⁸ können durch den Realteil komplexer Amplituden modelliert werden. Es gilt

$$E(t) = E_0 \cos(\omega t + \varphi) = \text{Re}\{\underline{E}_\omega \exp(j\omega t)\} \quad (2.15)$$

mit der komplexen Amplitude $\underline{E}_\omega = E_0 \exp(j\varphi)$. Bei Annahme harmonischer Zeitabhängigkeit entsteht aus der zeitabhängigen Wellengleichung (2.13) nach FOURIER⁹-Transformation in den Frequenzbereich die HELMHOLTZ¹⁰-Gleichung, eine partielle Differentialgleichung zweiter Ordnung aus der Klasse der elliptischen partiellen Differentialgleichungen

$$\Delta \vec{E}(\vec{r}, \omega) + k^2 \vec{E}(\vec{r}, \omega) = 0 \quad \text{mit} \quad k = \frac{\omega}{c_0} = \omega \sqrt{\mu_0 \varepsilon_0}, \quad (2.16)$$

unter der Voraussetzung von Freiraum mit den Materialparametern μ_0 und ε_0 . Die Wellenzahl k ist im Vakuum aus dem Quotient von Kreisfrequenz ω und Lichtgeschwindigkeit c_0 gegeben.

2.1.2.1. Lösungen der Helmholtz-Gleichung

Eine Lösung der HELMHOLTZ-Gleichung (2.16) in kartesischen Koordinaten ist durch ebene Wellen der Form

$$\vec{E}(\vec{r}) = E_0 e^{\pm j \vec{k} \cdot \vec{r}} \vec{e}_{\text{tan}} \quad \text{mit} \quad \vec{k} \perp \vec{e}_{\text{tan}} \quad \text{und} \quad |\vec{k}|^2 = k^2 \quad (2.17)$$

gegeben. Ihre Phasenfronten bilden Ebenen im Raum, woraus sich die Bezeichnung ableitet. Das elektrische Feld hat eine Amplitude von $|E_0|$ und ist in \vec{e}_{tan} -Richtung polarisiert. Da die Lösung im ganzen Raum \mathbb{R}^3 gilt, resultieren physikalische Lösungen erst durch die gewichtete Überlagerung verschiedener ebener Wellen.

Für die Beschreibung der zylindersymmetrischen Wellenausbreitung bieten die Zylinderwellen einen weiteren Lösungstyp der HELMHOLTZ-Gleichung (2.16). Da ein zylindrisches Koordinatensystem zugrunde liegt, sind die in der HELMHOLTZ-Gleichung verwendeten Vektordifferentialoperatoren aus Gleichung (2.14) anzupassen. Der skalare LAPLACE-Operator für E_z

⁷Pierre-Simon (Marquis de) Laplace (1749-1827), französischer Mathematiker und Astronom.

⁸In dieser Arbeit wird die in der Elektrotechnik übliche Zeitabhängigkeit $e^{j\omega t}$ verwendet. Insbesondere in der physikalischen Literatur wird häufig $e^{-i\omega t}$ angesetzt.

⁹Jean Baptiste Joseph Fourier (1768-1830), französischer Mathematiker und Physiker.

¹⁰Hermann Ludwig Ferdinand von Helmholtz (1821-1894), deutscher Physiker.

in zylindrischen Koordinaten lautet

$$\Delta = \frac{1}{\varrho} \frac{\partial}{\partial \varrho} \left(\varrho \frac{\partial}{\partial \varrho} \right) + \frac{1}{\varrho^2} \frac{\partial^2}{\partial \phi^2} + \frac{\partial^2}{\partial z^2}. \quad (2.18)$$

Lösungen der zylindrischen HELMHOLTZ-Gleichung sind die BESSEL-Funktion¹¹ J_n und HANKEL-Funktion¹² $H_n^{(1,2)}$ erster oder zweiter Art und Ordnung n .

2.1.2.2. Streuung am Zylinder im Freiraum

Da *photonische Kristalle* als Anordnung dielektrischer Zylinder in einem Hintergrundmedium angesehen werden können, wird an dieser Stelle zunächst die Streuung elektromagnetischer Wellen an einem einzelnen Zylinder betrachtet. Für das skalare Modellproblem eines in z -Richtung unendlich ausgedehnten Zylinders falle eine E_z polarisierte Welle mit Amplitude E_0 in x -Richtung ein. Die Lösung für das elektrische Feld ergibt sich als Superposition aus einfallendem Feld E_z^i und gestreutem Feld E_z^s zu $E_z = E_z^i + E_z^s$, wie in [2] hergeleitet wird. In zylindrischen Koordinaten werden die Feldanteile

$$E_z^i = E_0 e^{-jk_0 x} = E_0 e^{-jk_0 \varrho \cos \phi} = E_0 \sum_{n=-\infty}^{\infty} j^{-n} J_n(k_0 \varrho) e^{jn\phi}, \quad (2.19)$$

$$E_z^s = E_0 \sum_{n=-\infty}^{\infty} j^{-n} c_n H_n^{(2)}(k_0 \varrho) e^{jn\phi}, \quad (2.20)$$

mit der BESSEL-Funktion J_n und der HANKEL-Funktion zweiter Art $H_n^{(2)}$ in Abhängigkeit von der Freiraumwellenzahl k_0 beschrieben. Das elektrische Feld innerhalb eines nichtleitenden dielektrischen Zylinders der relativen Permittivität ε_r lässt sich als

$$E_z = E_0 \sum_{n=-\infty}^{\infty} j^n d_n J_n(k_0 \sqrt{\varepsilon_r} \varrho) e^{jn\phi} \quad (2.21)$$

angeben. Die Entwicklungskoeffizienten c_n und d_n ergeben sich jeweils aus der Erfüllung der Stetigkeitsbedingungen (2.9) bis (2.12) auf der Oberfläche des Zylinders.

2.1.3. Dispersion und Kausalität

Zur Beschreibung frequenzabhängiger Material- und Struktureigenschaften wird zunächst auf die Dispersion eingegangen, da sie für die Analyse periodischer Strukturen von zentraler Bedeutung ist. Unter Dispersion wird im allgemeinsten Fall die Abhängigkeit einer Größe der Wellengleichung von der Wellenlänge λ_{el} oder Frequenz f bzw. Kreisfrequenz ω verstanden. *Dispersionsfrei* sind Wellen, bei denen

$$c_0 = \frac{\omega}{k} = \lambda_{\text{el}} f = \text{const.} \quad \text{mit} \quad k := |\vec{k}| \quad (2.22)$$

gilt. Mathematisch entsteht die Verknüpfung von Kreiswellenzahl k und Kreisfrequenz ω als resultierende Bedingung aus dem Produktansatz zur Lösung der Wellengleichung [3]. Technisch gesehen ist die Pulsverzerrung ein Maß für die Dispersion, da im dispersionsbehafteten

¹¹Friedrich Wilhelm Bessel (1784-1846), deutscher Astronom, Mathematiker und Geodät.

¹²Hermann Hankel (1839-1873), deutscher Mathematiker.

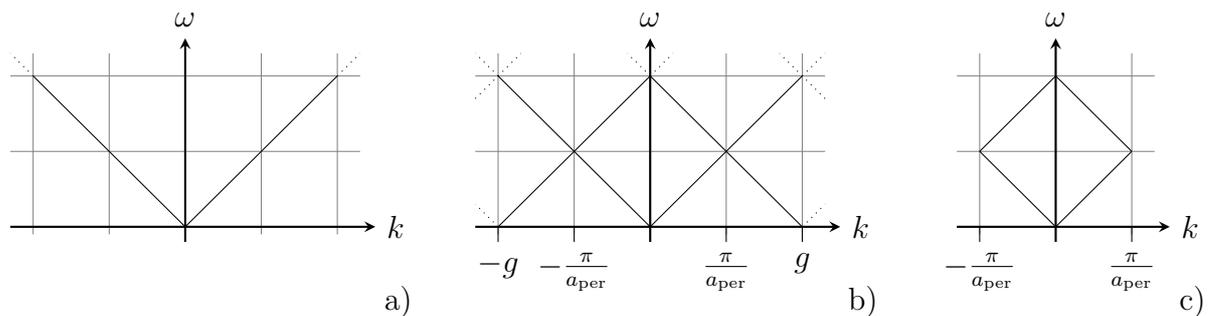


Abbildung 2.2.: a) Dispersionsdiagramm als ω - k -Diagramm für dispersionsfreie Wellenausbreitung gemäß (2.22). b) Dispersionsfreie Wellen, die periodisch bezüglich der Strukturgitterkonstanten a_{per} sind. c) Reduziertes Banddiagramm, BRILLOUIN-Zone, die alle Informationen aus b) enthält [4].

Fall Wellenpakete mit verschiedenen Frequenzen verschiedene Ausbreitungsgeschwindigkeiten erfahren. Die Definition von Phasengeschwindigkeit und Gruppengeschwindigkeit lauten

$$v_p = \frac{\omega}{k} \quad \text{und} \quad v_g = \frac{d\omega}{dk}. \quad (2.23)$$

Zur Veranschaulichung dispersiver Eigenschaften von Strukturen und Materialien wird im Dispersionsdiagramm nach Abbildung 2.2a die Kreiswellenzahl ω über der Wellenzahl k aufgetragen¹³. Rückwärtswellen oder rücklaufende Wellen zeichnen sich dadurch aus, dass für das Produkt [5, 9]

$$v_g v_p = \frac{d\omega}{dk} \frac{\omega}{k} < 0 \quad (2.24)$$

gilt. Sie treten beispielsweise in periodischen Medien auf. In den meisten Strukturgittern sind Rückwärtswellen keine fundamentalen Gitterwellen, sondern Gitterwellen höherer Ordnung. Eine wichtige Ausnahme bilden linkshändige *Metamaterial*-Strukturgitter, bei denen die fundamentale Gitterwelle eine Rückwärts-Welle ist [12].

2.1.3.1. Dispersion in periodischen Medien

Die Bezeichnung der zugrunde liegenden Mathematik von Differenzialgleichungen bzw. deren Lösungen mit translatorischer Periodizität wird je nach Anwendungsfeld mit den Namen BLOCH¹⁴ (1928, Elektronenleitung in kristallinen Festkörpern), HILL¹⁵ (1877, eindimensionale periodische Potentialfunktion), FLOQUET¹⁶ oder LJAPUNOW¹⁷ (1883 und 1892, gewöhnliche Differenzialgleichungen) assoziiert. Hier wird im Folgenden die Referenz auf BLOCH verwendet, da eine enge Verwandtschaft mit der Elektronenleitung in kristallinen Festkörpern besteht [13]. Im Folgenden werden die Grundlagen für periodische Strukturen in zwei Dimensionen wiederholt, dieser Abschnitt orientiert sich an der ausführlichen Darstellung in [4].

¹³Oder umgekehrt.

¹⁴Felix Bloch (1905-1983), schweizerisch-US-amerikanischer Physiker, 1952 Nobelpreis für Physik.

¹⁵George William Hill (1838-1914), US-amerikanischer Astronom und Mathematiker.

¹⁶Achille Marie Gaston Floquet (1847-1920), französischer Mathematiker.

¹⁷Alexander Michailowitsch Ljapunow (1857-1918), russischer Mathematiker und Physiker.

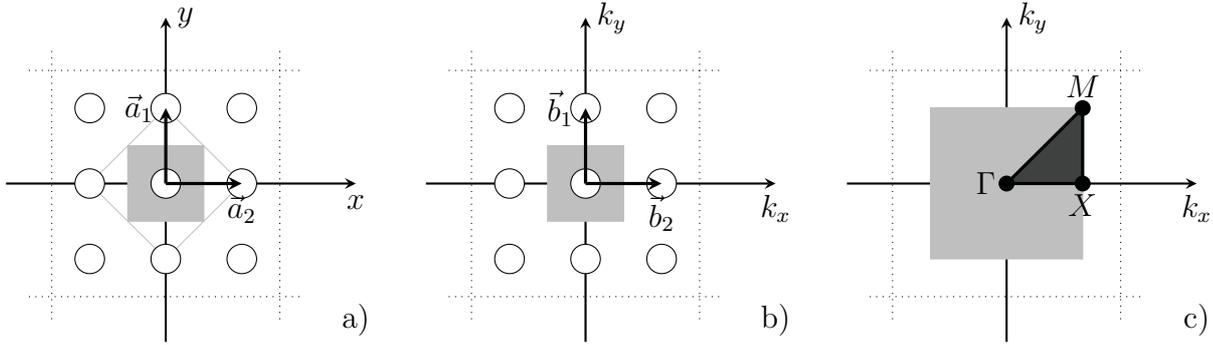


Abbildung 2.3.: Darstellung nach [4]: a) Grau schraffierte Elementarzelle eines periodischen Gitters (WIGNER–SEITZ–Zelle) im Ortsraum. b) Darstellung im reziproken Raum, grau schraffiert erste BRILLOUIN–Zone. c) Dunkelgrau schraffiert ist die irreduzible BRILLOUIN–Zone mit Mittelpunkt Γ , Seitenpunkt X und Eckpunkt M einer Elementarzelle.

Für eine unendlich ausgedehnte zweidimensional–periodische Materialverteilung $\varepsilon(x,y)$ nach Abbildung 2.3a soll die Lösung der Wellengleichung (2.16) für eine Welle mit E_z –Komponente bestimmt werden. Dazu wird ein Ansatz mit dem noch unbekanntem BLOCH–Wellenvektor \vec{K}_B gemacht

$$E_z(\vec{r}) = E_{\text{per}}(\vec{r})e^{-j\vec{K}_B \cdot \vec{r}} \quad \text{mit } \vec{r} = (x,y). \quad (2.25)$$

Dann lässt sich zeigen, dass die Funktion $E_{\text{per}}(\vec{r})$ der Periodizität

$$E_{\text{per}}(\vec{r}) = E_{\text{per}}(\vec{r} + \vec{R}) \quad (2.26)$$

unterliegt, was als BLOCH–Bedingung bezeichnet wird. Dabei ist \vec{R} der Translationsvektor, der für ein zweidimensionales Strukturgitter als

$$\vec{R} = m_1\vec{a}_1 + m_2\vec{a}_2 \quad \text{mit } \vec{a}_1 = (a_{\text{per}},0), \vec{a}_2 = (0,a_{\text{per}}), m_1,m_2 \in \mathbb{Z} \quad (2.27)$$

und der Strukturgitterkonstante $a_{\text{per}} \in \mathbb{R}^+$ notiert wird. Daraus folgt, dass die Funktion $E_{\text{per}}(\vec{r})$ selbst gitterperiodisch ist. Der Faktor $e^{-j\vec{K}_B \cdot \vec{r}}$ gibt den Phasenvorschub der Amplitude E_z über einer Elementarzelle in x –Richtung an.

Für den akademischen Fall ebener Wellen, die bezüglich der räumlichen Strukturgitterkonstante a_{per} die Periodizität der BLOCH–Bedingung erfüllen, folgt nach Abbildung 2.2b eine Kurvenschar $k \rightarrow k \pm mg \forall m \in \mathbb{N}$ mit $g = \frac{2\pi}{a_{\text{per}}}$. Für $|k| = \frac{\pi}{a_{\text{per}}}$ überlagern sich Wellen mit gegenläufiger Ausbreitung, so dass sich eine stehende Welle bildet und sich die Kurven im Dispersionsdiagramm schneiden. Das durch die Periodizität unendlich fortgesetzte Diagramm in Abbildung 2.2b lässt sich durch Reduktion auf die erste BRILLOUIN¹⁸–Zone kompakt als Abbildung 2.2c darstellen.

Zur geometrischen Konstruktion der ersten BRILLOUIN–Zone wird zunächst ein kartesisches¹⁹ Punktgitter (BRAVAIS²⁰–Gitter) in zwei Dimensionen betrachtet, das durch translatorische Periodizität der Punkte entsprechend (2.27) und Gitterkonstante $a_{\text{per}} \in \mathbb{R}^+$ gegeben ist. Die

¹⁸Léon Nicolas Brillouin (1889–1969), französisch-amerikanischer Physiker.

¹⁹Nach René Descartes (1596–1650), französischer Philosoph, Mathematiker und Naturwissenschaftler.

²⁰Auguste Bravais (1811–1863), französischer Physiker.

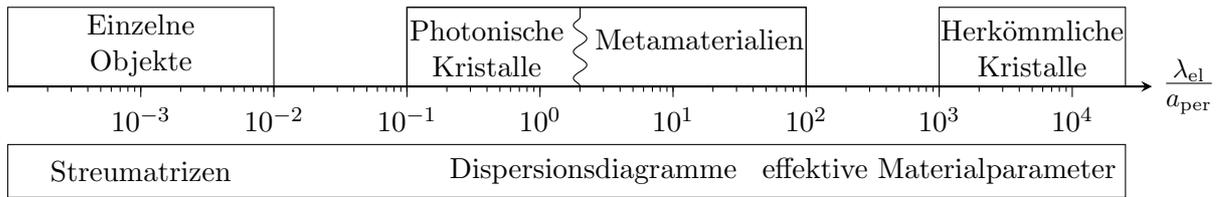


Abbildung 2.4.: Skala des Quotienten Wellenlänge pro Strukturgitterkonstante a_{per} mit Zuordnung typischer Strukturen und Methoden zur Charakterisierung.

Elementarzelle des Gitters, die nur einen Gitterpunkt in ihrem Zentrum enthält, wird WIGNER²¹–SEITZ²²–Zelle genannt (grau schraffiert in Abbildung 2.3a). Alle Orte innerhalb der WIGNER–SEITZ–Zelle liegen dem zugeordneten zentralen Gitterpunkt näher als den benachbarten Gitterpunkten. Das reziproke Gitter ist über die Menge jener Wellenvektoren definiert für die $\vec{K} \cdot \vec{R} = n2\pi$ mit $n \in \mathbb{Z}$ gilt. Entsprechend Abbildung 2.3b wird analog zur WIGNER–SEITZ–Zelle im BRAVAIS–Gitter im reziproken Gitter die erste BRILLOUIN–Zone konstruiert. Hierbei ist $\vec{K} = n_1\vec{b}_1 + n_2\vec{b}_2$ mit $\vec{b}_1 = (\frac{2\pi}{a_{\text{per}}}, 0)$, $\vec{b}_2 = (0, \frac{2\pi}{a_{\text{per}}})$, $n_1, n_2 \in \mathbb{Z}$ und Gitterkonstante $a_{\text{per}} \in \mathbb{R}^+$ gegeben. Durch Symmetriebetrachtung schrumpft die BRILLOUIN–Zone zur irreduziblen BRILLOUIN–Zone entsprechend Abbildung 2.3c. Konventionell wird der Ursprung mit Γ , der Seitenpunkt mit X und der Eckpunkt der irreduziblen BRILLOUIN–Zone mit M bezeichnet.

Mit diesen Überlegungen ist es möglich, das Dispersionsdiagramm eines unendlich ausgedehnten Strukturgitters durch die Analyse einer einzigen Elementarzelle zu gewinnen. Dabei muss die Periodizität der Feldlösung durch das numerische Modell gewährleistet werden.

2.1.3.2. Effektive Materialparameter periodischer Strukturgitter

Die Beschreibung eines periodischen Strukturgitters durch effektive Materialparameter bietet den Vorteil, dass die Substruktur des Strukturgitters nicht mehr detailliert beschrieben werden muss, was die Modellbildung erheblich vereinfacht. Dennoch müssen die effektiven Materialparameter das elektromagnetische Verhalten des Strukturgitters in jedem Fall korrekt modellieren. Abbildung 2.4 veranschaulicht die Größenverhältnisse zwischen Wellenlänge λ_{el} und Strukturgitterkonstante a_{per} . Im Fall kleiner Wellenlängen und in Relation dazu großer räumlicher Periodizität, lassen sich Problemstellungen mit *makroskopischen* Ansätzen wie der Strahlenoptik beschreiben. Im Fall großer Wellenlängen, die auf Strukturen sehr kleiner, *mikroskopischer*, räumlicher Periode treffen, ist eine Beschreibung des periodischen Mediums mit effektiven Materialparametern möglich. Dies gilt insbesondere für atomare Kristallgitter in einem großen Frequenzbereich, da die Strukturgitterkonstante kleiner als 1 nm ist²³. Der Bereich, in dem Wellenlängen und räumliche Strukturgitterkonstante in etwa gleich groß sind, wird als *mesoskopischer* Bereich bezeichnet, in dem Verfahren für effektive Materialparameter das elektromagnetische Verhalten nicht in jedem Fall hinreichend genau beschreiben, da die resultierende Gitterwelle aufgrund der relativ großen Substruktur sich nicht gleichförmig ausbreitet. Die Möglichkeit der Homogenisierung von *Metamaterial*–Strukturgittern, die nach Abbildung 2.4 zum *mesoskopischen* Bereich gehören, wurde in [22] untersucht.

²¹Eugene Paul Wigner (1902-1995), ungarisch-amerikanischer Physiker, 1963 Nobelpreis für Physik.

²²Frederick Seitz (1911-2008), US-amerikanischer Physiker.

²³Natriumchlorid hat beispielsweise eine Strukturgitterkonstante von $a_{\text{per}} = 0.562$ nm.

2.1.3.3. Kausalität

Der Prozess der Berechnung von effektiven Materialparametern für periodische Strukturgitter wird als Homogenisierung bezeichnet. Die Homogenisierung führt auf effektive Materialparameter, die das makroskopische elektromagnetische Verhalten modellieren und für ein bestimmtes Frequenzband gültig sind. Die berechneten effektiven Materialparameter müssen einem kausalen²⁴ Dispersionsmodell folgen, da sie das Verhalten eines physikalischen Systems beschreiben. Die KRAMERS²⁵–KRONIG²⁶–Relationen setzen Realteil und Imaginärteil der Funktionen von Dispersionsmodellen in Beziehung und sind ein Spezialfall der HILBERT²⁷–Transformation. Für ein kausales Dispersionsmodell für die komplexe Permittivität $\varepsilon_c = \varepsilon' - j\varepsilon''$ gilt [10, Seite 27][6]

$$\varepsilon'(\omega) = \varepsilon_\infty + \frac{2}{\pi} \text{CH} \int_0^\infty \frac{\omega' \varepsilon''(\omega')}{\omega'^2 - \omega^2} d\omega', \quad (2.28)$$

$$\varepsilon''(\omega) = -\frac{2}{\pi} \text{CH} \int_0^\infty \frac{\varepsilon'(\omega') - \varepsilon_\infty}{\omega'^2 - \omega^2} d\omega'. \quad (2.29)$$

CH bezeichnet hierbei den CAUCHY'schen²⁸ Hauptwert des auftretenden Integrals. Die Folgen dieses Ergebnisses sind weitreichend, da der frequenzabhängige Real- und Imaginärteil der Permittivität zum einen direkt in einander überführt werden können und zum anderen Verluste in Form von $\varepsilon'' > 0$ immer vorhanden sind. Die theoretische Betrachtung der Feldenergie führt auf die Relationen [6]

$$\frac{d(\omega\varepsilon'(\omega))}{d\omega} > 0, \quad \frac{d(\omega\mu'(\omega))}{d\omega} > 0. \quad (2.30)$$

Im Folgenden werden die benötigten Dispersionsmodelle vorgestellt, welche die KRAMERS–KRONIG–Relationen erfüllen.

2.1.3.4. Dispersion durch frequenzabhängige Materialeigenschaften

Physikalisch betrachtet entsteht die Dispersion in Materie aus der frequenzabhängigen makroskopischen Polarisation \vec{P} mit

$$\vec{P} = N\vec{p}, \quad (2.31)$$

wobei $\vec{p} = q\vec{r}$ die mikroskopischen Dipolmomente sind, die aus zwei Ladungen mit unterschiedlichen Vorzeichen vom Betrag q durch \vec{r} voneinander getrennt sind und N deren Dichte pro Volumeneinheit beschreibt. Die frequenzabhängige makroskopische Polarisation \vec{P} kann für viele Materialien relativ gut mit Modellen beschrieben werden, die sich polynomial in Abhängigkeit der Frequenz ω verhalten [3, Kapitel 7.5].

²⁴Die Impulsantwort eines kausalen verschiebungsinvarianten linearen Systems verschwindet für $t < 0$, da die Antwort des Systems nicht vor der Einspeisung des Signals beginnen kann.

²⁵Hendrik Anthony Kramers (1894-1952), niederländischer Physiker.

²⁶Ralph Kronig (1904-1995), deutsch-US-amerikanischer Physiker.

²⁷David Hilbert (1862-1943), deutscher Mathematiker.

²⁸Augustin Louis Cauchy (1789-1857), französischer Mathematiker.

Für das Modell von LORENTZ²⁹ wird angenommen, dass z.B. die Elektronen eines Materials als klassische, mechanische, harmonische Oszillatoren beschrieben werden können, die von einer externen Kraft – dem zeitabhängigen elektrischen Feld $\vec{E}(t)$ – ausgelenkt werden. Mit dieser Quelle ergibt sich die zeitabhängige Bewegungsgleichung für die Auslenkung \vec{r} eines einzelnen Elektrons zu

$$\frac{\partial^2 \vec{r}}{\partial t^2} + \alpha_t \frac{\partial \vec{r}}{\partial t} + \omega_0^2 \vec{r} = -\frac{q_e}{m_e} \vec{E}(t). \quad (2.32)$$

Hierbei ist α_t der Dämpfungsfaktor, ω_0 die Resonanzfrequenz, q_e die Elektronenladung und m_e die Elektronenmasse. Unter der Annahme einer bei der Frequenz ω zeitharmonisch schwingenden Anregung $E(t)$ ist die Lösung von (2.32) im Frequenzbereich

$$\vec{r}(\omega) = -\frac{1}{\omega_0^2 + j\omega\alpha_t - \omega^2} \frac{q_e}{m_e} \vec{E}(\omega). \quad (2.33)$$

Mit dem Dipolmoment eines Elektrons $\vec{p}(\omega) = q_e \vec{r}(\omega)$ und (2.31) lässt sich die dielektrische Verschiebungsstromdichte für ein homogenes Material entsprechend (2.6) als

$$\vec{D}(\omega) = \varepsilon_\infty \vec{E}(\omega) + \vec{P}(\omega) \quad (2.34)$$

schreiben. Die Polarisation kann in die frequenzabhängige Permittivität

$$\varepsilon(\omega) = \varepsilon_0 \varepsilon_\infty + \frac{Nq_e^2/m_e}{\omega_0^2 + j\omega\alpha_t - \omega^2} \quad (2.35)$$

eingeschlossen werden, was zum klassischen LORENTZ-Modell der dielektrischen Dispersion mit einer Polstelle führt

$$\frac{\varepsilon(\omega)}{\varepsilon_0} = \varepsilon_\infty + \frac{\omega_p^2}{\omega_0^2 + j\omega\alpha_t - \omega^2} \quad \text{mit} \quad \omega_p^2 = \frac{Nq_e^2}{\varepsilon_0 m_e}, \quad (2.36)$$

wobei ω_p die Plasmafrequenz ist und die Dispersionskurve für sehr hohe Frequenzen gegen den Wert ε_∞ konvergiert.

Ein Spezialfall des Modells von LORENTZ entsteht, wenn in den vorangegangenen Gleichungen (2.32) bis (2.36) die Resonanzfrequenz $\omega_0 = 0$ gesetzt wird. Das resultierende Modell von DRUDE³⁰ lautet

$$\frac{\varepsilon(\omega)}{\varepsilon_0} = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2 + j\omega\alpha_t} \quad \text{mit} \quad \omega_p^2 = \frac{Nq_e^2}{\varepsilon_0 m_e} \quad (2.37)$$

und dient zur Beschreibung von Metallen bei optischen Frequenzen.

Reichen (2.37) oder (2.36) allein nicht aus, um die breitbandige Frequenzabhängigkeit eines Materials zu interpolieren, können beispielsweise mehrere LORENTZ'sche Modelle mit unterschiedlichen Parametern $\omega_{L,i}^2, \omega_{0,i}$ kombiniert werden

$$\frac{\varepsilon(\omega)}{\varepsilon_0} = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2 + j\omega\alpha_t} + \sum_i \frac{\omega_{L,i}^2}{\omega_{0,i}^2 + j\omega\alpha_{t,i} - \omega^2}. \quad (2.38)$$

²⁹Hendrik Antoon Lorentz (1853-1928), niederländischer Mathematiker und Physiker.

³⁰Paul Karl Ludwig Drude (1863-1906), deutscher Physiker.

2.1.4. Energieerhaltung

Neben der Kausalität ist die Energieerhaltung eine weitere wichtige Eigenschaft physikalischer Systeme, die jegliche Modellierung einschließen sollte. Für Medien mit konstanten Materialparametern $\kappa, \varepsilon, \mu \in \mathbb{R}^+$ gibt der Satz von POYNTING³¹ die Energiebilanz elektromagnetischer Wellenausbreitung an. Im quellenfreien Raum lautet er in differenzieller komplexer Schreibweise

$$-\operatorname{div} \vec{S} = \bar{p}_v + j2\omega(\bar{w}_m - \bar{w}_e) \quad (2.39)$$

mit dem komplexen POYNTING'schen Vektor

$$\vec{S}(\vec{r}, \omega) = \frac{1}{2} \vec{E}(\vec{r}, \omega) \times \vec{H}(\vec{r}, \omega)^* \quad (2.40)$$

als zeitlich gemittelte Leistungsflussdichte und der zeitlich gemittelten elektrischen Verlustleistungsdichte

$$\bar{p}_v = \frac{\kappa}{2} \vec{E}(\vec{r}, \omega) \cdot \vec{E}(\vec{r}, \omega)^* \quad (2.41)$$

sowie der zeitlich gemittelten magnetischen und elektrischen Energiedichte

$$\bar{w}_m = \frac{\mu}{4} \vec{H}(\vec{r}, \omega) \cdot \vec{H}(\vec{r}, \omega)^*, \quad \bar{w}_e = \frac{\varepsilon}{4} \vec{E}(\vec{r}, \omega) \cdot \vec{E}(\vec{r}, \omega)^*. \quad (2.42)$$

OHM'sche³² Verluste hervorgerufen durch endliche Leitfähigkeiten $\kappa > 0$ verursachen dabei die Umwandlung von elektromagnetischer Feldenergie in Wärme. Durch die konstant vorausgesetzten Materialparameter können die endlichen Leitfähigkeiten als Imaginärteil einer komplexen Permittivität

$$\varepsilon_c = \varepsilon' - j\varepsilon'' \quad \text{mit} \quad \varepsilon'' = \kappa/\omega\varepsilon_0 \quad (2.43)$$

dargestellt werden³³. Die gesamte Energiebilanz, welche die Wärmeenergieanteile mit einschließt, bleibt dabei ausgeglichen, während die elektromagnetische Feldenergie bei der Existenz von OHM'schen Verlusten abnimmt [3, Kapitel 6.7].

2.1.4.1. Der Poynting'sche Satz für dispersive Medien

Etwas allgemeiner muss die Energiebilanz für isotrope dispersive Medien betrachtet werden, da die auftretenden Verluste in den Materialparametern frequenzabhängig beispielsweise nach den Modellen (2.36) und (2.37) sein können. Dazu wird zunächst der reelle zeitabhängige POYNTING-Vektor

$$\vec{S}(\vec{r}, t) = \vec{E}(\vec{r}, t) \times \vec{H}(\vec{r}, t) \quad (2.44)$$

eingeführt. Desweiteren wird die Voraussetzung verwendet, dass die elektrischen und magnetischen Feldstärken als $\vec{E} = \vec{E}_1(t) \cos(\omega_0 t + \varphi_E)$ bzw. $\vec{H} = \vec{H}_1(t) \cos(\omega_0 t + \varphi_H)$ darstellbar sind

³¹John Henry Poynting (1852-1914), englischer Physiker.

³²Georg Simon Ohm (1789-1854), deutscher Physiker.

³³Zeitabhängigkeit wie in (2.15) festgelegt.

und die zeitlichen Veränderungen von $\vec{E}_1(t)$ und $\vec{H}_1(t)$ klein gegenüber der reziproken Trägerfrequenz $1/\omega_0$ sind. Dann folgt für die effektive elektromagnetische Energiedichte gemäß [3, Kapitel 6.8] und [6, Kapitel 80]

$$w_{\text{eff,disp}}(\vec{r}) = \frac{1}{2} \operatorname{Re} \left\{ \frac{d(\omega\varepsilon)}{d\omega}(\omega_0) \right\} \left\langle \vec{E}(\vec{r},t) \cdot \vec{E}(\vec{r},t) \right\rangle + \frac{1}{2} \operatorname{Re} \left\{ \frac{d(\omega\mu)}{d\omega}(\omega_0) \right\} \left\langle \vec{H}(\vec{r},t) \cdot \vec{H}(\vec{r},t) \right\rangle. \quad (2.45)$$

Der Erwartungswert $\langle \cdot \rangle$ stellt dabei die Mittelung über eine Periode der Trägerfrequenz dar. Für dispersionsfreie Materialien vereinfachen sich die Ableitungen und es verbleiben lediglich die Materialparameter, so dass das reelle Pendant zu (2.42) mit

$$w_{\text{eff}}(\vec{r}) = \frac{1}{2} \left(\varepsilon \left\langle \vec{E}(\vec{r},t) \cdot \vec{E}(\vec{r},t) \right\rangle + \mu \left\langle \vec{H}(\vec{r},t) \cdot \vec{H}(\vec{r},t) \right\rangle \right) \quad (2.46)$$

resultiert. Der POYNTING'sche Satz in seiner reellen zeitabhängigen Darstellung lautet nun mit \vec{S} aus (2.44)

$$\begin{aligned} \frac{\partial}{\partial t} w_{\text{eff}}(\vec{r}) &= -\operatorname{div} \vec{S}(\vec{r},t) - \vec{J}(\vec{r},t) \cdot \vec{E}(\vec{r},t) \\ &\quad - \omega_0 \operatorname{Im} \{ \varepsilon(\omega_0) \} \left\langle \vec{E}(\vec{r},t) \cdot \vec{E}(\vec{r},t) \right\rangle - \omega_0 \operatorname{Im} \{ \mu(\omega_0) \} \left\langle \vec{H}(\vec{r},t) \cdot \vec{H}(\vec{r},t) \right\rangle. \end{aligned} \quad (2.47)$$

Im Term $\vec{J}(\vec{r},t) \cdot \vec{E}(\vec{r},t)$ sind analog zur Verlustleistungsdichte (2.41) in (2.39) die OHM'schen Verluste enthalten, während die beiden frequenzabhängigen Terme in der unteren Zeile die Absorptionsverluste im Material darstellen.

2.1.4.2. Gütefaktor

Der *Gütefaktor* Q (von engl. quality factor) ist sowohl ein Maß für die Breitbandigkeit einer Resonanz als auch ein Maß für Verluste. Der Gütefaktor eines Resonators ist definiert als Quotient von der im zeitlichen Mittelwert gespeicherten Energie und der Energie, die pro Sekunde dissipiert [3, Kapitel 8.8]

$$Q := \text{Kreisfrequenz} \times \frac{\text{zeitlich mittlere gespeicherte Energie}}{\text{Energieverlust pro Sekunde}}. \quad (2.48)$$

Da die Amplitude des z.B. elektrischen Feldes der Resonanzfrequenz ω_i über der Zeit abklingt, gilt $\vec{E} \propto e^{j\omega_i t - \alpha_{t,i} t}$ mit $\omega_i, \alpha_{t,i} \in \mathbb{R}^+$, so dass sich (2.48) in

$$Q_i = \frac{\omega_i}{2\alpha_{t,i}} \quad (2.49)$$

für die i -te Resonanz überführen lässt [1, Kapitel 5.10][11]. Der Exponent $\alpha_{t,i}$ ist hierbei die zeitliche Dämpfungskonstante. Der Gütefaktor dient desweiteren als Maß für das Ansprechverhalten eines Resonators, da die OHM'schen und dielektrischen Verluste, die im idealen Fall diskrete Resonanzfrequenz zu einem Frequenzband „verschmieren“, in welchem die Resonanz mehr oder minder stark durch einen breitbandigen Puls angeregt werden kann. Daher kann der Gütefaktor Q ebenso aus einer breitbandigen Resonanzkurve mit einem Maximum bei f_0 bzw. $\lambda_{\text{el},0}$ und der Halbleistungsbandbreite Δf bzw. der Linienbreite $\Delta\lambda_{\text{el}}$ gewonnen werden

$$Q = \frac{f_0}{\Delta f} = \frac{\Delta\lambda_{\text{el}}}{\lambda_{\text{el},0}}. \quad (2.50)$$

Voraussetzung dafür ist, dass die Resonanz entsprechend scharf – also Q hinreichend groß ist – und mit benachbarten Resonanzen eine vernachlässigbare Überlappung besteht [3, Kapitel 8.8].

2.1.5. Orthogonalität und Streumatrizen

In diesem Abschnitt wird gezeigt wie die Orthogonalität von Wellenleitermoden, die Eigenlösungen der MAXWELL'schen Gleichungen darstellen, für die Definition von Streumatrizen verwendet wird.

2.1.5.1. Orthogonalität von Wellenleitermoden

Bei geführter Wellenausbreitung entlang $\pm z$ -gerichteter längshomogener Wellenleiter haben die Feldamplituden eine Abhängigkeit vom Ort z und der Wellenzahl k_z der Form

$$\vec{E}, \vec{H} \propto e^{\mp j k_z z} \quad (2.51)$$

und lassen sich in die longitudinalen Komponenten $(\vec{E}, \vec{H})_z$ und die transversalen Komponenten $(\vec{E}, \vec{H})_{\text{tan}}$ zerlegen

$$\vec{E} = \vec{E}_{\text{tan}} + \vec{E}_z, \quad \vec{H} = \vec{H}_{\text{tan}} + \vec{H}_z. \quad (2.52)$$

Die spezifischen Lösungen des Eigenwertproblems der Wellendifferenzialgleichung³⁴ für die Querschnittsebene eines Wellenleiters werden als Wellenleitermoden bezeichnet. Die Feldlösung für den n -ten Wellenleitermode mit Ausbreitung in $+z$ -Richtung

$$\vec{E}_n(x, y, z) = a_n \vec{\mathcal{E}}_n(x, y) e^{-j k_{z,n} z}, \quad (2.53)$$

$$\vec{H}_n(x, y, z) = a_n \vec{\mathcal{H}}_n(x, y) e^{-j k_{z,n} z} \quad (2.54)$$

setzt sich aus der Wellenamplitude a_n , den transversalen Wellenleitermodenfunktion $\vec{\mathcal{E}}_n(x, y)$ und $\vec{\mathcal{H}}_n(x, y)$ sowie der komplexen Wellenzahl

$$k_{z,n} := \beta_{z,n} - j \alpha_{z,n} \quad \text{mit } \alpha_{z,n}, \beta_{z,n} \in \mathbb{R}_0^+ \quad (2.55)$$

zusammen. Hierbei wird $\beta_{z,n}$ als Ausbreitungs- und $\alpha_{z,n}$ als räumliche Dämpfungskonstante³⁵ bezeichnet. Für gedämpfte Wellenleitermoden gilt $\alpha_{z,n} > 0$, für evaneszente Wellenleitermoden $\alpha_{z,n} > 0$ und $\beta_{z,n} = 0$. Die allgemeine Lösung des Wellenleiterquerschnittsproblems der MAXWELL'schen Gleichungen lässt sich durch Superposition aller existierenden Wellenleitermoden schreiben

$$\vec{E}(x, y, z) = \sum_{n \in \mathbb{N}} \vec{\mathcal{E}}_n(x, y) (a_n e^{-j k_{z,n} z} + b_n e^{j k_{z,n} z}), \quad (2.56)$$

$$\vec{H}(x, y, z) = \sum_{n \in \mathbb{N}} \vec{\mathcal{H}}_n(x, y) (a_n e^{-j k_{z,n} z} - b_n e^{j k_{z,n} z}). \quad (2.57)$$

Hierbei ist b_n die Amplitude des n -ten Wellenleitermodes bei Ausbreitung in negative z -Richtung. Wellenleitermoden unterschiedlicher Ordnung n, m sind zueinander orthogonal bezüglich des Skalarprodukts [1, S.336]. Mit geeigneter Normierung gilt

$$\int_A (\vec{E}_n \times \vec{H}_m) \cdot \vec{e}_z dA = \begin{cases} 1, & \text{wenn } n = m, \\ 0, & \text{sonst,} \end{cases} \quad (2.58)$$

³⁴Ausführlichere Darstellung z.B. in [1, 3]

³⁵Im Gegensatz zur zeitlichen Dämpfungskonstante α_t aus Gleichung (2.49).

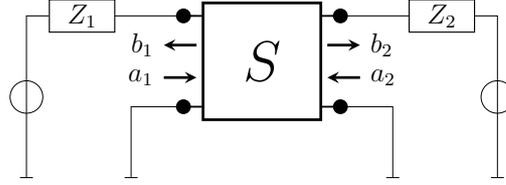


Abbildung 2.5.: Streumatrix eines Zweitors: Richtungsdefinitionen der Wellenamplituden.

wobei sich die Integration über die Querschnittsfläche A an einer beliebigen Position $z = z_0$ erstreckt. Für verlustlose Wellenleiter gilt darüber hinaus auch die Definition des Skalarproduktes über den POYNTING'schen Vektor entsprechend (2.40) [1, S.337] und passender Normierung

$$\int_A \left(\vec{E}_n \times \vec{H}_m^* \right) \cdot \vec{e}_z dA = \begin{cases} 1, & \text{wenn } n = m, \\ 0, & \text{sonst.} \end{cases} \quad (2.59)$$

2.1.5.2. Streumatrix fundamentaler Wellenleitermoden

Wellenleiter lassen sich mit Hilfe der Streumatrix als lineares Zweitor modellieren [1]. Die Streumatrizen kommen ursprünglich aus der Zweitorthorie und dienen entsprechend Abbildung 2.5 allgemein der Charakterisierung von Hochfrequenzeigenschaften. Sie beschreiben das Verhalten eines linearen Zweitors vollständig und verknüpfen einfallende und auslaufende Wellenamplituden. Bei der Herleitung der Streumatrizen werden die auslaufenden Wellenamplituden b_i als abhängige Variablen definiert und die einfallenden Wellenamplituden a_i als unabhängige Variablen.

Liegt die Referenzebene in (2.56) bei $z = 0$, so ergeben sich für die komplexe Amplituden der modalen Spannungen \mathcal{U}_n und Ströme \mathcal{I}_n am Tor n bei symmetrischer Normierung auf die Leitungsimpedanz $Z_{L,n}$

$$\mathcal{U}_n(\omega) = \sqrt{Z_{L,n}}(a_n(\omega) + b_n(\omega)), \quad (2.60)$$

$$\mathcal{I}_n(\omega) = \frac{1}{\sqrt{Z_{L,n}}}(a_n(\omega) - b_n(\omega)). \quad (2.61)$$

Die S -Parameter bilden die skalaren Einträge der Streumatrix und sind als Quotient von auslaufenden Wellen b_n von Tor n und einfallenden Wellen a_m an Tor m definiert, wobei einzig am Tor m eine einfallende Welle existiert

$$s_{nm}(\omega) = \left. \frac{b_n(\omega)}{a_m(\omega)} \right|_{a_l=0 \ \forall l \neq m}. \quad (2.62)$$

Für Zweitore folgt in Matrixform

$$\begin{pmatrix} b_1(\omega) \\ b_2(\omega) \end{pmatrix} = \begin{pmatrix} s_{11}(\omega) & s_{12}(\omega) \\ s_{21}(\omega) & s_{22}(\omega) \end{pmatrix} \begin{pmatrix} a_1(\omega) \\ a_2(\omega) \end{pmatrix}. \quad (2.63)$$

2.1.5.3. Streumatrix für Wellenleitermoden höherer Ordnung

Im vorherigen Abschnitt 2.1.5.2 wird die 2×2 -Streumatrix für einen einzigen fundamentalen Wellenleitermode eingeführt, jedoch lassen sich Wellenleitermoden höherer Ordnung sowie

die Kopplung von Wellenleitermoden untereinander ebenfalls mit Streumatrizen beschreiben. Mit den Blockmatrizen in der Form \mathcal{S}_{nm} gilt

$$\begin{pmatrix} \mathbf{b}_1(\omega) \\ \mathbf{b}_2(\omega) \end{pmatrix} = \begin{pmatrix} \mathcal{S}_{11}(\omega) & \mathcal{S}_{12}(\omega) \\ \mathcal{S}_{21}(\omega) & \mathcal{S}_{22}(\omega) \end{pmatrix} \begin{pmatrix} \mathbf{a}_1(\omega) \\ \mathbf{a}_2(\omega) \end{pmatrix}. \quad (2.64)$$

Dabei enthalten \mathbf{a}_n (\mathbf{b}_n) alle einfallenden (auslaufenden) Wellenamplituden von i verschiedenen Wellenleitermoden am Tor n und es gilt

$$\mathbf{a}_n(\omega) = (a_{n,1}(\omega), a_{n,2}(\omega), \dots, a_{n,i}(\omega))^T, \quad \mathbf{b}_n(\omega) = (b_{n,1}(\omega), b_{n,2}(\omega), \dots, b_{n,i}(\omega))^T. \quad (2.65)$$

2.1.5.4. Transfer–Streumatrix

Durch algebraische Umformung von (2.64) wird die Transfer–Streumatrix \mathcal{T} hergeleitet

$$\begin{pmatrix} \mathbf{b}_1(\omega) \\ \mathbf{a}_1(\omega) \end{pmatrix} = \begin{pmatrix} \mathcal{T}_{11}(\omega) & \mathcal{T}_{12}(\omega) \\ \mathcal{T}_{21}(\omega) & \mathcal{T}_{22}(\omega) \end{pmatrix} \begin{pmatrix} \mathbf{a}_2(\omega) \\ \mathbf{b}_2(\omega) \end{pmatrix}, \quad (2.66)$$

bei der beide Typen von Wellenamplituden jeweils pro Tor als Vektor zusammengefasst werden. Für die Umrechnung von \mathcal{S} nach \mathcal{T} wird die Relation

$$\begin{pmatrix} \mathcal{T}_{11} & \mathcal{T}_{12} \\ \mathcal{T}_{21} & \mathcal{T}_{22} \end{pmatrix} = \begin{pmatrix} \mathcal{S}_{12} - \mathcal{S}_{11}\mathcal{S}_{21}^{-1}\mathcal{S}_{22} & \mathcal{S}_{11}\mathcal{S}_{21}^{-1} \\ -\mathcal{S}_{21}^{-1}\mathcal{S}_{22} & \mathcal{S}_{21}^{-1} \end{pmatrix} \quad (2.67)$$

verwendet, wobei die Frequenzabhängigkeit nicht ausgeschrieben ist.

2.2. Methode der finiten Integration (FIT)

Die direkte Lösung der kontinuierlichen MAXWELL'schen Gleichungen aus 2.1 ist nur für elektromagnetische Aufgabenstellungen praktikabel, die aus relativ einfachen Geometrien bestehen. Für die meisten Anwendungsfälle ist es daher nur möglich, diskrete Formulierungen der MAXWELL'schen Gleichungen numerisch zu lösen. Im Rahmen dieser Arbeit wird die Methode der finiten Integration (FIT, von engl. finite integration technique) verwendet, die seit über 30 Jahren von WEILAND propagiert und seit dem weiterentwickelt wurde [38, 39]. Eine kommerzielle Implementierung existiert mit dem Software-Paket CST STUDIO SUITE [123]. Das Ziel ist die Berechnung der elektromagnetischen Feldverteilung im untersuchten Rechengebiet, aus der sekundäre Größen wie Energie- oder Leistungsgrößen gewonnen werden können. Im Folgenden wird die prinzipielle Methode skizziert. Die folgende Darstellung orientiert sich an [16, 21, 22, 25, 26, 28].

2.2.1. Räumlich diskrete Gitter–Maxwell–Gleichungen

Der Vorgang der Zuordnung der physikalischen Größen vom Kontinuum auf einen diskreten Unterraum wird als Diskretisierung bezeichnet. Dazu wird ein Ausschnitt aus \mathbb{R}^3 – namentlich das Rechengebiet Ω – in eine endliche abzählbare Menge disjunkter zusammenhängender Teilgebiete $\{\Omega_i\}$ unterteilt

$$\bigcup_{i \in I_V \subset \mathbb{N}} \Omega_i \subset \mathbb{R}^3, \quad (2.68)$$

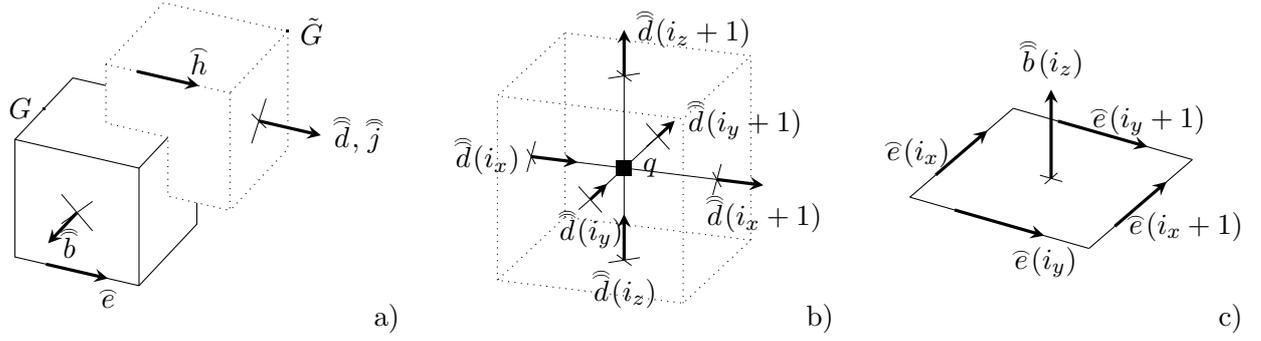


Abbildung 2.6.: a) Räumlicher Versatz von primärem Rechengitter G und dualem Rechengitter \tilde{G} . b) Die integralen Zustandsgrößen sind flächenallokierte Gitterflüsse, kantenallokierte Gitterspannungen sowie volumenallokierte Ladungen. c) Lokaler Umlauf der elektrischen Gitterspannungen.

die im Folgenden als Rechengitterzellen G bezeichnet werden. Hierbei ist die Indexmenge $I_v = \{1, \dots, n_p\}$ eine endliche Teilmenge aus \mathbb{N} , mit der Anzahl an Teilvolumina n_p . Prinzipiell ist die Geometrie der Rechengitterzellen in gewissen Grenzen beliebig, im Rahmen dieser Arbeit werden jedoch ausschließlich kartesische Rechengitter verwendet, die aus quaderförmigen Rechengitterzellen bestehen. Daraus wird sofort ersichtlich, dass beliebige Geometrien mit dem diskreten kartesischen Rechengitter im Allgemeinen nur noch näherungsweise korrekt dargestellt werden können. Für praktische Anwendungen hat sich die Indizierung der Rechengitterzellen mit dem kanonischen Index [28] bewährt, da dies zu effizient implementierbaren Algorithmen führt, wobei sich insbesondere die entstehenden Matrizen mit Bandstruktur effizient auf Rechnersystemen verarbeiten lassen. Das kanonische Indizierungsschema für xyz -Koordinaten eines kartesischen Rechengebiets mit

$$n_p = n_x n_y n_z \quad (2.69)$$

Punkten, die durch das Tripel (i_x, i_y, i_z) adressiert sind, ergibt sich zu

$$i = 1 + (i_x - 1) + (i_y - 1)n_x + (i_z - 1)n_x n_y \quad i_\zeta \in \{1, \dots, n_\zeta\} \wedge \zeta \in \{x, y, z\}. \quad (2.70)$$

In FIT werden direkt die kontinuierlichen MAXWELL'schen Gleichungen in Integralform (2.1) bis (2.4) diskretisiert. Dazu werden sie auf das dual-orthogonale Rechengitterpaar $\{G, \tilde{G}\}$ ³⁶ übertragen, welches das endliche Rechengebiet aus n_p Punkten im \mathbb{R}^3 darstellt. Aus der Diskretisierung resultieren die integralen Zustandsgrößen der elektrischen Gitterspannung \vec{e} und des elektrischen Gitterflusses \vec{d} , die als Integrale über primäre Kanten L_i bzw. duale Flächen \tilde{A}_i definiert sind. Ganz analog resultieren die integralen Zustandsgrößen der magnetischen Gitterspannung \vec{h} und des magnetischen Gitterflusses \vec{b} , die als Integrale über duale Kanten \tilde{L}_j bzw. primäre Flächen A_j definiert sind. Die Allokation der integralen Zustandsgrößen erfolgt nach Abbildung 2.6 auf den entsprechenden Kanten und Flächen zu

³⁶Primäres Rechengitter G und dazu duales Rechengitter \tilde{G} .

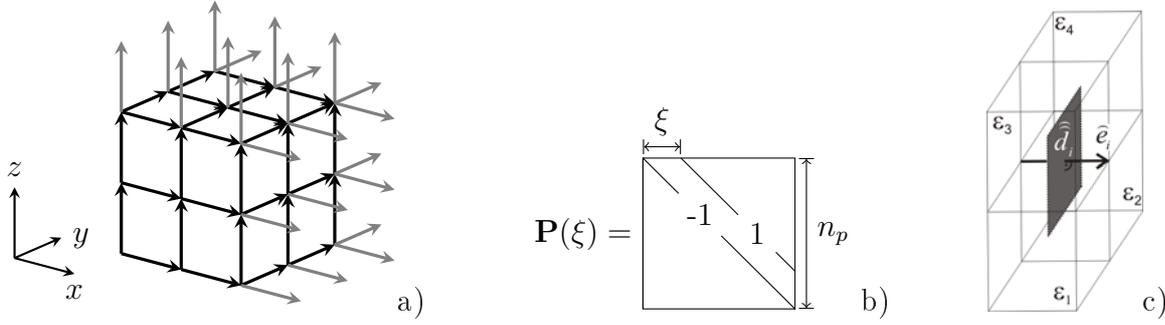


Abbildung 2.7.: a) Entartete Kanten (grau), die aufgrund der gewählten Indizierung außerhalb des Rechengitters liegen und nicht berücksichtigt werden. b) Aufbau des \mathbf{P} -Operators mit Hauptdiagonale und ξ -ter Nebendiagonale. c) Allokation der elektrischen Gitterspannung \hat{e}_i und des elektrischen Gitterflusses \hat{d}_i , sowie die daraus motivierte Mittelung der Permittivität $\bar{\epsilon}_i$.

$$\hat{e}_i(t) = \int_{L_i} \vec{E}(\vec{r}, t) \cdot d\vec{s}, \quad (2.71) \quad \hat{h}_j(t) = \int_{L_j} \vec{H}(\vec{r}, t) \cdot d\vec{s}, \quad (2.74)$$

$$\hat{d}_i(t) = \int_{\tilde{A}_i} \vec{D}(\vec{r}, t) \cdot d\vec{A}, \quad (2.72) \quad \hat{b}_j(t) = \int_{A_j} \vec{B}(\vec{r}, t) \cdot d\vec{A}, \quad (2.75)$$

$$\hat{j}_i(t) = \int_{\tilde{V}_i} \vec{J}(\vec{r}, t) \cdot d\vec{V}, \quad (2.73) \quad q_i(t) = \int_{\tilde{V}_i} \rho_L(\vec{r}, t) \cdot dV. \quad (2.76)$$

Die Allokation des elektrischen Gitterstroms $\hat{\mathbf{j}}$ erfolgt analog zum elektrischen Gitterfluss $\hat{\mathbf{d}}$. Die Ladung \mathbf{q} ist den primären Rechengitterknoten zugeordnet. Die gewählte Indizierung verursacht entsprechend Abbildung 2.7a, dass primäre Kanten über den Rand des Rechengebiets hinausragen (grau dargestellt). Diese Kanten werden als entartete Kanten bezeichnet. Matrixeinträge, die zu entarteten Kanten gehören, werden zu Null gesetzt, um die Bandstruktur der Matrizen nicht zu stören. Im kartesischen Rechengitter mit n_p primären Knoten wird die Indizierung der Kanten und Flächen derart gewählt, dass x -, y - und z -Komponenten Teilvektoren bilden, die zu Vektoren der Länge $3n_p$ entsprechend

$$\hat{\mathbf{e}} = \begin{pmatrix} \hat{e}_x \\ \hat{e}_y \\ \hat{e}_z \end{pmatrix}, \quad \hat{\mathbf{h}} = \begin{pmatrix} \hat{h}_x \\ \hat{h}_y \\ \hat{h}_z \end{pmatrix}, \quad \hat{\mathbf{d}} = \begin{pmatrix} \hat{d}_x \\ \hat{d}_y \\ \hat{d}_z \end{pmatrix}, \quad \hat{\mathbf{b}} = \begin{pmatrix} \hat{b}_x \\ \hat{b}_y \\ \hat{b}_z \end{pmatrix} \quad (2.77)$$

zusammengefasst werden. Während die vektoriellen Größen $3n_p$ Komponenten besitzen, lässt sich das skalare Feld der Ladung \mathbf{q} in einem Vektor der Länge n_p speichern. Die Gitter-Maxwell-Gleichungen (MGE, engl. Maxwell grid equations) ergeben sich zu

$$\mathbf{C}\hat{\mathbf{e}} = -\frac{d}{dt}\hat{\mathbf{b}}, \quad (2.78) \quad \tilde{\mathbf{C}}\hat{\mathbf{h}} = \frac{d}{dt}\hat{\mathbf{d}} + \hat{\mathbf{j}}, \quad (2.80)$$

$$\mathbf{S}\hat{\mathbf{b}} = 0, \quad (2.79) \quad \tilde{\mathbf{S}}\hat{\mathbf{d}} = \mathbf{q}. \quad (2.81)$$

Die Matrizen \mathbf{C} , $\tilde{\mathbf{C}}$ sowie \mathbf{S} , $\tilde{\mathbf{S}}$ sind algebraische Pendanten der topologischen Rotations- und Divergenzoperatoren (\mathbf{C} von engl. curl, Rotation; \mathbf{S} von engl. source, Quelle, Divergenz) und bestehen aus den diskreten partiellen Differenzierungsoperatoren $\mathbf{P}_x, \mathbf{P}_y, \mathbf{P}_z \in \{-1; 0; 1\}^{n_p \times n_p}$ deren struktureller Aufbau in Abbildung 2.7b angegeben ist. Für ein kartesisches Rechengitter

gilt bezogen auf Abbildung 2.7b $\mathbf{P}_x = \mathbf{P}(1)$, $\mathbf{P}_y = \mathbf{P}(n_x)$ und $\mathbf{P}_z = \mathbf{P}(n_x n_y)$, wobei Einträge zu entarteten Kanten entsprechend Abbildung 2.7a zu Null gesetzt werden. Daraus ergeben sich die Matrizen \mathbf{C} , $\tilde{\mathbf{C}}$ und \mathbf{S} , $\tilde{\mathbf{S}}$ zu

$$\mathbf{C} = \begin{pmatrix} \mathbf{0} & -\mathbf{P}_z & \mathbf{P}_y \\ \mathbf{P}_z & \mathbf{0} & -\mathbf{P}_x \\ -\mathbf{P}_y & \mathbf{P}_x & \mathbf{0} \end{pmatrix}, \quad \tilde{\mathbf{C}} = \mathbf{C}^T, \quad (2.82)$$

$$\mathbf{S} = (\mathbf{P}_x \ \mathbf{P}_y \ \mathbf{P}_z), \quad \tilde{\mathbf{S}} = (-\mathbf{P}_x^T \ -\mathbf{P}_y^T \ -\mathbf{P}_z^T). \quad (2.83)$$

Die diskreten Materialbeziehungen verknüpfen die kantenallokierten Gitterspannungen mit den flächenallokierten Gitterflüssen und somit die Gleichungspaare (2.78), (2.79) und (2.80), (2.81). Wird pro primärer Rechengitterzelle eine homogene Materialverteilung angenommen, können Materialverteilungen mit beliebiger Randkontur nur noch approximiert werden, da sich die Materialverteilung an den Grenzflächen benachbarter Rechengitterzellen sprunghaft ändert. Jedoch ist mit den Stetigkeitsbedingungen aus Gleichungen (2.9) bis (2.12) gegeben, dass die tangentialen elektrischen und magnetischen Feldstärken einen stetigen Verlauf über Materialgrenzen hinweg besitzen. Die dual-orthogonale Allokation der Gitterspannungen und Gitterflüsse gewährleistet daher eine eindeutige Definition. In Abbildung 2.7c schneidet die schraffierte duale Gitterfläche \tilde{A}_i die vier benachbarten primären Rechengitterzellen, denen homogene Materialparameter zugewiesen sind. Da der Gitterfluss \tilde{d}_i auf der dualen Fläche \tilde{A}_i allokiert ist und die Gitterspannung \tilde{e}_i senkrecht auf \tilde{A}_i steht, ist die Berechnung von \tilde{d}_i aus \tilde{e}_i nur über einen Mittelwert der beteiligten Permittivitäten möglich

$$\tilde{e}_i = \frac{\sum_{k=1}^4 \varepsilon_k \tilde{A}_k}{\sum_{k=1}^4 \tilde{A}_k}. \quad (2.84)$$

Neben der Verknüpfung von Gitterspannungen mit Gitterflüssen leisten die Materialmatrizen die Verknüpfung von primärem Rechengitter G und dualem Rechengitter \tilde{G} . Die Materialbeziehungen lauten analog zu Gleichungen (2.6) bis (2.7)

$$\hat{\mathbf{d}} = \mathbf{M}_\varepsilon \hat{\mathbf{e}}, \quad (2.85) \quad \hat{\mathbf{b}} = \mathbf{M}_\mu \hat{\mathbf{h}}, \quad (2.86) \quad \hat{\mathbf{j}} = \mathbf{M}_\kappa \hat{\mathbf{e}} \quad (2.87)$$

und bilden die räumliche Materialverteilung im Rechengebiet ab. Für dual-orthogonale Rechengitterpaare $\{G, \tilde{G}\}$ sind die Materialmatrizen diagonal und entsprechend leicht invertierbar. Für nichtorthogonale Diskretisierungsschemata [26] ist dies nicht der Fall. Die Gleichungen (2.85) bis (2.87) lassen sich komponentenweise schreiben als

$$\hat{d}_i = \mathbf{M}_{\varepsilon,ii} \hat{e}_i = \frac{\int_{\tilde{A}_i} \vec{D}(\vec{r},t) \cdot d\vec{A}}{\int_{L_i} \vec{E}(\vec{r},t) \cdot d\vec{s}} \hat{e}_i = \frac{\bar{\varepsilon} \tilde{A}_i}{L_i} \hat{e}_i + \mathcal{O}(L_i^\zeta), \quad (2.88)$$

$$\hat{h}_j = \mathbf{M}_{\mu^{-1},jj} \hat{b}_j = \frac{\int_{\tilde{L}_j} \vec{H}(\vec{r},t) \cdot d\vec{s}}{\int_{A_j} \vec{B}(\vec{r},t) \cdot d\vec{A}} \hat{b}_j = \frac{\tilde{L}_j}{\bar{\mu} A_j} \hat{b}_j + \mathcal{O}(\tilde{L}_j^\zeta), \quad (2.89)$$

$$\hat{j}_i = \mathbf{M}_{\kappa,ii} \hat{e}_i = \frac{\int_{\tilde{A}_i} \vec{J}(\vec{r},t) \cdot d\vec{A}}{\int_{L_i} \vec{E}(\vec{r},t) \cdot d\vec{s}} \hat{e}_i = \frac{\bar{\kappa} \tilde{A}_i}{L_i} \hat{e}_i + \mathcal{O}(L_i^\zeta), \quad (2.90)$$

wobei $\bar{\varepsilon}$ und $\bar{\kappa}$ flächengemittelte und $\bar{\mu}$ längengemittelte Größen sind. Für die Fehlerordnung gilt nach [26, Anhang A.1] $\zeta \geq 2$. Sie wird für äquidistante Rechengitter und homogene Materialverteilungen im Rechengebiet maximal. Die Erfüllung von wichtigen physikalischen wie auch vektoranalytischen Gesetzen bleibt hierbei erhalten. Im Einzelnen sind das

1. Quellenfreiheit von Wirbelfeldern ($\operatorname{div} \operatorname{rot} \equiv 0$),
2. Wirbelfreiheit von Gradientenfeldern ($\operatorname{rot} \operatorname{grad} \equiv 0$),
3. Ladungserhaltung [25] entsprechend (2.8),
4. Energieerhaltung [25] entsprechend (2.39),
5. Orthogonalität der Moden [40] entsprechend (2.58).

2.2.2. Zeitbereich

Aus Gleichungen (2.78) und (2.80) folgt im verlustlosen Fall für die Darstellung im Zeitbereich

$$\frac{d}{dt} \begin{pmatrix} \widehat{\mathbf{b}}(t) \\ \widehat{\mathbf{e}}(t) \end{pmatrix} = \begin{pmatrix} \mathbf{0} & -\mathbf{C} \\ \mathbf{M}_\varepsilon^{-1} \widetilde{\mathbf{C}} \mathbf{M}_\mu^{-1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \widehat{\mathbf{b}}(t) \\ \widehat{\mathbf{e}}(t) \end{pmatrix} + \begin{pmatrix} \mathbf{0} \\ -\mathbf{M}_\varepsilon^{-1} \widehat{\mathbf{j}}_e(t) \end{pmatrix}. \quad (2.91)$$

Mit dem zentralen Differenzenquotienten als Diskretisierung der Zeitableitung ergibt sich der Leapfrogalgorithmus für die MGE. Das entstehende Verfahren stimmt aus Implementierungssicht mit der Methode der finiten Differenzen im Zeitbereich (FDTD, von engl. finite differences time-domain) von YEE [43] überein.

Um frequenzabhängige Größen und insbesondere Streumatrizen zu erhalten, werden die Ergebnisse mit Hilfe der diskreten FOURIER-Transformation (DFT) in den Frequenzbereich transformiert. Als Anregungsfunktion bietet sich in vielen Fällen der modulierte GAUSS-Puls an, da er ein nahezu begrenztes Frequenzband anregt und seine FOURIER-Transformierte wiederum gaußförmig ist. Als Abbruchkriterium eignet sich je nach Problemstellung eine Begrenzung der Anzahl der zu simulierenden Pulsweiten, wie auch die im Rechengebiet enthaltene Energie, die unter einen angebbaren Wert abgeklungen sein muss, bevor die Simulation abgebrochen werden kann.

2.2.3. Frequenzbereich

Bei harmonischer Zeitabhängigkeit $e^{j\omega t}$ führt die Darstellung der Gitterzustandsgrößen nach (2.15) zur Frequenzbereichsdarstellung. Die Zeitableitung der Gitterzustandsgrößen geht in eine Multiplikation der komplexen Amplitude mit dem Faktor $j\omega$ über. Die diskrete Wellengleichung der FIT ist die Curl-Curl-Gleichung mit Anregungsterm $\widehat{\mathbf{j}}_s$. Im verlustlosen Fall ohne Leitungsströme lautet sie

$$(\widetilde{\mathbf{C}} \mathbf{M}_\mu^{-1} \mathbf{C} - \omega^2 \mathbf{M}_\varepsilon) \widehat{\mathbf{e}} = -j\omega \widehat{\mathbf{j}}_s. \quad (2.92)$$

Existiert mit $\widehat{\mathbf{j}}_s = 0$ keine Anregung in Gleichung (2.92), so geht die verlustlose diskrete Wellengleichung in die Eigenwertgleichung

$$\mathbf{M}_\varepsilon^{-1} \widetilde{\mathbf{C}} \mathbf{M}_\mu^{-1} \mathbf{C} \widehat{\mathbf{e}} = \omega^2 \widehat{\mathbf{e}} \quad (2.93)$$

für die elektrische Gitterspannung über. Bei Gleichung (2.93) handelt es sich um ein algebraisches Standard-Eigenwertproblem der Form $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$. Die Eigenwerte λ sind die quadrierten Resonanzfrequenzen ω^2 und die Feldlösungen entsprechen den Eigenvektoren $\widehat{\mathbf{e}}$. Auf die FIT-Eigenwertgleichung wird in Kapitel 3 ausführlich eingegangen.

2.2.4. Randbedingungen

Bisher wurde noch keine Aussage über das Verhalten an den Rändern des diskreten Rechengebiets getroffen. Die Modellierung der Berandung des finiten Rechengebiets hat jedoch maßgeblichen Einfluss auf die Qualität der resultierenden Feldlösung, da die Berandung die zu Grunde liegende Problemstellung nicht modifizieren darf und gleichzeitig das Rechengebiet so klein wie möglich halten soll, um den numerischen Aufwand begrenzt zu halten.

2.2.4.1. Konservative Randbedingungen

Konservative Randbedingungen erlauben keinen Leistungsfluss durch den Rand und im einfachsten Fall wird das Verschwinden einer tangentialen Gitterspannung gefordert. Interpretiert werden kann das im Fall der tangentialen elektrischen Gitterspannung $\widehat{\mathbf{e}}_{||}$ als Grenzfläche zu einer perfekt elektrischen Leitfähigkeit (PEC, von engl. perfect electric conductor) mit $\kappa \rightarrow \infty$ und im Fall der tangentialen magnetischen Gitterspannung $\widehat{\mathbf{h}}_{||}$ als Grenzfläche zu einer perfekt magnetischen Leitfähigkeit (PMC, von engl. perfect magnetic conductor) mit $\text{Im}\{\mu_c\} \rightarrow -\infty$ [28].

Am perfekt elektrisch leitenden Rand gilt neben $\widehat{\mathbf{e}}_{||} = 0$ für den zum Rand normalen magnetischen Gitterfluss $\widehat{\mathbf{b}}_{\perp} = 0$. In FIT kann der PEC Rand leicht dadurch implementiert werden, dass im Rotationsoperator \mathbf{C} die entsprechenden Einträge von $\widehat{\mathbf{e}}$ und $\widehat{\mathbf{b}}$ zu Null gesetzt werden. Am PMC Rand gilt für die tangential magnetische Gitterspannung $\widehat{\mathbf{h}}_{||} = 0$ für den zum Rand normalen elektrischen Gitterfluss $\widehat{\mathbf{d}}_{\perp} = 0$. Dieser Rand resultiert aus der virtuellen Verdopplung der halben tangential orientierten dualen Integrationsfläche \widehat{A}_i im AMPÈRE'schen Gesetz (2.80) [28].

2.2.4.2. Periodische Randbedingung

Die periodische Randbedingung wird verwendet, um Strukturgitter unendlicher Periodizität nach Abschnitt 2.1.3.1 zu modellieren. Sie setzt die tangentialen elektrischen Gitterspannungen zweier gegenüberliegender Randflächen in eine feste Phasenbeziehung $\widehat{\mathbf{e}}_{||,1} = \widehat{\mathbf{e}}_{||,2}e^{j\varphi}$ und wurde bereits 1986 implementiert [41]. Die Anzahl der Freiheitsgrade wird hierbei um die Unbekannten einer Randfläche reduziert, da diese durch die Phasenbeziehung a posteriori berechnet werden können. Aufgrund dieses Abhängigkeitsprinzips werden äquivalente Formulierungen in der Finite-Elemente-Methode als Master-Slave-Randbedingungen bezeichnet. Analog zur Darstellung in [21, S.62] kann ein Operator \mathbf{L}_{φ} eingeführt werden, der die tangentialen elektrischen Gitterspannungen am abhängigen Rand eliminiert. Die Menge der kanonischen Indices nach Gleichung (2.70) der Tangentialkomponenten der elektrischen Gitterspannung am abhängigen Rand (uv -Ebene bei w_{\max}) im kartesischen Rechengebiet lautet für die u -Komponenten

$$I_u = \{n_u n_v (n_w - 1) + 1, n_u n_v (n_w - 1) + 2, \dots, n_u n_v n_w\} \quad (2.94)$$

und für die v -Komponenten

$$I_v = \{n_p + n_u n_v (n_w - 1) + 1, n_p + n_u n_v (n_w - 1) + 2, \dots, n_p + n_u n_v n_w\}. \quad (2.95)$$

Der \mathbf{L}_{φ} -Operator ersetzt die Komponenten am abhängigen Rand w_{\max}

$$\widehat{\mathbf{e}} = \mathbf{L}_{\varphi} \widehat{\mathbf{e}}_{\text{red}} \quad (2.96)$$

mit dem Produkt der tangentialen Gitterspannungen am unabhängigen Rand w_{\min} und dem Phasenfaktor $e^{j\varphi}$. Demnach ist

$$\mathbf{L}_\varphi \in \mathbb{C}^{3n_p \times (3n_p - |I_u \cup I_v|)} \quad (2.97)$$

mit der Mächtigkeit $|I_{\hat{\mathbf{e}}, \text{tan}}|$ der Menge $I_{\hat{\mathbf{e}}, \text{tan}}$. Die Einträge von \mathbf{L}_φ lauten

$$\mathbf{L}_\varphi(i, l) = \begin{cases} e^{j\varphi} \delta_{i, i - n_u n_v (n_w - 1)} & \forall i \in I_u \\ e^{j\varphi} \delta_{i, i - n_u n_v (n_w - 1)} & \forall i \in I_v, \\ \delta_{i, l} & \text{sonst} \end{cases} \quad (2.98)$$

mit dem KRONECKER³⁷-Delta $\delta_{i, l}$. Hierbei sind Elemente, die aufgrund der Indizierung adressiert werden können, aber außerhalb des Rechengebiets liegen, nicht gesondert berücksichtigt und müssen wie in Abbildung 2.7a zu Null gesetzt werden. An den Rändern des Rechengitters haben die dualen Zellen nur halbe Volumina, weshalb die halbierten Volumina am unabhängigen Rand verdoppelt werden müssen, um die Materialverteilung am abhängigen Rand akkurat modellieren zu können. Voraussetzung ist allerdings, dass das (halbe) duale Zellvolumen an abhängigem wie unabhängigem Rand gleich ist. Die resultierende Materialmatrix für die Permittivität wird als $\mathbf{M}_{\varepsilon, \text{per}}$ bezeichnet.

2.2.4.3. Wellenleiterrand

Zur Simulation einer geführten Wellenausbreitung wird der sogenannte Wellenleiterrand verwendet, der es ermöglicht, Wellenpakete ein- und auszukoppeln und so den Anschluss einer Zuleitung an das Rechengebiet darstellt. Darüber hinaus dient das Konzept des Wellenleiterrands zur Aufzeichnung der Wellenamplituden (2.53) bis (2.57), die für die Berechnung der Streumatrix (2.63) aus Abschnitt 2.1.5.2 benötigt werden. Wie für die Orthogonalität der Wellenleitermoden im kontinuierlichen Raum aus Abschnitt 2.1.5.1 entsprechend Gleichungen (2.58) und (2.59) vorausgesetzt, lässt sich ein beliebiges diskretes Feldbild $\hat{\mathbf{e}}_{\text{tan}}(\omega)$ in eine Reihe orthogonaler Moden $\hat{\mathbf{e}}_i$

$$\hat{\mathbf{e}}_{\text{tan}}(\omega) = \sum_{i=0}^n c_i(\omega) \hat{\mathbf{e}}_i \quad (2.99)$$

entwickeln³⁸. Die Entwicklungskoeffizienten $c_i(\omega)$ ergeben sich in einem kartesischen Rechengebiet mit Ausbreitungsrichtung z mit dem diskreten Analogon des Skalarprodukts (2.58) zu

$$c_i(\omega) = \hat{\mathbf{e}}_x \cdot \hat{\mathbf{h}}_{y, i} - \hat{\mathbf{e}}_y \cdot \hat{\mathbf{h}}_{x, i}. \quad (2.100)$$

Da die Extraktion von Streuparametern in dieser Arbeit unverändert durchgeführt wird, sei auf die Literatur zur Behandlung transversal homogener in [18, 32] und breitbandig inhomogener Wellenleiter in [20] verwiesen. Von Interesse ist an dieser Stelle die Einführung des diskreten Eigenwertproblems für die Berechnung der diskreten Wellenleitermoden $\hat{\mathbf{e}}_i$ wie in [24, 26, 40] gezeigt wurde. Das diskrete Eigenwertproblem für den Wellenleiterrand wird in Abschnitt 3.3 ausführlich behandelt.

³⁷Leopold Kronecker (1823-1891), deutscher Mathematiker.

³⁸Bei der Diskretisierung einer Fläche mit n_p Gitterpunkten stehen maximal $2n_p$ Wellenleitermoden für die Reihenentwicklung zur Verfügung.

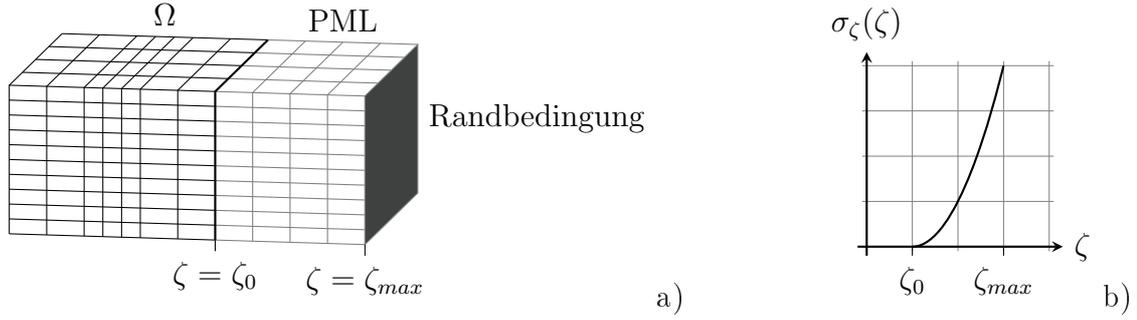


Abbildung 2.8.: a) Ursprüngliches Rechengebiet Ω , das um einige Schichten PML ab $\zeta = \zeta_0$ erweitert wird. b) Polynomiales Dämpfungsprofil $\sigma_\zeta(\zeta)$. Durch $\sigma_\zeta(\zeta_0) = 0$ wird gewährleistet, dass die reelle Ortskoordinate stetig in die komplexe Ortskoordinate übergeht.

2.2.4.4. Transparente Randbedingungen

Transparente Randbedingungen für Rechengebiete müssen eine Vielzahl von Anforderungen erfüllen. Der Rand soll nicht nur durchlässig für propagierende und evaneszente Wellen sein, sondern Durchlässigkeit für nahezu beliebige Frequenzen und Einfallswinkel bieten. Während in [23] noch eine echte (ebene) Randbedingung vorgestellt wurde, hat sich seit [14] eine Modellierung durchgesetzt, bei der das ursprüngliche Rechengebiet im Volumen um mehrere Lagen eines impedanzangepassten leitfähigen zusätzlichen Rechengebiets erweitert wird (siehe Abbildung 2.8a), in dem die Welle näherungsweise komplett absorbiert wird. Da im kontinuierlichen Modell die Impedanzanpassung perfekt ist, hat sich die Bezeichnung *perfectly matched layers* (PML) durchgesetzt. Eine FIT-Implementierung basierend auf der Idee aus [44] ist in [20] beschrieben. Im Folgenden wird die Vorgehensweise aus [31] beschrieben, die unabhängig vom Koordinatensystem immer möglich ist. Analytisch betrachtet wird die reelle Ortskoordinate ζ am Rand $\zeta = \zeta_0$ des Rechengebiets in die komplexe Ebene entsprechend der analytischen Funktion

$$\zeta \rightarrow \tilde{\zeta} = \int_{\zeta_0}^{\zeta} s_\zeta(\zeta') d\zeta' \quad \text{für } \zeta > \zeta_0 \quad \text{und} \quad \zeta \in \{x, y, z\} \quad (2.101)$$

fortgesetzt. Dabei geht ζ bei ζ_0 stetig in die komplexe Ortskoordinate $\tilde{\zeta}$ über, wobei s_ζ der neue, in der PML geltende, Streckungsfaktor ist. Für die Dämpfung propagierender Wellen wird für den Streckungsfaktor $s_\zeta(\zeta)$

$$s_\zeta(\zeta) = 1 + \frac{\sigma_\zeta(\zeta)}{j\omega} \quad \text{für } \sigma_\zeta \in \mathbb{R}^+, \quad (2.102)$$

gefordert, wobei $\sigma_\zeta(\zeta)$ ein näher zu spezifizierendes Dämpfungsprofil ist (siehe unten). Der Imaginärteil der Ortskoordinate innerhalb der PML verursacht das exponentielle Abklingen der Welle. In Gleichung (2.102) ist der ursprüngliche reelle Anteil des Streckungsfaktors als erster Summand erkennbar, während der zweite Summand mit dem Imaginärteil die Erweiterung in die komplexe Ebene darstellt. Für quasistatische Problemstellungen hat sich die Wahl des Streckungsfaktors zu

$$s_\zeta(\zeta) = 1 + \frac{\sigma_\zeta(\zeta)}{\alpha_{\text{PML}} + j\omega} \quad \text{für } \alpha_{\text{PML}} \in \mathbb{R}^+ \quad (2.103)$$

bewährt [31], da durch den zusätzlichen Koeffizienten α_{PML} auch für $\omega \rightarrow 0$ die Dämpfungseigenschaft der PML erhalten bleibt. Das Abklingen originär evaneszenter Wellenanteile lässt sich mit der Forderung an den Streckungsfaktor

$$\text{Re}\{s_\zeta(\zeta)\} > 1 \quad (2.104)$$

innerhalb der PML beschleunigen, wofür ggf. ein weiterer Freiheitsgrad eingeführt wird [161]. Für die oben eingeführten Dämpfungsprofile $\sigma_\zeta(\zeta)$ wird üblicherweise ein Polynom vom Grad $m \in \mathbb{N}$ mit

$$\sigma_\zeta(\zeta) = \sigma_{\text{max}} \left(\frac{\zeta - \zeta_0}{\zeta_{\text{max}} - \zeta_0} \right)^m \quad (2.105)$$

gewählt. So wird gewährleistet, dass am Übergang vom Rechengebiet in die PML $\sigma_\zeta(\zeta_0) = 0$ und am äußeren Rand der PML (also dem neuen erweiterten Rechengebietsrand) $\sigma_\zeta(\zeta_{\text{max}}) = \sigma_{\text{max}}$ gilt, wie in Abbildung 2.8b veranschaulicht ist.

Im kartesischen Fall lässt sich nach Gleichung (2.101) die Abbildung des Ortsvektors \vec{r} auf den komplexen Ortsvektor $\tilde{\vec{r}}$ als

$$\vec{r} \rightarrow \tilde{\vec{r}} = \bar{\bar{\mathbf{G}}} \cdot \vec{r} \quad \text{mit} \quad \bar{\bar{\mathbf{G}}} = \begin{bmatrix} \tilde{x}/x & 0 & 0 \\ 0 & \tilde{y}/y & 0 \\ 0 & 0 & \tilde{z}/z \end{bmatrix} \quad (2.106)$$

schreiben. Geometrisch betrachtet lässt sich die Fortsetzung der reellen Ortskoordinate in der komplexen Ebene als Metriktransformation auffassen [31]. Der Metriktensor des EUKLID'schen³⁹ Raums

$$\bar{\bar{\mathbf{G}}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.107)$$

eines Rechengebiets Ω erfährt an dessen Rand $\zeta = \zeta_0$ die komplexe Streckung der Form

$$\tilde{\bar{\bar{\mathbf{G}}}} = \bar{\bar{\mathbf{S}}}(x,y,z) \cdot \bar{\bar{\mathbf{G}}} \cdot \bar{\bar{\mathbf{S}}}(x,y,z) \quad \text{mit} \quad \bar{\bar{\mathbf{S}}}(x,y,z) = \begin{bmatrix} s_x(x) & 0 & 0 \\ 0 & s_y(y) & 0 \\ 0 & 0 & s_z(z) \end{bmatrix}, \quad (2.108)$$

wobei die Einträge in $\bar{\bar{\mathbf{S}}}(x,y,z)$ aus (2.101) stammen und als Metrikkoeffizienten identifiziert werden. In der PML-Region gelten die Gleichungen [30]

$$\bar{\bar{\mathbf{S}}}^{-1} \cdot \text{rot } \vec{E}^c(\tilde{\vec{r}}) = -j\omega \vec{B}^c(\tilde{\vec{r}}) \quad \text{und} \quad \bar{\bar{\mathbf{S}}}^{-1} \cdot \text{rot } \vec{H}^c(\tilde{\vec{r}}) = j\omega \vec{D}^c(\tilde{\vec{r}}), \quad (2.109)$$

welche die Abhängigkeiten der Feldgrößen von den komplexen Ortskoordinaten nach Gleichung (2.106) darstellen. Die Umrechnung von (2.109) in reelle Orte führt auf Felder, welche die MAXWELL'schen Gleichungen (2.1) bis (2.4) innerhalb der PML tatsächlich erfüllen⁴⁰ [30]:

$$\vec{E}^a(\vec{r}) = \bar{\bar{\mathbf{S}}} \vec{E}^c(\bar{\bar{\mathbf{G}}} \cdot \vec{r}), \quad \vec{H}^a(\vec{r}) = \bar{\bar{\mathbf{S}}} \vec{H}^c(\bar{\bar{\mathbf{G}}} \cdot \vec{r}), \quad (2.110)$$

$$\vec{D}^a(\vec{r}) = \det(\bar{\bar{\mathbf{S}}}) \bar{\bar{\mathbf{S}}}^{-1} \vec{D}^c(\bar{\bar{\mathbf{G}}} \cdot \vec{r}), \quad \vec{B}^a(\vec{r}) = \det(\bar{\bar{\mathbf{S}}}) \bar{\bar{\mathbf{S}}}^{-1} \vec{B}^c(\bar{\bar{\mathbf{G}}} \cdot \vec{r}). \quad (2.111)$$

³⁹Euklid von Alexandria (ca. 360 v. Chr. - ca. 280 v. Chr.), griechischer Mathematiker.

⁴⁰Im Folgenden mit 'a' indiziert.

Diese Transformation erhält die Stetigkeit für \vec{E} - und \vec{H} -Komponenten, die bezüglich der Fläche zwischen innerem Rechengebiet und PML bei $\zeta = \zeta_0$ tangential liegen. Entsprechendes gilt für die Normalkomponenten von \vec{D} und \vec{B} , jedoch nicht für die jeweils anderen Komponenten, deren Stetigkeit gemäß den MAXWELL'schen Gleichungen nicht erforderlich ist. Aus (2.110) und (2.111) lassen sich die Verknüpfungen von Feldern und Flussdichten herleiten, die auf

$$\vec{D}^a(\vec{r}) = \det(\bar{\mathbf{S}})(\bar{\mathbf{S}}^{-1} \cdot \bar{\epsilon} \cdot \bar{\mathbf{S}}^{-1})\vec{E}^a \quad \text{und} \quad \vec{B}^a(\vec{r}) = \det(\bar{\mathbf{S}})(\bar{\mathbf{S}}^{-1} \cdot \bar{\mu} \cdot \bar{\mathbf{S}}^{-1})\vec{H}^a \quad (2.112)$$

führen. Für allgemeine lineare Medien⁴¹ mit $\bar{\epsilon}$ und $\bar{\mu}$ am Rechengebietsrand ergeben sich die PML-Parameter zu

$$\bar{\epsilon}_{\text{PML}} = \det(\bar{\mathbf{S}})\bar{\mathbf{S}}^{-1} \cdot \bar{\epsilon} \cdot \bar{\mathbf{S}}^{-1} \quad \text{und} \quad \bar{\mu}_{\text{PML}} = \det(\bar{\mathbf{S}})\bar{\mathbf{S}}^{-1} \cdot \bar{\mu} \cdot \bar{\mathbf{S}}^{-1}. \quad (2.113)$$

Dabei ist neben der Stetigkeit des Übergangs gewährleistet, dass die MAXWELL'schen Gleichungen ohne Modifikationen erfüllt sind. Die Formulierung in der Methode der finiten Integration ergibt sich aus der Diskretisierung der Gleichungen

$$\text{rot } \vec{E}^a = -j\omega\bar{\mu}_{\text{PML}}\vec{H}^a, \quad \text{rot } \vec{H}^a = j\omega\bar{\epsilon}_{\text{PML}}\vec{E}^a \quad (2.114)$$

analog zu (2.78) und (2.80) zu

$$\mathbf{C}\hat{\mathbf{e}}^a = -j\omega\mathbf{M}_{\mu}^{\text{PML}}(\omega)\hat{\mathbf{h}}^a, \quad \tilde{\mathbf{C}}\hat{\mathbf{h}}^a = j\omega\mathbf{M}_{\epsilon}^{\text{PML}}(\omega)\hat{\mathbf{e}}^a \quad (2.115)$$

mit den Diagonalmatrizen $\mathbf{M}_{\epsilon}^{\text{PML}}(\omega)$ und $\mathbf{M}_{\mu}^{\text{PML}}(\omega)$. Hierbei ist die Frequenzabhängigkeit der Materialmatrizen zu beachten, die in (2.101) unter Verwendung von (2.102) oder (2.103) einbezogen wird und sich über (2.108) bis in (2.113) fortpflanzt.

2.2.5. Fehler

Bei der Messung physikalischer Größen treten Fehler auf und auch bei der Simulation physikalischer Größen gibt es Fehler. Diese Fehler lassen sich folgendermaßen klassifizieren:

- **Klasse der Modellierungsfehler.** Alle Fehler, die auf dem Weg von einer physikalischen Problemstellung zum abstrahierten mathematischen Modell entstehen, zählen zu den Modellierungsfehlern. Dazu gehören vernachlässigte physikalische Effekte (Nichtlinearitäten, kleine Inhomogenitäten der Materialparameter, Temperaturabhängigkeiten,...), ungenaue oder idealisierte Materialparameter und Geometriedaten. Desweiteren stellen die Ränder des Rechengebiets und Symmetriebedingungen eine potentielle Quelle für Modellierungsfehler dar, ebenso wie die Verwendung nicht struktur-konformer Rechengitter.
- **Klasse der Verfahrensfehler.** Verfahrensfehler sind einem numerischen Verfahren immanent. Bei FIT entstehen Verfahrensfehler beispielsweise bei der Materialdiskretisierung nach (2.84) oder durch den der Zeitintegration (2.91) zugrunde liegenden zentralen Differenzenquotienten.
- **Klasse der numerischen Fehler.** Die Darstellung mancher rationalen und aller irrationalen Zahlen ist in primitiven Datentypen moderner Rechnersystemen nur mit endlicher Genauigkeit möglich, wodurch es zum datentypabhängigen Rundungsfehler kommt.

⁴¹Gilt nicht für nichtlineare, jedoch für bianisotrope Materialien mit äquivalenter Umrechnung.

Er kann als Quantisierungsfehler verstanden werden, da nur eine endliche Auswahl des kontinuierlichen Zahlenraums dargestellt werden kann. Für 64-bit Gleitkommawerte beträgt er 2.220446×10^{-16} , kann sich aber bei numerisch ungünstigen Verfahren (z.B. GAUSS-Elimination für schlecht konditionierte Gleichungssysteme) signifikant akkumulieren. Darüber hinaus entstehen numerische Fehler durch den Abbruch von iterativen Algorithmen, die insbesondere zum Lösen von Gleichungssystemen zum Einsatz kommen. Allerdings ist dieser Abbruchfehler in Form eines Residuums meistens gut bekannt und auch in gewissen Grenzen kontrollierbar, zumindest solange er deutlich größer als der datentypabhängige Rundungsfehler ist.

2.3. Methode der flexiblen lokalen Approximation (FLAME)

Die TREFFTZ⁴²-Methode der flexiblen lokalen Approximation (TREFFTZ-FLAME von engl. Flexible Local Approximation Methods) wird von TSUKERMAN seit dem Jahr 2004 propagiert [19]. Die hier verwendete Darstellung orientiert sich an [34, 35, 37]. Anders als bei der Methode der finiten Integration aus Abschnitt 2.2 wird FLAME auf abgeleitete, der jeweiligen Aufgabenstellung angepasste Differenzialgleichungen angewendet. Den Hauptanwendungsbereich stellen dabei die Randwertprobleme dar. Engere Verwandtschaft ist zur Methode der finiten Differenzen gegeben, bei der diskrete Formulierungen durch TAYLOR'sche⁴³ Reihenentwicklung der Differenzialoperatoren gewonnen werden.

2.3.1. Differenzenschemata in FLAME

Das Rechengebiet Ω wird in überlappende Teilgebiete $\Omega^{(i)}$ entsprechend Abbildung 2.9 zerlegt. Jedes Teilgebiet bekommt einen Differenzenstern (engl. stencil) zugewiesen, der benachbarte Knoten des globalen (aus Anschauungsgründen meist kartesischen) Rechengitters miteinander verknüpft. Die verschiedenen Knoten innerhalb eines Differenzensterns sind mit einem eindeutigen Index versehen. Die zugrunde liegende Differenzialgleichung wird innerhalb eines Teilgebiets $\Omega^{(i)}$ als homogen mit dem Differenzialoperator L und der Lösung u

$$Lu = 0 \quad \text{in} \quad \Omega^{(i)} \subset \mathbb{R}^{2,3} \quad (2.116)$$

angenommen. Auf dem Teilgebiet (i) , mit dem äquidistanten Knotenabstand h , wird die Approximation der exakten Lösung $u_h^{(i)}$ als Linearkombination von m dem Differenzenstern zugeordneten Basisfunktionen $\{\psi_l^{(i)}\}$ angesetzt

$$u_h^{(i)} = \sum_{l=1}^m c_l^{(i)} \psi_l^{(i)}. \quad (2.117)$$

Der Idee von TREFFTZ [33] folgend werden nur solche Basisfunktionen $\{\psi_l^{(i)}\}$ zugelassen, die eine analytisch exakte Lösung der zugrunde liegenden Differenzialgleichung (2.116) darstellen. Als weitere Voraussetzung an die Basisfunktionen wird gefordert, dass diese die exakte Lösung

⁴²Erich Immanuel Trefftz (1888-1937), deutscher Mechaniker, Mathematiker.

⁴³Brook Taylor (1685-1731), britischer Mathematiker.

u_{ex} auf einem lokalen Differenzenstern hinreichend gut approximieren. Dies ist in erster Linie dann möglich, wenn die Eigenschaften der exakten Lösung u_{ex} zumindest lokal bekannt sind. Für viele fundamentale geometrische Formen gibt es analytische Ausdrücke, die lokal (also innerhalb oder in der Nähe des jeweiligen Objekts) gute Approximationen der gesuchten Lösung darstellen. Die Beziehung des Koeffizientenvektors $\mathbf{c}^{(i)} \in \mathbb{R}^{m \times 1}$ mit den Knotenwerten der Lösung $\mathbf{u}^{(i)}$ ist durch

$$\mathbf{N}^{(i)} \mathbf{c}^{(i)} = \mathbf{u}^{(i)} \quad (2.118)$$

gegeben. Dabei ist $\mathbf{N}^{(i)}$ die Matrix der Knotenwerte, in der alle m Basisfunktionen $\{\psi_l^{(i)}\}$ an den n Stützstellen des Differenzensterns von Teilgebiet $\Omega^{(i)}$ ausgewertet werden

$$\mathbf{N}^{(i)} = \begin{pmatrix} \psi_1^{(i)}(\vec{r}_1) & \psi_2^{(i)}(\vec{r}_1) & \dots & \psi_m^{(i)}(\vec{r}_1) \\ \psi_1^{(i)}(\vec{r}_2) & \psi_2^{(i)}(\vec{r}_2) & \dots & \psi_m^{(i)}(\vec{r}_2) \\ \dots & \dots & \dots & \dots \\ \psi_1^{(i)}(\vec{r}_n) & \psi_2^{(i)}(\vec{r}_n) & \dots & \psi_m^{(i)}(\vec{r}_n) \end{pmatrix}. \quad (2.119)$$

Im Allgemeinen werden für Differenzensterne über die n Stützstellen $m = n - 1$ Basisfunktionen gewählt, so dass die Matrix der Knotenwerte rechteckig wird [36]. Für einen n -Punkt Differenzenstern ist ein Differenzenschema $\mathbf{s}^{(i)} \in \mathbb{R}^{n \times 1}$ gesucht, das

$$\mathbf{s}^{(i)T} \mathbf{u}^{(i)} = 0 \quad (2.120)$$

für Knotenwerte $\mathbf{u}^{(i)}$ jeglicher Funktionen $u_h^{(i)}$ aus (2.117) erfüllt. Aus (2.118) und (2.120) folgt unmittelbar

$$\mathbf{s}^{(i)T} \mathbf{N}^{(i)} \mathbf{c}^{(i)} = 0. \quad (2.121)$$

Um (2.121) für beliebige Koeffizientenvektoren $\mathbf{c}^{(i)}$ erfüllen zu können, muss

$$\mathbf{s}^{(i)} \in \ker \mathbf{N}^{(i)T} \quad (2.122)$$

gelten. Das lokale Differenzenschema⁴⁴ $\mathbf{s}^{(i)}$ eines zugeordneten Teilgebiets $\Omega^{(i)}$ enthält jene Koeffizienten mit denen benachbarte Knotenwerte für die lokale Approximation der Lösung gewichtet werden, so dass das lokale ortsabhängige Differenzenschema mit der i -ten Zeile der Systemmatrix koinzidiert. Ist der Defekt⁴⁵ $\text{def}(\mathbf{N}^{(i)T})$ aus (2.122) Eins und ist $\mathbf{s}^{(i)}$ nicht-trivial, so repräsentiert $\mathbf{s}^{(i)}$ das gesuchte TREFFTZ-FLAME Differenzenschema bis auf einen beliebigen Vorfaktor. Genau dann ist durch die Gleichungen in $\mathbf{N}^{(i)T} \mathbf{s}^{(i)} = 0$ gewährleistet, dass die angesetzten Basisfunktionen $\{\psi_l^{(i)}\}$ das Differenzenschema erfüllen, da sie mit $\mathbf{s}^{(i)}$ gewichtet sind. Neben dieser erforderlichen Eigenschaft im *diskreten* Modell lässt der verwendete TREFFTZ-Ansatz nur solche Basisfunktionen zu, welche die zugrunde liegende *kontinuierliche* Differenzialgleichung erfüllen.

2.3.2. Eigenschaften

Für harmonische Polynome als Basisfunktionen⁴⁶ resultiert mit TREFFTZ-FLAME, bis auf einen beliebigen Vorfaktor, das gleiche Differenzenschema und somit die gleichen Matrixeinträge, auf die auch die taylorreihenbasierte Formulierung der Methode der finiten Differenzen

⁴⁴Die resultierenden Schemata sind immer eine Differenz bzw. algebraische Summe.

⁴⁵ $\text{def}(\mathbf{A}) = \dim(\ker(\mathbf{A}))$ entsprechend Satz 4.1.10

⁴⁶Beispielsweise zur Lösung der LAPLACE'schen Differenzialgleichung.

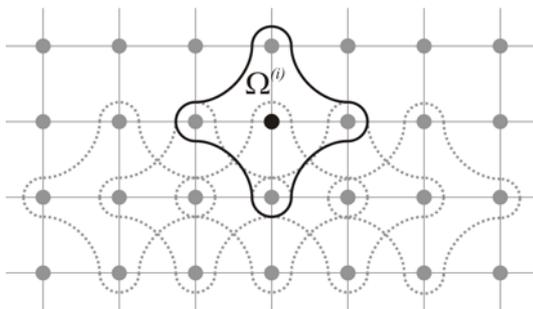


Abbildung 2.9.: Überlappende Teilgebiete $\Omega^{(i)}$ am Beispiel des 5-Punkt-Differenzensterns.

führt. In verschiedenen Teilgebieten können unterschiedliche TREFFTZ'sche Basisfunktionen zugelassen werden⁴⁷. Da die Stetigkeitsbedingungen an Materialübergängen algebraisch statt geometrisch ausgedrückt werden können, beschreiben lokale Basisfunktionen das Feldverhalten an der Materialgrenze, so dass ein zur Geometrie konformes Rechengitter nicht notwendig ist. Randbedingungen können in lokalen Teilgebieten berücksichtigt werden: Für Randwertprobleme sind das meist NEUMANN'sche⁴⁸ und DIRICHLET'sche⁴⁹ Randbedingungen. Zusammenfassend lassen sich drei vorteilhafte Hauptmerkmale angeben:

1. Basisfunktionen sind nicht auf Polynome limitiert.
2. Die Lösung wird approximiert – nicht der Differenzialoperator.
3. Die Auswertung der Basisfunktionen ist auch zwischen den Knoten möglich.

Nachteilhaft ist, dass die Forderung $\text{def}(\mathbf{N}^{(i)T}) = 1$ nicht a priori feststellbar ist. Für Fälle, in denen $\text{def}(\mathbf{N}^{(i)T}) = 0$ entsteht, kann die Anzahl der Stützstellen erhöht oder die Anzahl der Basisfunktionen reduziert werden. Für $\text{def}(\mathbf{N}^{(i)T}) > 1$ kann entweder das jeweils Genteilige durchgeführt werden oder mit zusätzlichen Kriterien eine Linearkombination der unabhängigen Kernvektoren verwendet werden. Darüber hinaus hat ebenso die Auswahl der in Frage kommenden Basisfunktionen als auch die tatsächlichen Positionen der Stützstellen eines Differenzensterns im Raum einen Einfluss auf $\text{def}(\mathbf{N}^{(i)T})$.

Die algebraische Beschreibung von graduellen Materialübergängen⁵⁰ kann dazu führen, dass lokale analytische Approximationen für ein Teilgebiet $\Omega^{(i)}$ nicht existieren. Als Ausweg bieten sich numerisch erzeugte Basisfunktionen⁵¹ oder aber hinreichende Rechengitterverfeinerung an [17]. Aufgrund der lokal unterschiedlichen Basisfunktionen wird die Systemmatrix im Allgemeinen – auch für selbstadjungierte Operatoren – unsymmetrisch.

2.3.3. Fehler und Konvergenz

TREFFTZ-FLAME ist zunächst ein heuristischer Ansatz in dem Sinne, dass der Methode kein globaler Konvergenzbeweis zugrunde liegt. Zwei wichtige Fehler-Definitionen kommen zur Anwendung.

⁴⁷Theoretisch darf auch die Anzahl der Stützstellen pro lokalem Differenzenstern variieren.

⁴⁸Carl Gottfried Neumann (1832-1925), deutscher Mathematiker.

⁴⁹Johann Peter Gustav Lejeune Dirichlet (1805-1859), deutscher Mathematiker.

⁵⁰z.B. in geschichteten Medien oder Körpern mit geringem Abstand.

⁵¹Die numerisch generierten Basisfunktionen müssen entsprechend dem TREFFTZ-Ansatz die zugrunde liegende Differenzialgleichung erfüllen.

2.3.3.1. Approximationsfehler

Der Vektor $\mathbf{u}_{ex}^{(i)}$ enthält die Knotenwerte der exakten Lösung u_{ex} an den Stützstellen des Differenzensterns (i) . Der lokale Approximationsfehler am Differenzenstern (i) ist dann durch die Vektornorm der Differenz von exakter Lösung $\mathbf{u}_{ex}^{(i)}$ und berechneter Lösung $\mathbf{u}_h^{(i)}$ gegeben durch

$$\epsilon_a^{(i)} = \|\mathbf{u}_h^{(i)} - \mathbf{u}_{ex}^{(i)}\|. \quad (2.123)$$

Der Vektor $\mathbf{u}_h^{(i)}$ hängt hierbei von der gewählten Rechengitterschrittweite h ab und für die Konvergenz ist notwendig, dass $\epsilon_a^{(i)} \rightarrow 0$ gilt, wenn $h \rightarrow 0$ geht. Der Approximationsfehler (2.123) ist somit ein Vorwärtsfehler ⁵².

2.3.3.2. Konsistenzfehler

Während zur Bestimmung des Approximationsfehlers die Lösung $\mathbf{u}_h^{(i)}$ für das gesamte Rechengebiet vorliegen muss, lässt sich der Konsistenzfehler lokal mit dem Differenzenschema $\mathbf{s}^{(i)T}$ und der exakten Lösung $\mathbf{u}_{ex}^{(i)}$ auf dem Teilgebiet $\Omega^{(i)}$ auswerten. Der lokale Konsistenzfehler ergibt sich nach der Definition von TSUKERMAN entsprechend [34, 35, 37] durch die Gewichtung der knotenweisen exakten Lösung $\mathbf{u}_{ex}^{(i)}$ mit dem berechneten Differenzenschema (2.122) zu

$$\epsilon_k^{(i)} = \mathbf{s}^{(i)T} \mathbf{u}_{ex}^{(i)}. \quad (2.124)$$

Ein Differenzenschema $\mathbf{s}^{(i)}$ wird konsistent genannt, wenn der Konsistenzfehler für sukzessive Rechengitterverfeinerung in h (bei äquidistanten) oder $\max\{h_x, h_y, h_z\}$ (für raumrichtungsabhängig verschiedene Rechengitterschrittweiten) gegen Null strebt

$$\lim_{h \rightarrow 0} \epsilon_k^{(i)} = 0. \quad (2.125)$$

Da die Differenzenschemata in (2.122) nur bis auf einen beliebigen Vorfaktor bestimmt sind, geht dieser Vorfaktor auch in die Berechnung des Konsistenzfehlers (2.124) mit ein. Die Performanz verschiedener Differenzenschemata lässt sich dennoch vergleichen, wenn sie normiert werden, d.h. der Vorfaktor wird so gewählt, dass $\|\mathbf{s}\| = 1$ gilt. Für diesen Fall der normierten Differenzenschemata sind Approximationsfehler (2.123) und Konsistenzfehler nach (2.124) identisch [15] und es gilt

$$\|\mathbf{s}^{(i)T} \mathbf{u}_{ex}^{(i)}\| = \|\mathbf{s}^{(i)T} \mathbf{u}_h^{(i)} + \mathbf{s}^{(i)T} (\mathbf{u}_{ex}^{(i)} - \mathbf{u}_h^{(i)})\| \leq \|\mathbf{0}\| + \|\mathbf{s}\| \|\mathbf{u}_{ex}^{(i)} - \mathbf{u}_h^{(i)}\| \quad (2.126)$$

$$= \|\mathbf{u}_{ex}^{(i)} - \mathbf{u}_h^{(i)}\| \rightarrow 0. \quad (2.127)$$

Allgemein ($\|\mathbf{s}\| \neq 1$) ist der Konsistenzfehler durch den Approximationsfehler beschränkt. Ausgehend von diesen Ergebnissen lässt sich die Konvergenz der Methode für die Klasse der Differenzenschemata mit monotonem Differenzenoperator beweisen [36].

2.3.4. Minimalbeispiel

Die Vorgehensweise zur Erzeugung des numerischen Modells wird im Folgenden an einem Minimalbeispiel demonstriert. Dazu ist in Abbildung 2.10a ein äquidistantes Rechengitter

⁵²Siehe Definition 4.1.23.

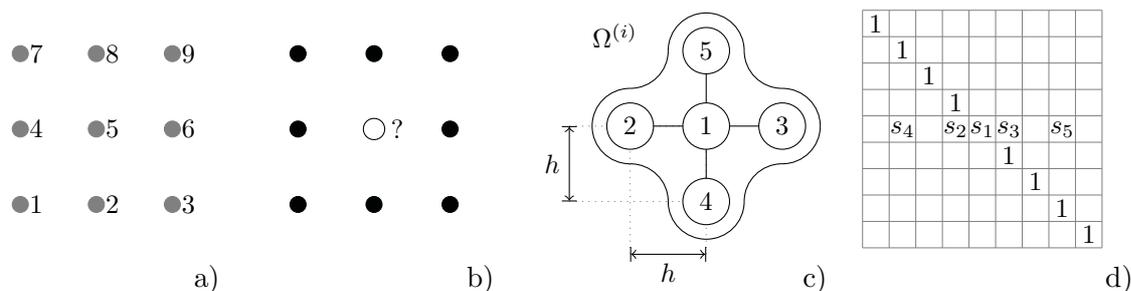


Abbildung 2.10.: TREFFTZ-FLAME von der Diskretisierung zur Systemmatrix: a) und b) Globales Indizierungsschema eines Randwertproblems mit unbekanntem Knotenwert 5. c) 5-Punkt Differenzenstern für den unbekannten mittleren Knoten in a) und b). Das lokale Indizierungsschema in c) führt auf die Spaltenpositionen der Elemente des Kernvektors in der resultierenden Systemmatrix d).

mit 3×3 Punkten und zugehörigem Indizierungsschema angegeben. Die Formulierung als Randwertproblem führt nach Abbildung 2.10b auf einen unbekanntem Gitterpunkt (Index 5). Das Rechengebiet unbekannter Gitterpunkte wird im nächsten Schritt vollständig mit überlappenden Teilgebieten $\Omega^{(i)}$ abgedeckt. Hierzu ist in diesem Minimalbeispiel lediglich ein Teilgebiet notwendig. Entsprechend Abbildung 2.10c wird für jeden Gitterpunkt (i) ein 5-Punkt-Differenzenstern auf den Stützstellen des Rechengitters gebildet, der mit einem lokalen Indizierungsschema versehen wird. Abhängig von der im Rechengitter zu lösenden Differenzialgleichung werden für jedes Teilgebiet $\Omega^{(i)}$ Basisfunktionen $\psi_m^{(i)}$ ausgewählt, welche einerseits die erwartete Lösung auf dem lokalen Teilgebiet gut approximieren und andererseits die analytische Differenzialgleichung selbst erfüllen. In Abbildung 2.11 ist dargestellt, wie für jedes Teilgebiet $\Omega^{(i)}$ eine Zeile in der Systemmatrix $\mathbf{A}_{\text{FLAME}}$ erzeugt wird. Zunächst ist zu entscheiden, ob der aktuelle Gitterpunkt zu den Unbekannten oder zum Rand gehört, so dass gegebenenfalls die rechte Seite des linearen Gleichungssystems ergänzt wird. Handelt es sich um Gitterpunkte, die Unbekannte repräsentieren, werden die m Basisfunktionen an den n Knoten des Differenzensterns ausgewertet und als Knotenmatrix $\mathbf{N}^{(i)}$ nach Gleichung (2.119) angeordnet. Im nächsten Schritt wird der Kern der Knotenmatrix $\mathbf{s}^{(i)} = \ker \mathbf{N}^{(i)T}$ nach Gleichung (2.122) bestimmt. Der Vektor $\mathbf{s}^{(i)}$ enthält unmittelbar die Einträge einer Zeile der aufzustellenden Systemmatrix $\mathbf{A}_{\text{FLAME}}$. Die Elemente von $\mathbf{s} = [s_1, s_2, s_3, s_4, s_5]^T$ sind entsprechend ihrer zugehörigen Position im globalen Punktgitter in die Systemmatrix einzutragen. Abbildung 2.10d zeigt die Zuordnung für das betrachtete Minimalbeispiel⁵³. Eine Vielzahl weiterer Beispiele sowie die Konstruktion von inhomogenen und nichtlinearen Differenzenschemata sind in der eingangs zitierten Literatur enthalten.

⁵³Selbstverständlich lassen sich die hier nicht benötigten Randwerte der Ecken aus der Matrix eliminieren, worauf aber aus Anschauungsgründen verzichtet wird.

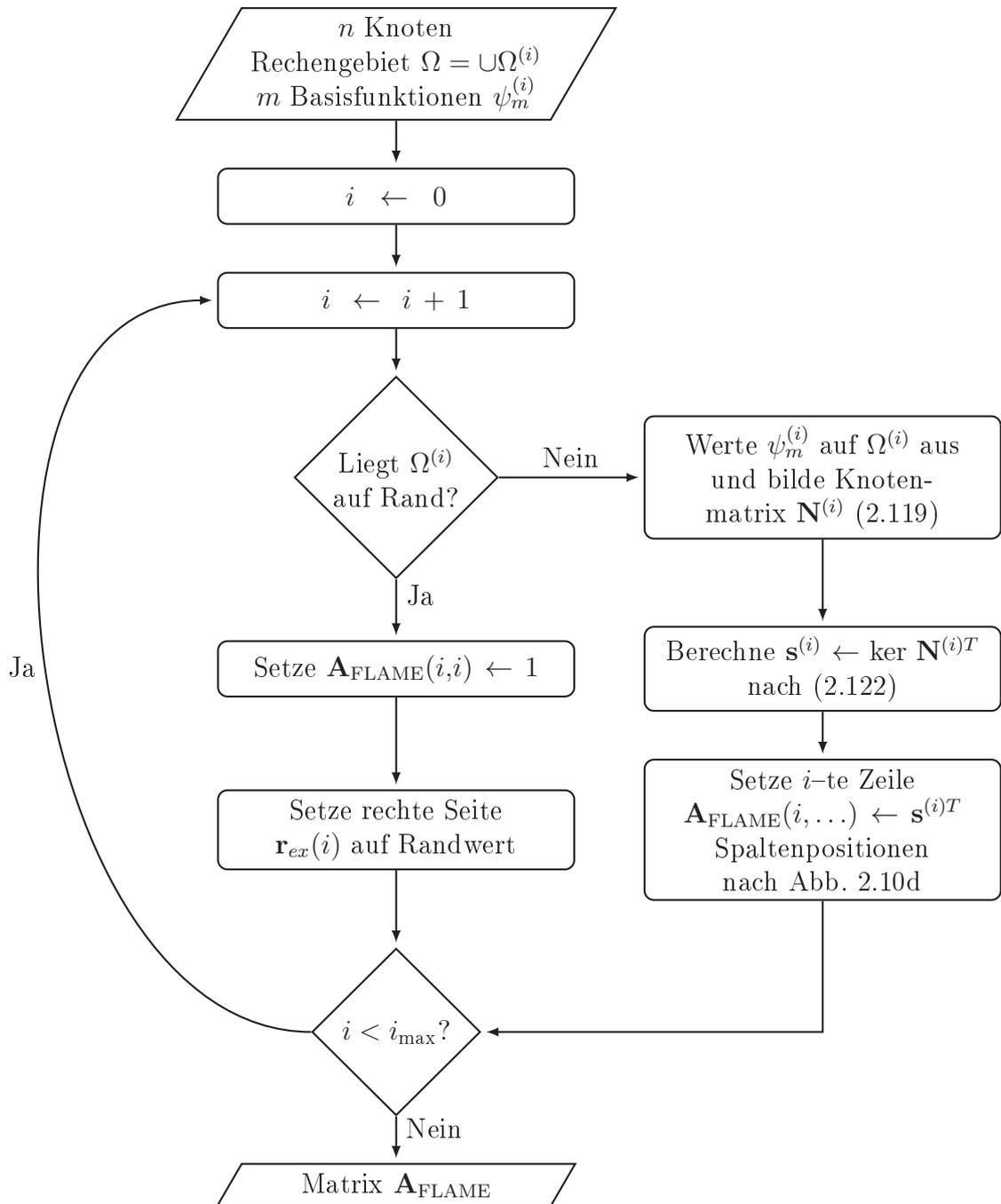


Abbildung 2.11.: Vorgehensweise zur Erzeugung der Systemmatrix und der rechten Seite für das Randwertproblem $\mathbf{A}_{\text{FLAME}}(h)\mathbf{u}_h = \mathbf{r}_{ex}$ in der FLAME-Diskretisierung.

3. Formulierungen für periodische und isolierte Nanostrukturen

In diesem Kapitel werden die algebraischen Problemstellungen zur Berechnung der elektromagnetischen Eigenschaften der betrachteten Nanostrukturen formuliert. Zum einen werden periodische Nanostrukturen mit adäquater Randbedingung betrachtet und zum anderen werden Formulierungen für einzelne Nanostrukturen im Freiraum aufgestellt, zu deren Modellierung eine transparente Randbedingung notwendig ist. Das numerische Gitter wird im Folgenden als Rechengitter bezeichnet, während die periodische Struktur das Strukturgitter bildet. Die Nomenklatur der Zustandsgrößen der Methode der finiten Integration bleibt hiervon unberührt.

3.1. Streumatrix–Ansatz zur Berechnung der Dispersion in periodischen Strukturen

Ausgangspunkt für die Berechnung der Dispersionseigenschaften periodischer Strukturen ist die Streumatrix $\mathcal{S}(\omega)$ nach Gleichung (2.63), die im Rahmen dieser Arbeit entweder aus der Simulation einer einzelnen Elementarzelle einer periodischen Anordnung mit CST MICROWAVE STUDIO im Zeitbereich (2.91) oder im Frequenzbereich (2.92) gewonnen wird. Für die Simulation der Streumatrix ist die Verwendung der Methode der finiten Integration aus Abschnitt 2.2 jedoch nicht unabdingbar, da die Streumatrix auch mit anderen Formulierung, wie z.B. der Finite-Elemente-Methode berechnet werden kann. Eine Elementarzelle ist der Baustein, aus dem durch Translation entsprechend Abbildung 2.3a das periodische Strukturgitter erzeugt wird. In Abbildung 3.1a ist dieser Sachverhalt für ein dreidimensionales Strukturgitter skizziert. Die Randbedingungen einer Parallelplattenleitung sind für das Si-

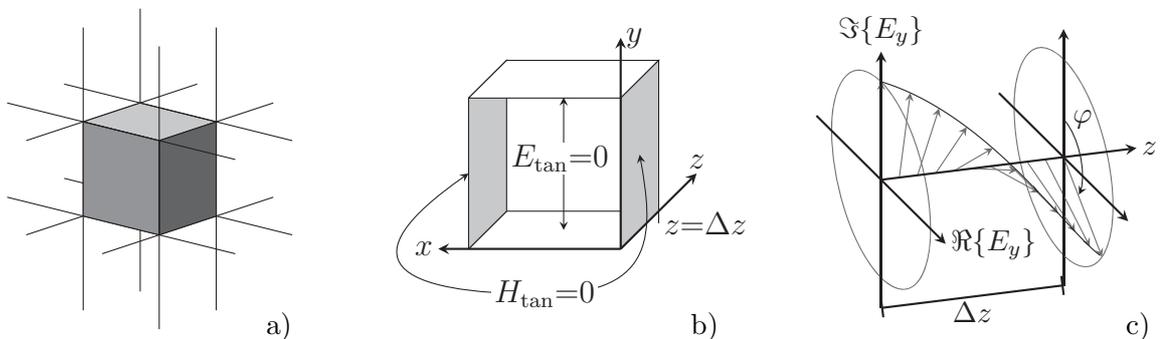


Abbildung 3.1.: a) Elementarzelle als Ausschnitt eines unendlich ausgedehnten periodischen Strukturgitters. b) Koordinatensystem und gewählte Randbedingungen. c) Veranschaulichung des Phasenvorschubs von \vec{E} einer Welle bei Propagation in einem verlustlosen Medium über die Länge Δz .

mulationsmodell einer einzelnen Elementarzelle in Abbildung 3.1b angegeben. Grund für die Wahl dieser Randbedingungen ist, dass dieses Wellenleitermodensystem die für die Analyse der Strukturen notwendige Periodizität hinreichend genau modelliert. In z -Richtung, der Ausbreitungsrichtung, hat die Elementarzelle eine Länge von Δz , was auch gleichzeitig die Strukturgitterkonstante a_{per} ist.

3.1.1. Streumatrix für fundamentale Wellenleitermoden

Liegt die Streumatrix einer Elementarzelle für den fundamentalen Wellenleitermode vor, so kann a posteriori die Periodizität gefordert werden. Mit der Festlegung der Wellenamplituden entsprechend Abbildung 2.5 soll für die Propagation durch eine Elementarzelle der Dicke Δz entsprechend Abbildung 3.1c die BLOCH-Bedingung (2.26) für die Wellenamplituden gelten

$$b_1(\omega) = a_2(\omega)e^{j\varphi(\omega)}, \quad b_2(\omega) = a_1(\omega)e^{-j\varphi(\omega)}. \quad (3.1)$$

Hierbei propagieren $a_1(\omega)$ und $b_2(\omega)$ in Richtung der aufsteigenden z -Koordinate. Für den Phasenvorschub $\varphi(\omega)$ gilt

$$\varphi(\omega) = k_{\text{per}}(\omega)\Delta z = (\beta_{\text{per}}(\omega) - j\alpha_{\text{per}}(\omega))\Delta z \quad \text{und} \quad \varphi(\omega), k_{\text{per}}(\omega) \in \mathbb{C} \quad (3.2)$$

mit der unbekanntenen Blochwellenzahl $k_{\text{per}}(\omega)$, die sich aus der periodischen Ausbreitungskonstante $\beta_{\text{per}}(\omega)$ und der Dämpfungskonstante $\alpha_{\text{per}}(\omega)$ zusammensetzt. Mit der BLOCH-Bedingung für die Wellenamplituden (3.1) folgt für die Streumatrix des Zweitorts nach Gleichung (2.63)

$$\begin{pmatrix} b_1(\omega) \\ b_2(\omega) \end{pmatrix} = \begin{pmatrix} s_{11}(\omega) & s_{12}(\omega) \\ s_{21}(\omega) & s_{22}(\omega) \end{pmatrix} \begin{pmatrix} a_1(\omega) \\ a_2(\omega) \end{pmatrix} = \begin{pmatrix} a_2(\omega)e^{+j\varphi(\omega)} \\ a_1(\omega)e^{-j\varphi(\omega)} \end{pmatrix}. \quad (3.3)$$

Dieses 2×2 Gleichungssystem lässt sich nach $e^{j\varphi(\omega)}$ auflösen, wie im Anhang A.1 gezeigt ist. Unter der Annahme der Berücksichtigung eines einzigen Wellenleitermodes ist die Problemstellung äquivalent zur Berechnung der Eigenwerte der Transfer-Streumatrix \mathcal{T} nach (2.66), die im folgenden Abschnitt beschrieben wird.

3.1.2. Streumatrix für mehrere Wellenleitermoden

Der Ansatz aus dem vorherigen Abschnitt kann auf Wellenleitermoden höherer Ordnung verallgemeinert werden¹[1]. Dazu wird die in Abschnitt 2.1.5.3 eingeführte multimodale Streumatrix verwendet. Am Tor n werden die Vektoren $\mathbf{a}_n(\omega)$ und $\mathbf{b}_n(\omega)$ aus Gleichung (2.65) gebildet, die jeweils i berücksichtigte Wellenleitermoden umfassen. Um nach dem unbekanntenen Phasenvorschub φ bei einer gegebenen Frequenz ω auflösen zu können, wird die Transfer-Streumatrix \mathcal{T} nach (2.66) verwendet, in welche die BLOCH-Bedingung für die Wellenamplituden aus Gleichung (3.1) eingesetzt wird

$$\begin{pmatrix} \mathbf{b}_1(\omega) \\ \mathbf{a}_1(\omega) \end{pmatrix} = \begin{pmatrix} \mathcal{T}_{11}(\omega) & \mathcal{T}_{12}(\omega) \\ \mathcal{T}_{21}(\omega) & \mathcal{T}_{22}(\omega) \end{pmatrix} \begin{pmatrix} \mathbf{a}_2(\omega) \\ \mathbf{b}_2(\omega) \end{pmatrix} = e^{j\varphi} \begin{pmatrix} \mathbf{a}_2(\omega) \\ \mathbf{b}_2(\omega) \end{pmatrix}. \quad (3.4)$$

¹Die Untersuchungen zur Berücksichtigung Moden höherer Ordnung im Streumatrixansatz ist in Kooperation mit Dr.-Ing. Grzegorz Lubkowski und Prof. Dr.-Ing. Thomas Weiland vom Institut Theorie elektromagnetischer Felder an der TU Darmstadt entstanden [142].

Gesucht wird in Gleichung (3.4) ein Skalar $e^{j\varphi}$, der die gleichen Abbildungseigenschaften wie die Transfer–Streumatrix \mathcal{T} besitzt. Diese Problemstellung ist äquivalent zur Lösung des komplexen Eigenwertproblems

$$\mathcal{T}\mathbf{v} = \gamma\mathbf{v} \quad (3.5)$$

für jeden einzelnen Frequenzpunkt ω . Das Eigenpaar (\mathbf{v}_m, γ_m) wird im Folgenden als Eigenzustand bezeichnet². Hierbei gewichtet der Eigenvektor \mathbf{v}_m genau diejenigen Wellenamplituden, die superponiert werden müssen, um die Wellenamplitudenverteilung der periodischen Strukturgitterwelle am Tor n zu erhalten. Demnach ist der Eigenvektor des m -ten periodischen Eigenzustands als

$$\mathbf{v}_m = (\dots a_{nl} \dots, \dots b_{nl} \dots)^T \quad l \in [1, i] \quad (3.6)$$

zu interpretieren. Die Dimension des Eigenwertproblems entspricht der Anzahl der Tore multipliziert mit der Anzahl i der berücksichtigten Wellenleitermoden. Für die hier betrachteten Zweitore gilt $\dim(\mathcal{T}) = 2i$. Im Fall verlustfreier Elementarzellen sind die ungedämpft ausbreitungsfähigen Eigenzustände anhand $|\gamma_m| = 1$ identifizierbar und für evaneszente Eigenzustände $|\gamma_m| \neq 1$, jeweils an einem Frequenzpunkt ω . Dementsprechend gilt für ausbreitungsfähige Eigenzustände verlustbehafteter Elementarzellen $|\gamma_m| \cong 1$. Die Eigenzustände von \mathcal{T} treten mindestens³ paarweise auf: Für jeden Eigenzustand γ , der für die Transmission in eine Koordinatenrichtung durch das Strukturgitter steht, existiert ein Eigenzustand γ^{-1} , der für die Transmission in die entgegengesetzte Richtung steht. Eine Möglichkeit, die Ausbreitungsrichtung der Eigenzustände festzulegen, ist die Berechnung des jeweiligen Leistungsflusses aus den einzelnen Wellenamplituden am Tor n . In

$$P_n(\mathbf{v}) = \sum_{l=1}^r (|a_{nl}|^2 - |b_{nl}|^2) + 2 \operatorname{Im} \left\{ \sum_{l=r+1}^i (a_{nl} b_{nl}^*) \right\} \quad (3.7)$$

gibt r die Anzahl der ausbreitungsfähigen und i die Anzahl der insgesamt berücksichtigten Wellenleitermoden an [1]. Die gesuchte Dispersionsrelation berechnet sich aus den Eigenzuständen γ_m mit $|\gamma_m| \cong 1$ durch $\varphi(\omega) = \angle(\gamma_m)$.

3.2. Eigenwertformulierungen zur Berechnung der Dispersion in periodischen Strukturen

Eigenwertprobleme, die in der Methode der finiten Integration formuliert werden, haben die elektrischen oder magnetischen Gitterspannungen als Freiheitsgrade, wie beispielsweise Gleichung (2.93). Daraus folgt bei einer Diskretisierung mit n_p Raumpunkten nach Gleichung (2.69) und drei Vektorkomponenten pro Raumpunkt im Allgemeinen eine Dimension der Systemmatrix⁴ von $3n_p$. Während das resultierende Eigenwertproblem (2.93) für geschlossene Randbedingungen nach Abschnitt 2.2.4.1 ausführlich in [24] untersucht wurde, wird in diesem Abschnitt ausführlich auf das Eigenwertproblem mit periodischen Rändern eingegangen [41], wobei unterschiedliche Anforderungen und Struktureigenschaften berücksichtigt werden.

²Um eine Abgrenzung zur Bezeichnung „Mode“ zu schaffen.

³Für ausbreitungsfähige orthogonale Polarisierungen auch Vielfache davon.

⁴Konservative Randbedingungen nach Abschnitt 2.2.4.1 können die Anzahl der Freiheitsgrade geringfügig reduzieren.

3.2.1. Verlustfreie Elementarzellen

Die Periodizität des zugrunde liegenden Rechengebiets kann nach Abschnitt 2.2.4.2 und Gleichung (2.98) in den \mathbf{L}_φ -Operator und die Materialmatrix $\mathbf{M}_{\varepsilon,\text{per}}$ implementiert werden, so dass die algebraische Eigenwertgleichung

$$\mathbf{L}_\varphi^H \tilde{\mathbf{C}} \mathbf{M}_\mu^{-1} \mathbf{C} \mathbf{L}_\varphi \hat{\mathbf{e}} = \omega^2 \mathbf{M}_{\varepsilon,\text{per}} \hat{\mathbf{e}} \quad \text{mit } \omega \in \mathbb{R} \text{ und } \varphi \in \mathbb{R} \quad (3.8)$$

folgt. In den Sonderfällen $\varphi = 0$ und $\varphi = \pi$ wird die Systemmatrix $\mathbf{L}_\varphi^H \tilde{\mathbf{C}} \mathbf{M}_\mu^{-1} \mathbf{C} \mathbf{L}_\varphi$ reell symmetrisch, ansonsten komplex hermitesch. Die Eigenwerte ω^2 geben Frequenzen an, bei denen eine Welle im unendlich ausgedehnten Strukturgitter existieren kann. Das Dispersionsdiagramm wird durch das sukzessive Lösen von (3.8) für verschiedene Phasenvorschübe φ erhalten, was für Elementarzellen beispielsweise in [29] und [41] gezeigt wurde.

Analog zur vektoranalytischen Identität (2.14) lässt sich (3.8) um einen diskreten grad-div-Term zur Nabla-Quadrat-Eigenwertgleichung

$$\mathbf{L}_\varphi^H (\tilde{\mathbf{C}} \mathbf{M}_\mu^{-1} \mathbf{C} + \sigma_S \mathbf{M}_\varepsilon \tilde{\mathbf{S}}^T \mathbf{D}_N \tilde{\mathbf{S}} \mathbf{M}_\varepsilon) \mathbf{L}_\varphi \hat{\mathbf{e}} = \omega^2 \mathbf{M}_{\varepsilon,\text{per}} \hat{\mathbf{e}} \quad (3.9)$$

ergänzen [24]. Vorteilhaft ist, dass durch die eingearbeitete Divergenzbeziehung $\tilde{\mathbf{S}} \mathbf{M}_\varepsilon \hat{\mathbf{e}} = 0$ der in Gleichung (2.93) n_p -fach vorhandene statische Eigenwert Null um den Faktor σ_S derart im Spektrum verschoben werden kann, dass der neue kleinste Eigenwert in (3.9) gerade derjenige ist, der von technischem Interesse ist. Die $n_p \times n_p$ -Matrix \mathbf{D}_N implementiert die ideal elektrisch leitfähigen Rand- und Oberflächenbedingungen für die Divergenzbeziehung (2.81).

3.2.2. Verlustbehaftete Elementarzellen, zeitliches Abklingverhalten

Zur Berücksichtigung von schwachen Verlusten, die z.B. in Form von endlichen Leitfähigkeiten bei Beschleunigerkavitäten auftreten, existieren a posteriori Verfahren, die auf der Störungstheorie basieren, um die Güte (2.49) zu bestimmen, obwohl zunächst ein reelles Eigenwertproblem gelöst wird [42]. Sind die Leitfähigkeiten jedoch so gering, dass die Eindringtiefe größer als die eigentliche Struktur wird, ist eine vollständige Rechnung notwendig, was im Folgenden für periodische Strukturen gezeigt wird.

Für die Dispersionsmodelle (2.36) und (2.37) aus Abschnitt 2.1.3, bei denen die Frequenzabhängigkeit sowohl der Materialparameter als auch der daraus resultierenden Verluste explizit bekannt sind, gilt ganz allgemein

$$\varepsilon(\omega) = \varepsilon_0 \left(\varepsilon_\infty + \frac{\beta_0 + j\omega\beta_1}{\alpha_0 + j\omega\alpha_1 - \omega^2} \right), \quad (3.10)$$

mit den reellwertigen Parametern ε_∞ , α_0 , α_1 , β_0 und β_1 bei vorausgesetzter $e^{j\omega t}$ -Zeitabhängigkeit nach Gleichung (2.15). Die kontinuierlichen MAXWELL'schen Gleichungen (2.1) bis (2.4) lassen sich im Frequenzbereich als Eigenwertgleichung für das elektrische Feld \vec{E}

$$\text{rot } \frac{1}{\mu} \text{rot } \vec{E} = \omega^2 \varepsilon(\omega) \vec{E} \quad (3.11)$$

herleiten. Einsetzen von Gleichung (3.10) in (3.11) führt auf

$$\text{rot } \frac{1}{\mu} \text{rot } \vec{E} = \omega^2 \varepsilon_0 \left(\varepsilon_\infty + \frac{\beta_0 + j\omega\beta_1}{\alpha_0 + j\omega\alpha_1 - \omega^2} \right) \vec{E}, \quad (3.12)$$

was sich nach der Sortierung bezüglich der Potenzen in ω als polynomiales Eigenwertproblem (PEP) mit den komplexen Eigenwerten ω als

$$(\omega^4 A_4 + \omega^3 A_3 + \omega^2 A_2 + \omega A_1 + A_0) \vec{E} = 0 \quad (3.13)$$

darstellt. Dabei sind die Koeffizientenoperatoren A_i

$$\begin{aligned} A_{cc} &= \text{rot} \frac{1}{\mu} \text{rot}, & A_0 &= -\alpha_0 A_{cc}, & A_1 &= -j\alpha_1 A_{cc}, \\ A_2 &= \varepsilon_0(\varepsilon_\infty \alpha_0 + \beta_0) + A_{cc}, & A_3 &= j\varepsilon_0(\alpha_1 \varepsilon_\infty + \beta_1), & A_4 &= -\varepsilon_0 \varepsilon_\infty. \end{aligned} \quad (3.14)$$

Mit der Methode der finiten Integration kann dieses PEP direkt in die diskrete Form für die elektrischen Gitterspannungen überführt werden. Für die Koeffizienten $\varepsilon_0, \varepsilon_\infty, \alpha_0, \alpha_1, \beta_0, \beta_1$ des dispersiven Modells aus (3.10) werden jeweils Diagonalmatrizen der Dimension $3n_p$ eingeführt, die den entsprechenden Rechengitterpunkten die Koeffizienten zuweisen

$$\varepsilon_0 \rightarrow \mathbf{M}_{\varepsilon_0}, \quad \alpha_0 \rightarrow \mathbf{M}_{\alpha_0}, \quad \alpha_1 \rightarrow \mathbf{M}_{\alpha_1}, \quad \varepsilon_\infty \rightarrow \mathbf{M}_{\varepsilon_\infty}, \quad \beta_0 \rightarrow \mathbf{M}_{\beta_0}, \quad \beta_1 \rightarrow \mathbf{M}_{\beta_1}, \quad (3.15)$$

so dass (3.10) in die Form

$$\mathbf{M}_\varepsilon(\omega) = \mathbf{M}_{\varepsilon_0} \left(\mathbf{M}_{\varepsilon_\infty} + (\mathbf{M}_{\beta_0} + j\omega \mathbf{M}_{\beta_1}) (\mathbf{M}_{\alpha_0} + j\omega \mathbf{M}_{\alpha_1} - \omega^2 \mathbf{I})^{-1} \right) \quad (3.16)$$

übergeht. Das diskrete PEP ist vom Matrixpolynomgrad vier und lautet

$$\Psi(\omega) \vec{e} = (\omega^4 \mathbf{A}_4 + \omega^3 \mathbf{A}_3 + \omega^2 \mathbf{A}_2 + \omega \mathbf{A}_1 + \mathbf{A}_0) \vec{e} = 0 \quad \text{mit } \omega \in \mathbb{C} \text{ und } \varphi \in \mathbb{R}, \quad (3.17)$$

wobei die Koeffizientenmatrizen \mathbf{A}_i aus den folgenden Parametern und Operatoren bestehen

$$\mathbf{A}_{cc,per} = \mathbf{L}_\varphi^H \tilde{\mathbf{C}} \mathbf{M}_{\mu^{-1}} \mathbf{C} \mathbf{L}_\varphi, \quad \mathbf{A}_0 = -\mathbf{M}_{\alpha_0} \mathbf{A}_{cc,per}, \quad (3.18)$$

$$\mathbf{A}_1 = -j \mathbf{M}_{\alpha_1} \mathbf{A}_{cc,per}, \quad \mathbf{A}_2 = \mathbf{M}_{\varepsilon_0} (\mathbf{M}_{\varepsilon_\infty} \mathbf{M}_{\alpha_0} + \mathbf{M}_{\beta_0}) + \mathbf{A}_{cc,per}, \quad (3.19)$$

$$\mathbf{A}_3 = j \mathbf{M}_{\varepsilon_0} (\mathbf{M}_{\alpha_1} \mathbf{M}_{\varepsilon_\infty} + \mathbf{M}_{\beta_1}), \quad \mathbf{A}_4 = -\mathbf{M}_{\varepsilon_0} \mathbf{M}_{\varepsilon_\infty}. \quad (3.20)$$

Bei beliebig inhomogenen Materialverteilungen enthält $\mathbf{M}_{\varepsilon_\infty}$ alle im Rechengebiet auftretenden relativen Permittivitäten, während $\mathbf{M}_{\alpha_0}, \mathbf{M}_{\beta_0}, \mathbf{M}_{\alpha_1}$ und \mathbf{M}_{β_1} nur in Regionen dispersiver Permittivitäten von Null verschieden sind. Dabei ist an Materialübergängen die flächengewichtete Mittelung der Permittivität entsprechend (2.84) unverändert anzuwenden. Die periodische Randbedingung ist in $\mathbf{A}_{cc,per}$ analog zu (3.8) eingearbeitet, wobei dort alle halbierten dualen Flächen, die in die Berechnung der gemittelten Permittivität am unabhängigen periodischen Rand eingehen, entsprechend Abschnitt 2.2.4.2 verdoppelt werden müssen. Sind im Rechengebiet keine dispersiven Permittivitäten vorhanden, fallen die Matrizen $\mathbf{A}_0, \mathbf{A}_1$ und \mathbf{A}_3 weg. Division durch die quadrierte Frequenz ω^2 erzeugt in diesem Fall Gleichung (3.8). Das resultierende PEP (3.17) ist komplex und entsprechende Lösungsstrategien werden in Kapitel 4 diskutiert. Aus den komplexen Eigenwerten ω lässt sich abschließend entsprechend (2.49) die Güte berechnen.

3.2.3. Verlustbehaftete Elementarzellen, räumliches Abklingverhalten

Im vorherigen Abschnitt 3.2.2 wird die unbekannte Kreisfrequenz $\omega \in \mathbb{C}$ durch den Ansatz eines bekannten Phasenvorschubs $\varphi \in \mathbb{R}$ berechnet. Im Folgenden wird der umgekehrte Fall

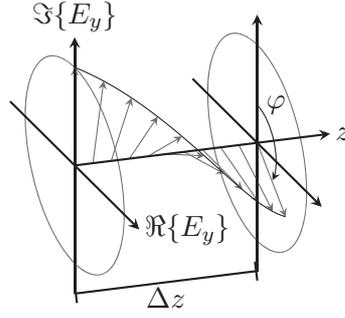


Abbildung 3.2.: Veranschaulichung der Propagation einer Welle \vec{E} durch ein verlustbehaftetes Medium. Das Abklingen über der Länge Δz kann als komplexer Phasenverschiebung modelliert werden (vgl. Abbildung 3.1c).

betrachtet, in dem die Frequenz $\omega \in \mathbb{R}$ vorgegeben wird, woraus sich der komplexe Phasenverschiebung $\varphi \in \mathbb{C}$ berechnet. Durch die im Material einer Elementarzelle vorhandenen Leitfähigkeiten entstehen OHM'sche Verluste, so dass mit $\varphi \in \mathbb{C}$ das Abklingen der Amplitude der Feldverteilung über der Länge Δz modelliert wird. Abbildung 3.2 veranschaulicht, dass die Amplitude einer einfallenden Welle neben dem Phasenverschiebung über Δz auch eine Dämpfung erfährt. Zur numerischen Beschreibung wird Gleichung (3.8) herangezogen, wobei $\mathbf{M}_{\varepsilon, \text{per}}(\omega)$ in diesem Fall einer Frequenzabhängigkeit nach (3.10) unterliegen darf, die zu Beginn der Berechnung evaluiert wird

$$\mathbf{L}_{\varphi}^{(q)H} \mathbf{C}^T \mathbf{M}_{\mu^{-1}} \mathbf{C} \mathbf{L}_{\varphi}^{(q)} \hat{\mathbf{e}} = \omega^2 \mathbf{M}_{\varepsilon, \text{per}}(\omega) \hat{\mathbf{e}} \quad \text{mit } \omega \in \mathbb{R} \text{ und } \varphi \in \mathbb{C}. \quad (3.21)$$

Hierbei ist der Operator $\mathbf{L}_{\varphi}^{(q)}$ im Gegensatz zu \mathbf{L}_{φ} aus (2.98) als quadratische Matrix angesetzt, die Nulleinträge auf der Hauptdiagonalen für abhängige Kanten umfasst. So lässt sich $\mathbf{L}_{\varphi}^{(q)}$ in

$$\mathbf{L}_{\varphi}^{(q)} = \mathbf{I}_L + e^{j\varphi} \mathbf{N}_{\text{per}} \quad (3.22)$$

zerlegen, wobei \mathbf{I}_L eine Einheitsmatrix mit oben genannten Nulleinträgen auf der Hauptdiagonalen für abhängige Kanten ist. \mathbf{N}_{per} besteht aus einzelnen Eins-Einträgen jenseits der Hauptdiagonalen, die abhängige und unabhängige Komponenten (vgl. Abschnitt 2.2.4) durch den skalaren Phasenfaktor $e^{j\varphi}$ verknüpfen. Durch Einsetzen von Gleichung (3.22) in (3.21) folgt

$$((\mathbf{I}_L + e^{-j\varphi} \mathbf{N}_{\text{per}}^T) \mathbf{A}_{\text{cc}} (\mathbf{I}_L + e^{j\varphi} \mathbf{N}_{\text{per}}) - \omega^2 \mathbf{M}_{\varepsilon, \text{per}}(\omega)) \hat{\mathbf{e}} = 0 \quad \text{mit } \mathbf{A}_{\text{cc}} = \tilde{\mathbf{C}} \mathbf{M}_{\mu^{-1}} \mathbf{C}. \quad (3.23)$$

Nach dem Ausmultiplizieren ergibt sich

$$\underbrace{(e^{j\varphi} \mathbf{I}_L \mathbf{A}_{\text{cc}} \mathbf{N}_{\text{per}})}_{\lambda^1 \mathbf{A}_1} + \underbrace{\mathbf{I}_L \mathbf{A}_{\text{cc}} \mathbf{I}_L + \mathbf{N}_{\text{per}}^T \mathbf{A}_{\text{cc}} \mathbf{N}_{\text{per}} - \omega^2 \mathbf{M}_{\varepsilon, \text{per}}(\omega)}_{\lambda^0 \mathbf{A}_0} + \underbrace{e^{-j\varphi} \mathbf{N}_{\text{per}}^T \mathbf{A}_{\text{cc}} \mathbf{I}_L}_{\lambda^{-1} \mathbf{A}_{-1}} \hat{\mathbf{e}} = 0, \quad (3.24)$$

was sich mit der Schreibweise $\lambda = e^{j\varphi}$ und den Koeffizientenmatrizen \mathbf{A}_i als polynomiales Eigenwertproblem mit Matrixpolynomgrad zwei in λ darstellt. Das PEP (3.24) ist komplexwertig und adäquate Lösungsstrategien werden in Kapitel 4 diskutiert. Aus den komplexen Eigenwerten $\lambda = e^{j\varphi}$ lassen sich abschließend der Phasenverschiebung $\varphi(\omega) = \angle(\lambda)$ und die Dämpfungskonstante aus Gleichung (2.55) als $\alpha = \text{Re}\{\ln(\lambda)\}/\Delta z$ bestimmen.

3.3. Berechnung von Ausbreitungskonstanten in längshomogenen Wellenleitern

Die Darstellung zur Berechnung von Ausbreitungskonstanten in längshomogenen Wellenleitern lehnt sich an die Herleitungen in [21, 24, 26] an.

Die Längshomogenität der Wellenleiterstrukturen erlaubt die Betrachtung lediglich der Querschnittsfläche. In jedem Rechengitterpunkt der zweidimensionalen Querschnittsfläche einer Wellenleiterstruktur haben die Zustandsgrößen nach wie vor drei verschiedene räumliche Komponenten entsprechend (2.77). Jedoch wird zur Lösung der MAXWELL'schen Gleichungen die partielle Ableitung in Ausbreitungsrichtung (hier z) benötigt. Die Feldabhängigkeit in Ausbreitungsrichtung ist nach (2.51) gegeben, was eine kontinuierlich-analytische Betrachtung des diskreten partiellen Differenzationsoperator \mathbf{P}_z erlaubt. In der Querschnittsebene $z = z_0$ gilt

$$\mathbf{P}_z \widehat{\mathbf{e}} \rightarrow (e^{-jk_z \Delta z} - 1) \widehat{\mathbf{e}}|_{z_0} \quad (3.25)$$

$$= (-jk_z \Delta z) \widehat{\mathbf{e}}|_{z_0} + \mathcal{O}(\Delta z^2). \quad (3.26)$$

Für kleine Gitterschrittweiten Δz kann der Operator $\mathbf{P}_z = -jk_z \Delta z \mathbf{I}$ in \mathbf{C} und $\widetilde{\mathbf{S}}$ aus (2.82) und (2.83) ersetzt werden und ist darüber hinaus mit $\mathbf{P}_z^{-1} = \frac{j}{k_z \Delta z} \mathbf{I}$ trivial invertierbar. Da der Wellenleiterquerschnitt keine freien Ladungen enthält, gilt

$$\widetilde{\mathbf{S}} \widehat{\mathbf{d}} = 0 \Rightarrow -\mathbf{P}_x^T \widehat{\mathbf{d}}_x - \mathbf{P}_y^T \widehat{\mathbf{d}}_y - jk_z \Delta z \widehat{\mathbf{d}}_z = 0 \quad (3.27)$$

$$\widehat{\mathbf{d}}_z = \frac{1}{jk_z \Delta z} (-\mathbf{P}_x^T \widehat{\mathbf{d}}_x - \mathbf{P}_y^T \widehat{\mathbf{d}}_y). \quad (3.28)$$

Damit lassen sich die zum Wellenleiterquerschnitt normalen Komponenten $\widehat{\mathbf{d}}_z$ durch die tangentialen Komponenten $\widehat{\mathbf{d}}_x, \widehat{\mathbf{d}}_y$ bestimmen, wodurch sich die Anzahl der Freiheitsgrade auf 2/3 der ursprünglichen Anzahl reduziert. Desweiteren existieren in dieser Formulierung keine statischen Wellenleitermoden, sofern es sich nicht um TEM-Wellenleiter handelt. Für die Eigenwertgleichung der Tangentialkomponenten $\widehat{\mathbf{d}}_t$ gilt durch (3.28)

$$\begin{pmatrix} \mathbf{I} & 0 & 0 \\ 0 & \mathbf{I} & 0 \end{pmatrix} \widetilde{\mathbf{C}}_{k_z} \mathbf{M}_{\mu^{-1}} \mathbf{C}_{k_z} \mathbf{M}_{\varepsilon^{-1}} \begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{I} \\ -\mathbf{P}_x^T / jk_z \Delta z & -\mathbf{P}_y^T / jk_z \Delta z \end{pmatrix} \widehat{\mathbf{d}}_t = \omega^2 \widehat{\mathbf{d}}_t. \quad (3.29)$$

In dieser Schreibweise enthalten die Curl-Operatoren $\widetilde{\mathbf{C}}_{k_z}$ und \mathbf{C}_{k_z} den \mathbf{P}_z -Operator aus (3.25). Auflösen nach k_z liefert die Form

$$(\mathbf{A} - (k_z \Delta z)^2 \mathbf{B} - \omega^2 \mathbf{I}) \widehat{\mathbf{d}}_t = 0. \quad (3.30)$$

Für die Matrizen \mathbf{A} und \mathbf{B} gilt schließlich

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \quad \text{und} \quad \mathbf{B} = \begin{pmatrix} \mathbf{B}_{11} & 0 \\ 0 & \mathbf{B}_{22} \end{pmatrix} \quad (3.31)$$

mit

$$\mathbf{A}_{11} = \mathbf{M}_{\mu^{-1}y} \mathbf{P}_x \mathbf{M}_{\varepsilon^{-1}z} \mathbf{P}_x^T + \mathbf{P}_y^T \mathbf{M}_{\mu^{-1}z} \mathbf{P}_y \mathbf{M}_{\varepsilon^{-1}x}, \quad (3.32)$$

$$\mathbf{A}_{12} = \mathbf{M}_{\mu^{-1}y} \mathbf{P}_x \mathbf{M}_{\varepsilon^{-1}z} \mathbf{P}_y^T - \mathbf{P}_y^T \mathbf{M}_{\mu^{-1}z} \mathbf{P}_x \mathbf{M}_{\varepsilon^{-1}y}, \quad (3.33)$$

$$\mathbf{A}_{21} = \mathbf{M}_{\mu^{-1}x} \mathbf{P}_y \mathbf{M}_{\varepsilon^{-1}z} \mathbf{P}_x^T - \mathbf{P}_x^T \mathbf{M}_{\mu^{-1}z} \mathbf{P}_y \mathbf{M}_{\varepsilon^{-1}x}, \quad (3.34)$$

$$\mathbf{A}_{22} = \mathbf{M}_{\mu^{-1}x} \mathbf{P}_y \mathbf{M}_{\varepsilon^{-1}z} \mathbf{P}_y^T + \mathbf{P}_x^T \mathbf{M}_{\mu^{-1}z} \mathbf{P}_x \mathbf{M}_{\varepsilon^{-1}y} \quad (3.35)$$

und den Diagonalmatrizen

$$\mathbf{B}_{11} = \mathbf{M}_{\mu^{-1}y} \mathbf{M}_{\varepsilon^{-1}x}, \quad (3.36)$$

$$\mathbf{B}_{22} = \mathbf{M}_{\mu^{-1}x} \mathbf{M}_{\varepsilon^{-1}y}. \quad (3.37)$$

Die Abhängigkeit von $(\Delta z)^2$ lässt sich in Gleichung (3.30) herauskürzen, da mit denen in Gleichung (2.88) eingeführten Mittelungen der Materialparameter sowohl $\mathbf{M}_{\mu^{-1}y} \mathbf{M}_{\varepsilon^{-1}x} \propto 1/(\Delta z)^2$ als auch $\mathbf{M}_{\mu^{-1}x} \mathbf{M}_{\varepsilon^{-1}y} \propto 1/(\Delta z)^2$ in (3.36) gilt.

Gleichung (3.30) kann als Eigenwertproblem für bekannte Ausbreitungskonstanten k_z und unbekannte Kreisfrequenzen ω formuliert werden. Der hier betrachtete umgekehrte Fall führt auf das Eigenwertproblem

$$\mathbf{B}^{-1}(\mathbf{A} - \omega^2 \mathbf{I}) \widehat{\mathbf{d}}_t = k_z^2 \widehat{\mathbf{d}}_t. \quad (3.38)$$

3.3.1. Wellenleiter mit transversal transparenten Rändern

Bei Wellenleitern mit transversal transparenter Berandung kann wieder das Verfahren der PML aus Abschnitt 2.2.4.4 herangezogen werden. Wie in (3.40) für den Fall dreidimensionaler Strukturen weisen auch die Materialmatrizen für Wellenleiterquerschnitte eine Frequenzabhängigkeit auf. Bei der Berechnung von Ausbreitungskonstanten für gegebene Frequenzen stellt das jedoch kein Problem dar, da die Auswertung $\mathbf{M}_{\varepsilon^{-1}}(\omega)$ und $\mathbf{M}_{\mu^{-1}}(\omega)$ im ersten Schritt erfolgt. Das resultierende Eigenwertproblem wird komplex und ist entsprechend aufwendiger zu lösen.

3.4. Eigenwerte isolierter Strukturen im Freiraum

Die Simulation von Strukturen im Freiraum erfordert im Allgemeinen eine adäquate Berandung des Rechengebiets, welche transparent bezüglich der Propagation abgestrahlter Wellen sein muss. Die in Abschnitt 2.2.4.4 eingeführte transparente Randbedingung der *perfectly matched layers*, führt durch komplexe Koordinatenstreckung eine erweiterte Frequenzabhängigkeit in die Materialmatrizen \mathbf{M}_ε und \mathbf{M}_μ ein. Die Modellierung nach Abschnitt 2.2.4.4 führt auf ein Eigenwertproblem der Form

$$\widetilde{\mathbf{C}} \mathbf{M}_{\mu^{-1}}(\omega) \mathbf{C} \widehat{\mathbf{e}} = \omega^2 \mathbf{M}_\varepsilon(\omega) \widehat{\mathbf{e}}, \quad (3.39)$$

mit komplexen, frequenzabhängigen Materialmatrizen sowie der komplexen Eigenfrequenz ω^2 . Zur genaueren Untersuchung wird ein Dämpfungsprofil (2.105) in Gleichung (2.102) eingesetzt und mittels Gleichung (2.108) die Materialparameter der PML nach Gleichung (2.113) bestimmt. In FIT-Diskretisierung lautet das Durchflutungsgesetz

$$\widetilde{\mathbf{C}} \widehat{\mathbf{h}} = j\omega \underbrace{\begin{pmatrix} \left(\mathbf{I} + \mathbf{M}_{\sigma,x} \frac{1}{j\omega} \right)^{-1} \mathbf{M}_{\varepsilon,xx} \left(\mathbf{I} + (\mathbf{M}_{\sigma,y} + \mathbf{M}_{\sigma,z}) \frac{1}{j\omega} - \mathbf{M}_{\sigma,y} \mathbf{M}_{\sigma,z} \frac{1}{\omega^2} \right) \\ \left(\mathbf{I} + \mathbf{M}_{\sigma,y} \frac{1}{j\omega} \right)^{-1} \mathbf{M}_{\varepsilon,yy} \left(\mathbf{I} + (\mathbf{M}_{\sigma,x} + \mathbf{M}_{\sigma,z}) \frac{1}{j\omega} - \mathbf{M}_{\sigma,x} \mathbf{M}_{\sigma,z} \frac{1}{\omega^2} \right) \\ \left(\mathbf{I} + \mathbf{M}_{\sigma,z} \frac{1}{j\omega} \right)^{-1} \mathbf{M}_{\varepsilon,zz} \left(\mathbf{I} + (\mathbf{M}_{\sigma,x} + \mathbf{M}_{\sigma,y}) \frac{1}{j\omega} - \mathbf{M}_{\sigma,x} \mathbf{M}_{\sigma,y} \frac{1}{\omega^2} \right) \end{pmatrix}}_{\mathbf{M}_\varepsilon(\omega)} \widehat{\mathbf{e}}, \quad (3.40)$$

in dem die diagonalen $n_p \times n_p$ Blockmatrizen $\mathbf{M}_{\sigma,x}$, $\mathbf{M}_{\sigma,y}$ und $\mathbf{M}_{\sigma,z}$ die Werte des Dämpfungsprofils (2.105) innerhalb der PML für die jeweilige Raumrichtung beinhalten. Die Frequenzabhängigkeit ist hierbei maximal quadratisch. Gleichung (3.39) lässt sich mit $\mathbf{M}_\varepsilon(\omega)$ entsprechend (3.40) (und $\mathbf{M}_\mu(\omega)$ mit (2.113) analog dazu) nicht in ein polynomiales Eigenwertproblem umformen, da dazu die Inverse von $\tilde{\mathbf{C}}$ benötigt würde. Eine Formulierung, die dennoch zu einem PEP zweiter Ordnung führt, lässt sich als

$$\left[\begin{pmatrix} \mathbf{C} & 0 \\ 0 & \tilde{\mathbf{C}} \end{pmatrix} - j\omega \begin{pmatrix} 0 & -\mathbf{M}_\mu(\omega) \\ \mathbf{M}_\varepsilon(\omega) & 0 \end{pmatrix} \right] \begin{pmatrix} \hat{\mathbf{e}} \\ \hat{\mathbf{h}} \end{pmatrix} = 0 \quad (3.41)$$

herleiten, wobei $\mathbf{M}_\varepsilon(\omega)$ und $\mathbf{M}_\mu(\omega)$ auszuschreiben sind, wie in (3.40) gezeigt ist.

In der Praxis ist die Lösung von (3.41) verzichtbar, da für die hier betrachteten Problemstellungen iterative Löser zum Einsatz kommen, die Eigenwerte nahe einem Schätzwert ω_e liefern. Für kleine Abweichungen von ω_e zum Eigenwert ω genügt es daher

$$\tilde{\mathbf{C}}\mathbf{M}_{\mu^{-1}}(\omega_e)\mathbf{C}\hat{\mathbf{e}} = \omega^2\mathbf{M}_\varepsilon(\omega_e)\hat{\mathbf{e}} \quad (3.42)$$

für ω_e auszuwerten und dann lineare Eigenwertlöser zu verwenden. Der durch diese Linearisierung eingeführte Fehler wird in Abschnitt 5.4.3 am Beispiel untersucht. Dennoch ist der Rechenaufwand, den die komplexwertigen Komponenten in der PML erzeugen, nicht zu unterschätzen, da die resultierenden Systeme komplex und unsymmetrisch werden.

3.5. Eigenwerte skalarer 2D Problemstellungen mit FLAME

Für die Simulation mit einer FLAME-Diskretisierung werden in dieser Arbeit Problemstellungen betrachtet, die in einer Koordinatenrichtung als unendlich ausgedehnt angenommen werden können und sich somit zweidimensional in der x - y -Ebene modellieren lassen.

In der vektoriellen HELMHOLTZ-Gleichung (2.16) wird nur eine einzige Vektorkomponente zugelassen, so dass aus

$$\Delta E_z + k^2 E_z = 0 \quad \text{und} \quad \Delta H_z + k^2 H_z = 0 \quad \text{mit} \quad k = \frac{\omega}{c_0} = \omega\sqrt{\mu\varepsilon} \quad (3.43)$$

die skalaren HELMHOLTZ-Gleichungen für E_z bzw. H_z folgen. Mit diesem Ansatz lassen sich beispielsweise Transmissions- und Reflexionseigenschaften ebener Trennflächen oder die Streueigenschaften von unendlich ausgedehnten dielektrischen Zylindern aus Abschnitt 2.1.2.2 modellieren.

Die in FLAME angesetzten Basisfunktionen müssen die skalare HELMHOLTZ-Gleichung erfüllen und müssen darüber hinaus so gewählt werden, dass sie lokal eine sinnvolle Approximation der erwarteten Lösung darstellen. Der Übergang zwischen Regionen unterschiedlicher Basisfunktionen verursacht dabei den Approximationsfehler aus Abschnitt 2.3.3.1. Für 3×3 -Differenzensterne werden typischerweise acht Basisfunktionen gewählt, um sicherzustellen, dass sich das lokale Differenzenschema $\mathbf{s}^{(i)}$ aus dem Kern der Knotenmatrix $\mathbf{s}^{(i)} \in \ker \mathbf{N}^{(i)\text{T}}$ auf dem Teilgebiet $\Omega^{(i)}$ berechnen lässt [37].

Als acht Basisfunktionen im homogenen Dielektrikum ε_r eignen sich ebene Wellen mit jeweils um 45° versetzten Ausbreitungsrichtungen

$$\psi_m^{pw}(\omega, x, y) = e^{j\frac{\omega\sqrt{\varepsilon_r}}{c_0}(\cos(m\frac{\pi}{4})x + \sin(m\frac{\pi}{4})y)} \quad m \in \{0, \dots, 7\}. \quad (3.44)$$

Für die Streuung in der Nähe zylindrischer Streukörper werden die in Abschnitt 2.1.2.2 eingeführten Zylinderfunktionen der Ordnung m als Basisfunktionen verwendet

$$\psi_m^{cyl}(\omega, \varrho, \phi) = j^{-m} J_m(\omega/c_0\varrho) e^{jm\phi} \quad m \in \mathbb{Z}. \quad (3.45)$$

Während die lokalen Differenzenschemata $\mathbf{s}^{(i)}$ für den Ansatz von Zylinderfunktionen nach Gleichung (3.45) numerisch zur Laufzeit berechnet werden, lässt sich für die Verwendung ebener Wellen als Basisfunktionen auf dem 3×3 -Differenzenstern ein analytischer Ausdruck berechnen.

3.5.1. Struktur der Systemmatrix

Sind die Basisfunktionen ebene Wellen im homogenen Dielektrikum nach Gleichung (3.44), besteht $\mathbf{s}^{(i)} \in \ker \mathbf{N}^{(i)T}$ entsprechend (2.122) auf dem äquidistanten 3×3 -Differenzenstern auf dem Teilgebiet $\Omega^{(i)}$ aus folgenden Komponenten [36]:

$$\begin{aligned} s_1 &= (\mathbf{e}_{\frac{1}{2}} + 1) \frac{\mathbf{e}_{\frac{1}{2}}\mathbf{e}_1 + 2\mathbf{e}_{\frac{1}{2}}\mathbf{e}_0 - 4\mathbf{e}_{-\frac{1}{2}}\mathbf{e}_1 + \mathbf{e}_{\frac{1}{2}} - 4\mathbf{e}_{-\frac{1}{2}} + \mathbf{e}_1 + 2\mathbf{e}_0 + 1}{(\mathbf{e}_0 - 1)^2(\mathbf{e}_{-\frac{1}{2}} - 1)^4} \\ s_{2\dots 5} &= -\frac{\mathbf{e}_{\frac{3}{2}}\mathbf{e}_0 - 2\mathbf{e}_{\frac{1}{2}}\mathbf{e}_1 + 2\mathbf{e}_{\frac{1}{2}}\mathbf{e}_0 - 2\mathbf{e}_{\frac{1}{2}} + \mathbf{e}_0}{(\mathbf{e}_0 - 1)^2(\mathbf{e}_{-\frac{1}{2}} - 1)^4} \\ s_{6\dots 9} &= \mathbf{e}_{-\frac{1}{2}} \frac{2\mathbf{e}_{\frac{1}{2}}\mathbf{e}_0 - \mathbf{e}_{-\frac{1}{2}}\mathbf{e}_1 - 2\mathbf{e}_{-\frac{1}{2}}\mathbf{e}_0 - \mathbf{e}_{-\frac{1}{2}} + 2\mathbf{e}_0}{(\mathbf{e}_0 - 1)^2(\mathbf{e}_{-\frac{1}{2}} - 1)^4}. \end{aligned} \quad (3.46)$$

Dabei gilt für die Schreibweise $\mathbf{e}_\varsigma = \exp(2\varsigma jhk)$, $\varsigma \in \{-\frac{1}{2}, 0, \frac{1}{2}, 1, \frac{3}{2}\}$ mit der Rechengitterschrittweite h und der Wellenzahl $k = \frac{\omega\sqrt{\varepsilon_r}}{c}$. s_1 ist dem zentralen und $s_{2\dots 5}$ sind den seitlichen Knoten analog Abbildung 2.10 aus Abschnitt 2.3.4 zugeordnet. Den Eckknoten im 3×3 -Differenzenstern sind die Werte $s_{6\dots 9}$ zugeordnet. Da (3.46) unmittelbar die Einträge der Systemmatrix darstellt, ist ersichtlich, dass die Wellenzahl k als Argument in den Exponentialfunktionen erhalten bleibt.

3.5.2. Eigenwertformulierung in FLAME

Wie in Abschnitt 2.3.4 motiviert wird, ergibt sich die Systemmatrix in FLAME nicht aus dem Produkt verschiedener Operatoren, sondern wird zeilenweise aus den Komponenten des Vektors \mathbf{s} aus (2.122) aufgebaut, welcher den Kern der Knotenmatrix $\mathbf{N}^{(i)T}$ nach Gleichung (2.119) der gewählten Basisfunktionen $\{\psi_m^{(i)}(\omega)\}$ darstellt. Die Bestimmung von Eigenwerten kann in FLAME als

$$\mathbf{A}_{\text{FLAME}}(\omega_e)\mathbf{x}_e = \mathbf{0} \quad (3.47)$$

formuliert werden. Hierbei ist das Eigenpaar (ω_e, \mathbf{x}_e) gesucht. Der Hauptunterschied zu den nichtlinearen Eigenwertproblemen aus der FIT-Diskretisierung aus Gleichungen (3.17) und

(3.24) besteht darin, dass sich für die FLAME–Eigenwertformulierung (3.47) der Eigenwert nicht mehr als Faktor vor eine Koeffizientenmatrix ziehen lässt. In der Praxis wird (2.122) numerisch gelöst, so dass die Frequenzabhängigkeit in der Systemmatrix $\mathbf{A}_{\text{FLAME}}$ analytisch nicht mehr auftaucht. Im folgenden Kapitel werden Methoden vorgestellt, mit denen sich die FLAME–Eigenwertformulierung (3.47) iterativ lösen lässt.

4. Numerische Methoden zur Berechnung algebraischer Eigenwertprobleme

Dieses Kapitel stellt neben den benötigten Grundlagen zur numerischen Berechnung algebraischer Eigenwertprobleme die Löser für lineare und nichtlineare Eigenwertprobleme vor. Den Schwerpunkt bildet dabei das JACOBI–DAVIDSON–Verfahren. Abschnitte 4.1 bis 4.3 rekapitulieren überwiegend bekannte Ergebnisse zur Berechnung von Eigenwertproblemen mit einem Fokus auf Matrixpolynomen. Abschnitt 4.4 gibt einen Überblick über die verwendeten Löser für lineare Gleichungssysteme, sowie die Umsortierung von FIT–Matrizen für die ILU–Zerlegung. In den Abschnitten 4.5 bis 4.7 liegt der Schwerpunkt auf Erweiterungen der zuvor vorgestellten Algorithmen, die im Rahmen dieser Arbeit entstanden sind.

Der Weg zum Begriff des Eigenwertproblems

Die Evolution im 19. Jahrhundert hin zum Begriff des Eigenwertproblems wird im Folgenden anhand einer sinngemäßen Übersetzung von [89] dargestellt¹.

Eigenwerte werden oft im Kontext der linearen Algebra oder Matrixtheorie eingeführt. Historisch betrachtet hatten jedoch die quadratischen Formen ebenso wie die Theorie der Differenzialgleichungen maßgeblichen Einfluss in der Entstehung.

Bereits EULER² untersucht im 18. Jahrhundert die Rotationsbewegung starrer Körper und entdeckt dabei die Relevanz der Hauptträgheitsachsen. LAGRANGE³ erkennt in den Hauptträgheitsachsen das, was heute als Eigenvektoren der Trägheitsmatrix bezeichnet wird. Im frühen 19. Jahrhundert untersucht CAUCHY⁴ quadratische Formen und verallgemeinert die Ergebnisse von EULER und LAGRANGE. CAUCHY nennt die Eigenwerte seinerzeit noch *charakteristische Wurzel* und prägt so den bis heute bestehenden Begriff des *charakteristischen Polynoms*, dessen Lösung die Eigenwerte sind.

Zur Lösung der Wärmeleitungsgleichung verwendet FOURIER⁵ 1822 die Ergebnisse von LAPLACE und LAGRANGE, um durch Separation der Variablen zum Ziel zu kommen. STURM⁶ entwickelt FOURIER's Ideen weiter und macht CAUCHY auf sie aufmerksam, der wiederum zu dem Schluss kommt, dass reelle symmetrische Matrizen auch reelle Eigenwerte haben. 1855 verallgemeinert HERMITE⁷ dies auf symmetrisch konjugiert-komplexe Matrizen, woraus sich der Begriff der *hermiteschen* Matrix ableitet. Zu dieser Zeit beweist BRIOSCHI⁸, dass die Eigen-

¹Die Zusammenstellung historischer Werke soll nicht den Eindruck erwecken, es seien die ersten oder einzigen im Kontext relevanten Fachpublikationen zur damaligen Zeit gewesen.

²Leonhard Euler (1707-1783), schweizerischer Mathematiker.

³Joseph-Louis de Lagrange (1736-1813), italienischer Mathematiker und Astronom.

⁴Augustin Louis Cauchy (1789-1857), französischer Mathematiker.

⁵Jean Baptiste Joseph Fourier (1768-1830), französischer Mathematiker und Physiker.

⁶Jacques Charles François Sturm (1803-1855), französischer Mathematiker.

⁷Charles Hermite (1822-1901), französischer Mathematiker.

⁸Francesco Brioschi (1824-1894), italienischer Mathematiker.

werte orthogonaler Matrizen auf dem Einheitskreis der komplexen Ebene liegen und CLEBSCH⁹ findet das entsprechende Ergebnis für schief-symmetrische Matrizen. WEIERSTRASS¹⁰ ergänzt die LAPLACE'sche Stabilitätstheorie um die wichtige Erkenntnis, dass defekte Matrizen Instabilität verursachen können.

Etwa parallel dazu studieren LIOUVILLE¹¹ und STURM Eigenwertprobleme linearer Differenzialgleichungen, woraus schließlich die STURM-LIOUVILLE-Theorie erwächst. SCHWARZ¹² betrachtet Ende des 19. Jahrhunderts die Eigenwerte der LAPLACE-Gleichung in höheren Dimensionen, während POINCARÉ¹³ dies für die POISSON¹⁴-Gleichung einige Jahre später durchführt.

Eines der ersten iterativen Verfahren für die Berechnung von Eigenwerten fällt ebenso in diese Epoche und stammt von JACOBI¹⁵ aus dem Jahr 1846 [66]. Ursprüngliche Intention von JACOBI war die Vereinfachung des Lösens von linearen Gleichungssystemen. Die JACOBI-Methode reduziert eine reelle symmetrische Matrix auf Diagonalform durch eine Folge ebener Rotationen, die heute GIVENS¹⁶-Rotationen heißen, Details sind im Anhang A.3 beschrieben.

Über Evolution im 20. Jahrhundert heißt es weiter in [89]:

Zu Beginn des 20. Jahrhunderts analysiert HILBERT¹⁷ die Eigenwerte von Integraloperatoren, indem er die (kontinuierlichen) Operatoren als unendlich dimensionale Matrizen betrachtet. HILBERT wird seit 1904 die Propagation des deutschen Wortes *eigen* zugeschrieben, obwohl er möglicherweise von HELMHOLTZ inspiriert wurde, der einige Jahre zuvor *Eigentöne* schwingender Membrane untersuchte. Die Bezeichnung *eigen* hat sich seitdem in der Mathematik international durchgesetzt, so dass es im Englischen die Begriffe *eigenvalue*, *eigenvector* und *eigenpair* gibt. Einer der ersten numerischen Algorithmen zur Berechnung von Eigenwerten und Eigenvektoren¹⁸ erscheint mit der *Potenzmethode* 1929 von VON MISES¹⁹.

1944 erweitert WIELANDT²⁰ die *Potenzmethode* zur *inversen Iteration*, mit der sich Eigenwerte nahe einer vorgegebenen Zahl berechnen lassen. Beide Verfahren bilden die Einführung im nächsten Abschnitt 4.1.

Zur weiteren historischen Entwicklung der Verfahren zur Eigenwertberechnung während des 20. Jahrhunderts sei an dieser Stelle auf die Publikation [59] verwiesen, die sich sehr ausführlich diesem Abschnitt der Historie widmet. Der historische Abschnitt von Mitte der 1980er bis Mitte der 1990er Jahre in [87] bietet für iterative Methoden großer schwachbesetzter Eigenwertprobleme einen sehr guten Überblick.

⁹Rudolf Friedrich Alfred Clebsch (1833-1872), deutscher Mathematiker.

¹⁰Karl Theodor Wilhelm Weierstraß (1815-1897), deutscher Mathematiker.

¹¹Joseph Liouville (1809-1882), französischer Mathematiker.

¹²Hermann Amandus Schwarz (1843-1921), deutscher Mathematiker.

¹³Jules Henri Poincaré (1854-1912), französischer Mathematiker, theoretischer Physiker und Philosoph.

¹⁴Siméon-Denis Poisson (1781-1842), französischer Physiker und Mathematiker.

¹⁵Carl Gustav Jacob Jacobi (1804-1851), deutscher Mathematiker.

¹⁶James Wallace Givens, Jr. (1910-1993), US-amerikanischer Mathematiker und Informatiker.

¹⁷David Hilbert (1862-1943), deutscher Mathematiker.

¹⁸Das Verfahren von JACOBI zielte ursprünglich auf die Lösung eines linearen Gleichungssystems ab.

¹⁹Richard Edler von Mises (1883-1953), österreichischer Mathematiker.

²⁰Helmut Wielandt (1910-2001), deutscher Mathematiker.

4.1. Einführung

Zum Verständnis von Eigenwerten und Eigenvektoren wird der geometrische Fall betrachtet, dass eine lineare Abbildung (die z.B. durch eine Matrix gegeben ist) Vektoren skalieren kann, ohne die Richtung der Vektoren zu verändern. Da diese Eigenschaft nicht jeder beliebige Vektor mit einer Abbildung verbindet, handelt es sich dabei um besondere Vektoren, die Eigenvektoren der linearen Abbildung. Die Skalierung der Länge eines Vektors lässt sich ebenfalls durch eine Zahl beschreiben, die einem bestimmten Eigenvektor zugeordnet ist. Der Skalierungsfaktor, welcher die gleiche Skalierung wie die Abbildung durchführen kann, heißt Eigenwert. Zusammen bilden Eigenwert und Eigenvektor ein Eigenpaar der linearen Abbildung.

Klassische Lehrbücher zur Lösung von Eigenwertproblemen sind die Werke von WILKINSON²¹ [90] und WATSON [61] sowie das Buch von GOLUB und VAN LOAN [60]. Speziell für große, schwachbesetzte Matrizen gibt es mit den Arbeiten von VAN DER VORST [86] und SAAD [72] gute Nachschlagewerke. Für die praktische Implementierung zeigt [46] viele Vorlagen gängiger Lösungsstrategien. Darüber hinaus existieren zur numerischen Mathematik einführende Werke in deutscher Sprache, wie z.B. [53, 62, 85]. Zunächst werden einige Definitionen und Sätze aus der Eigenwerttheorie wiederholt. Der Aufbau der folgenden Einführung orientiert sich an [68]. Die jeweiligen Beweise sind in der Literatur [53, 60, 62] enthalten.

4.1.1. Klassifizierung von Eigenwertproblemen

Da Eigenwerte nicht zwangsläufig endlich sein müssen, wird zunächst eine „um Unendlich“ erweiterte komplexe Menge definiert.

Definition 4.1.1: Die Erweiterung von \mathbb{C} um ∞ wird definiert als

$$\bar{\mathbb{C}} := \mathbb{C} \cup \{z \cdot \infty \mid z \in \mathbb{C}, |z| = 1\}.$$

Zunächst wird der allgemeinste Typ von Eigenwertproblem eingeführt, der in dieser Arbeit benötigt wird. Im nichtlinearen Eigenwertproblem hängen eine oder mehrere Matrixoperatoren nichtlinear vom Eigenwert λ ab.

Definition 4.1.2 (Nichtlineares Eigenwertproblem): Sei $\Xi : \bar{\mathbb{C}} \rightarrow \bar{\mathbb{C}}^{n \times n}$ eine beliebige nichtlineare Funktion auf der Menge der komplexen quadratischen Matrizen. Das Paar (λ, \mathbf{x}) mit $\lambda \in \bar{\mathbb{C}}$ und nichttrivialem $\mathbf{x} \in \mathbb{C}^n \setminus \mathbf{0}$ wird Eigenpaar von Ξ genannt, wenn

$$\Xi(\lambda)\mathbf{x} = \mathbf{0},$$

mit dem Eigenwert λ und dem zugehörigen rechten Eigenvektor²² \mathbf{x} .

Die Aufgabe, ein Eigenpaar von Ξ zu finden wird als nichtlineares Eigenwertproblem bezeichnet. Eine Unterklasse sind Matrixpolynome, die zwar nichtlinear von Eigenwert λ abhängen, für die aber eine Reihe vorteilhafter Eigenschaften bekannt sind. Vergleichbar mit skalaren Polynomen ist das Matrixpolynom als die Summe von Koeffizientenmatrizen \mathbf{A}_i multipliziert mit der jeweiligen Potenz der skalaren Variablen definiert.

²¹James Hardy Wilkinson (1919-1986), britischer Mathematiker.

²²Wenn im Folgenden von *Eigenvektor* ohne Attribut die Rede ist, ist der *rechte Eigenvektor* \mathbf{x} gemeint.

Definition 4.1.3 (Matrixpolynom): Sei $\mathbf{A}_i \in \mathbb{C}^{n \times n}$, $i = 0, \dots, \ell$, $n \in \mathbb{N}$ und $\lambda \in \mathbb{C}$. Die Funktion $\Psi : \bar{\mathbb{C}} \rightarrow \bar{\mathbb{C}}^{n \times n}$ mit

$$\Psi(\lambda) := \sum_{i=0}^{\ell} \lambda^i \mathbf{A}_i = \mathbf{A}_0 + \lambda \mathbf{A}_1 + \dots + \lambda^{\ell-1} \mathbf{A}_{\ell-1} + \mathbf{A}_{\ell} \lambda^{\ell}$$

heißt Matrixpolynom vom Polynomgrad ℓ . Ψ heißt singular, wenn für die Determinante $\det(\Psi(\lambda)) = 0$ gilt und regulär sonst.

Die Eigenpaare des Matrixpolynoms Ψ sind als Paare $(\lambda_i, \mathbf{x}_i)$ mit $\mathbf{x}_i \neq \mathbf{0}$ definiert, die Ψ auf den Nullvektor $\mathbf{0}$ abbilden.

Definition 4.1.4 (Polynomiales Eigenwertproblem): Sei $\Psi : \bar{\mathbb{C}} \rightarrow \bar{\mathbb{C}}^{n \times n}$ ein Matrixpolynom. Das Paar (λ, \mathbf{x}) mit $\lambda \in \bar{\mathbb{C}}$ und $\mathbf{x} \in \mathbb{C}^n \setminus \mathbf{0}$ wird Eigenpaar von Ψ genannt, wenn

$$\Psi(\lambda) \mathbf{x} = \mathbf{0},$$

mit dem Eigenwert λ und dem zugehörigen rechten Eigenvektor \mathbf{x} .

Die Aufgabe, ein Eigenpaar von Ψ zu finden wird als polynomiales Eigenwertproblem bezeichnet. Der linke Eigenvektor \mathbf{y} ist durch

$$\mathbf{y}^H \Psi(\lambda) = \mathbf{0}$$

definiert und gehört ebenfalls zum Eigenwert λ . Durch die Forderung an den Matrixpolynomgrad $\ell = 1$ entsteht das allgemeine Eigenwertproblem, eine Form des linearen Eigenwertproblems.

Bemerkung 4.1.5 (Allgemeines Eigenwertproblem): Für $\ell = 1$ und $\mathbf{A}_0 = \mathbf{A}$ und $\mathbf{A}_1 = -\mathbf{B}$ ergibt sich aus dem polynomialen Eigenwertproblem das allgemeine Eigenwertproblem

$$\mathbf{A} \mathbf{x} = \lambda \mathbf{B} \mathbf{x}.$$

Wird neben $\ell = 1$ noch zusätzlich gefordert, dass \mathbf{A}_1 den Wert der negativen Identitätsmatrix annimmt $\mathbf{A}_1 = -\mathbf{I}$, ergibt sich das Standardeigenwertproblem für die Matrix \mathbf{A} .

Bemerkung 4.1.6 (Standardeigenwertproblem): Für Matrixpolynomgrad $\ell = 1$ und $\mathbf{A}_0 = \mathbf{A}$ und $\mathbf{A}_1 = -\mathbf{I}$ mit der Identitätsmatrix \mathbf{I} ergibt sich aus dem polynomialen Eigenwertproblem das Standardeigenwertproblem

$$\mathbf{A} \mathbf{x} = \lambda \mathbf{x}.$$

Die Beziehung zwischen allgemeinem Eigenwertproblem und Standardeigenwertproblem ist für reguläre Matrizen leicht nachvollziehbar.

Bemerkung 4.1.7: Ein allgemeines Eigenwertproblem kann als Standardeigenwertproblem ausgedrückt werden, wenn \mathbf{A} oder \mathbf{B} regulär und damit invertierbar ist.

Der Beweis verwendet die Inverse von \mathbf{A} oder \mathbf{B} , die aufgrund der vorausgesetzten Regularität existiert. Somit lassen sich beide Matrizen auf eine Seite der Gleichung bringen.

Polynomiale Eigenwertprobleme nach Definition 4.1.4 können auf allgemeine Eigenwertprobleme nach Bemerkung 4.1.5 linearisiert werden, wobei das linearisierte Eigenwertproblem die Dimension des Ausgangsproblems n multipliziert mit dem höchsten Polynomgrad ℓ erhält [58].

Satz 4.1.8: *Polynomiale Eigenwertprobleme mit Polynomgrad $\ell > 1$ und Dimension $n \times n$ können auf allgemeine Eigenwertprobleme der Dimension $\ell n \times \ell n$ linearisiert werden. Dabei treten ℓn Eigenpaare auf.*

Eine „Linearisierung“²³ $L(\lambda)$ eines polynomialen Eigenwertproblems wird mit der Hilfe regulärer Matrizen gebildet, jedoch gibt es zwei verschiedene Linearisierungen, die mit der Verwendung der Identitätsmatrix \mathbf{I} auskommen [46]. Daraus folgt für das polynomiale Eigenwertproblem nach Definition 4.1.4 mit Polynomgrad ℓ

$$\begin{aligned} \Psi(\lambda)\mathbf{x} &= \sum_{i=0}^{\ell} \lambda^i \mathbf{A}_i \mathbf{x} = \mathbf{0} \\ \Leftrightarrow &\left(\begin{array}{c} \left[\begin{array}{cccc} \mathbf{A}_1 & \mathbf{A}_2 & \mathbf{A}_3 & \dots & \mathbf{A}_\ell \\ \mathbf{I} & & & & \\ \lambda & & \mathbf{I} & & \\ & & & \ddots & \\ & & & & \mathbf{I} \end{array} \right] + \left[\begin{array}{cccc} \mathbf{A}_0 & & & \\ & -\mathbf{I} & & \\ & & -\mathbf{I} & \\ & & & \ddots \\ & & & & -\mathbf{I} \end{array} \right] \end{array} \right) \left(\begin{array}{c} \mathbf{x} \\ \lambda \mathbf{x} \\ \lambda^2 \mathbf{x} \\ \vdots \\ \lambda^{\ell-1} \mathbf{x} \end{array} \right) = \mathbf{0} \quad (4.1) \end{aligned}$$

und

$$\Leftrightarrow \left(\begin{array}{c} \left[\begin{array}{cccc} \mathbf{A}_\ell & & & \\ & \mathbf{I} & & \\ \lambda & & \mathbf{I} & \\ & & & \ddots \\ & & & & \mathbf{I} \end{array} \right] + \left[\begin{array}{cccc} \mathbf{A}_{\ell-1} & \mathbf{A}_{\ell-2} & \mathbf{A}_{\ell-3} & \dots & \mathbf{A}_0 \\ & -\mathbf{I} & & & \\ & & -\mathbf{I} & & \\ & & & \ddots & \\ & & & & -\mathbf{I} \end{array} \right] \end{array} \right) \left(\begin{array}{c} \lambda^{\ell-1} \mathbf{x} \\ \lambda^{\ell-2} \mathbf{x} \\ \lambda^{\ell-3} \mathbf{x} \\ \vdots \\ \mathbf{x} \end{array} \right) = \mathbf{0}. \quad (4.2)$$

Obwohl die beiden Linearisierungen (4.1) und (4.2) äquivalent sind, bestehen Unterschiede in der numerischen Stabilität und Effizienz. Insbesondere für den Fall, dass \mathbf{A}_ℓ regulär ist und eine kleine Konditionszahl besitzt, lässt sich (4.2) in ein Standardeigenwertproblem nach Bemerkung 4.1.6 überführen.

Im Folgenden werden die Abbildungseigenschaften der Koeffizientenmatrizen eines Matrixpolynoms genauer betrachtet. Die Definition von Kern und Bild gilt für Matrizen.

Definition 4.1.9 (Kern und Bild): *Sei $\mathbf{A} \in \mathbb{C}^{n \times n}$. Der Kern von \mathbf{A} ist definiert als*

$$\ker(\mathbf{A}) := \{\mathbf{v} \in \mathbb{C}^n \mid \mathbf{A}\mathbf{v} = \mathbf{0}\}$$

und für den Bildbereich gilt

$$\operatorname{im}(\mathbf{A}) := \{\mathbf{z} \in \mathbb{C}^n \mid \mathbf{z} = \mathbf{A}\mathbf{v}, \mathbf{v} \in \mathbb{C}^n\}.$$

Der Zusammenhang zwischen der Dimension von Kern und Bild ist durch den Dimensionssatz gegeben.

Satz 4.1.10 (Dimensionssatz): *Ist eine lineare Abbildung durch die Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ gegeben, gilt*

$$\dim \mathbb{C}^{n \times n} = n = \dim \operatorname{im}(\mathbf{A}) + \dim \ker(\mathbf{A}).$$

²³„Linearisierung“ ist auf die Form der resultierenden Darstellung bezogen. Der Eigenwert selbst taucht in evtl. höheren Potenzen im Eigenvektor der Linearisierung auf.

In anderer Schreibweise gilt die äquivalente Aussage

$$\dim \mathbb{C}^{n \times n} = n = \text{rg}(\mathbf{A}) + \text{def}(\mathbf{A}),$$

mit dem Defekt der Matrix $\text{def}(\mathbf{A}) := \dim \ker(\mathbf{A})$ und dem Rang der Matrix $\text{rg}(\mathbf{A}) := \dim \text{im}(\mathbf{A})$.

Definition 4.1.11 (Charakteristisches Polynom): Das charakteristische Polynom p_Ψ eines Matrixpolynoms Ψ ist definiert als

$$p_\Psi(\lambda) := \det(\Psi(\lambda)).$$

Für den maximalen Polynomgrad ergibt sich mit Hilfe des Dimensionssatzes 4.1.10:

Bemerkung 4.1.12: Eigenwerte des polynomialen Eigenwertproblems $\Psi(\lambda)\mathbf{x} = \mathbf{0}$ können unendlich sein. Der Polynomgrad $g := \deg p_\Psi$ des charakteristischen Polynoms ist maximal ℓn und die fehlenden $\ell n - g$ Nullstellen werden als unendliche Eigenwerte definiert.

Um zu zeigen, an welcher Stelle unendliche Eigenwerte entstehen können, wird das Ergebnis aus Satz 4.1.8 verwendet, wonach sich polynomiale Eigenwertprobleme stets als allgemeine Eigenwertprobleme $\mathbf{A}\mathbf{x} = \lambda\mathbf{B}\mathbf{x}$ mit $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n \times n}$ aus Bemerkung 4.1.5 ausdrücken lassen. Dazu gelte $\ker(\mathbf{A}) = \{\mathbf{0}\}$ und somit $\text{def}(\mathbf{A}) = 0$. Mit dem Dimensionssatz 4.1.10 folgt \mathbf{B} : $\text{def}(\mathbf{B}) = n - \text{rg}(\mathbf{B})$. Anders ausgedrückt bedeutet dies, dass Matrix \mathbf{A} keine und Matrix \mathbf{B} eine Anzahl von $n - \text{rg}(\mathbf{B})$ Nulleigenwerten besitzt. Sei $\mathbf{0} \neq \mathbf{v} \in \ker(\mathbf{B})$ und $\mathbf{v} \notin \ker(\mathbf{A})$. Dann gilt für das allgemeine Eigenwertproblem mit der Spektraltransformation $v = 1/\lambda$

$$v \underbrace{\mathbf{A}\mathbf{v}}_{\neq \mathbf{0}} = \mathbf{B}\mathbf{v} = \mathbf{0} \quad \Rightarrow \quad 0 = v = \frac{1}{\lambda} \quad \Rightarrow \quad \lambda = \infty.$$

Die unendlichen Eigenwerte treten demnach auf, wenn $\text{def}(\mathbf{B}) > 0$ gilt. Darüber hinaus ist der Sonderfall zu betrachten, in dem die Matrizen \mathbf{A} und \mathbf{B} gemeinsame Kerne besitzen. Aus der vorherigen Betrachtung folgen undefinierte Eigenwerte der Form „0/0“.

Bemerkung 4.1.13: Falls $\ker(\mathbf{A}) \cap \ker(\mathbf{B}) \neq \{\mathbf{0}\}$ werden diese undefinierten Eigenwerte auch unendlich genannt.

Daraus folgt eine wichtige Eigenschaft von Standard eigenwertproblemen.

Bemerkung 4.1.14: Standard eigenwertprobleme haben keine unendlichen Eigenwerte, da $\mathbf{B} = \mathbf{I}$ und $\ker(\mathbf{I}) = \{\mathbf{0}\}$.

4.1.2. Kenngrößen

Im Folgenden werden einige wichtige Kenngrößen eingeführt, die zur Analyse von Eigenwertproblemen notwendig sind.

Definition 4.1.15 (Spektrum): Die Menge der Eigenwerte einer Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ wird Spektrum genannt und mit

$$\sigma(\mathbf{A}) = \{\lambda \in \mathbb{C} \mid \exists \mathbf{x} \neq \mathbf{0} : \mathbf{A}\mathbf{x} = \lambda\mathbf{x}\}$$

bezeichnet.

Definition 4.1.16 (Spektralradius): Der Spektralradius einer Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ ist der Betrag des betragsgrößten Eigenwerts

$$\rho(\mathbf{A}) := \max_i |\lambda_i(\mathbf{A})|.$$

Die Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ kann mehrere identische Eigenwerte haben. Die Vielfachheit eines Eigenwerts als Nullstelle des charakteristischen Polynoms wird als algebraische Vielfachheit bezeichnet.

Definition 4.1.17 (Algebraische Vielfachheit): Die natürliche Zahl n_i heißt algebraische Vielfachheit des Eigenwerts λ_i . Im Fall von $n_i = 1$ wird der zugehörige Eigenwert einfacher Eigenwert genannt.

Die geometrische Vielfachheit des Eigenwerts λ_i ist die maximale Anzahl linear unabhängiger Eigenvektoren, die zu diesem Eigenwert existiert.

Definition 4.1.18 (Geometrische Vielfachheit): Die Dimension

$$m_i := \dim \ker(\mathbf{A} - \lambda_i \mathbf{I}) = n - \text{rg}(\mathbf{A} - \lambda_i \mathbf{I})$$

heißt geometrische Vielfachheit des Eigenwerts λ_i .

Die geometrische Vielfachheit m_i ist stets kleiner oder gleich der algebraischen Vielfachheit n_i jedes Eigenwerts $\lambda_i \in \sigma(\mathbf{A})$, so dass die Beziehung

$$1 \leq m_i \leq n_i$$

gilt. Stimmen geometrische und algebraische Vielfachheit aller Eigenwerte einer Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ überein, bilden die Eigenvektoren \mathbf{x}_i von \mathbf{A} eine Basis des \mathbb{C}^n .

Definition 4.1.19 (Defekter Eigenwert): Ist die geometrische Vielfachheit eines Eigenwerts $\lambda_i \in \sigma(\mathbf{A})$ kleiner als seine algebraische Vielfachheit, gilt $m_i < n_i$ und λ_i wird als defekter Eigenwert bezeichnet.

Eine Matrix mit mindestens einem defekten Eigenwert wird analog zum Dimensionssatz 4.1.10 defekte Matrix genannt. Beim Auftreten defekter Eigenwerte können die „fehlenden“ Eigenvektoren durch $l = n_i - m_i$ Hauptvektoren $\mathbf{x}_{H,l}$ ergänzt werden, die jeweils $(\mathbf{A} - \lambda_i \mathbf{I})\mathbf{x}_{H,l} = 0$ erfüllen [85]. Eine Matrix kann jedoch n linear unabhängige Eigenvektoren haben, obwohl sie mehrfache Eigenwerte besitzt. Als Beispiel dient die Einheitsmatrix \mathbf{I} . Bei der Berechnung von Eigenvektoren mehrfacher Eigenwerte kann die bei hermiteschen Matrizen vorhandene Orthogonalität der Eigenvektoren ausgenutzt werden.

Definition 4.1.20 (Rayleigh-Quotient [82]): Mit $\mathbf{A} \in \mathbb{C}^{n \times n}$ und $\mathbf{v} \in \mathbb{C}^n$ ist der RAYLEIGH-Quotient als

$$\rho_{\mathbf{A}}(\mathbf{v}) := \frac{\mathbf{v}^H \mathbf{A} \mathbf{v}}{\mathbf{v}^H \mathbf{v}}$$

definiert.

Die tatsächliche Tragweite des RAYLEIGH-Quotienten offenbart sich für Eigenwerte mit einfacher Vielfachheit, wenn \mathbf{v} ein Eigenvektor \mathbf{x} ist, da in diesem Fall dann der direkte Zusammenhang zwischen Eigenwert und Eigenvektor entsteht.

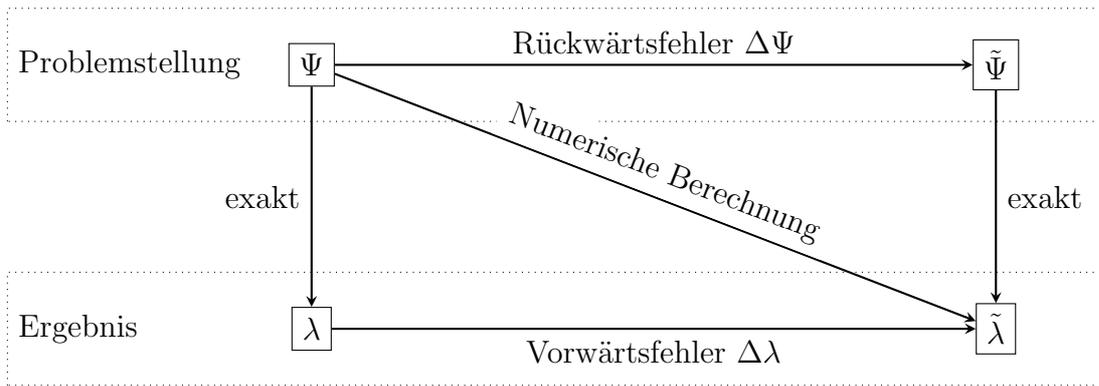


Abbildung 4.1.: Veranschaulichung zum Vorwärts- und Rückwärtsfehler.

Bemerkung 4.1.21: Ist $\mathbf{x} \in \mathbb{C}^n$ ein Eigenvektor von $\mathbf{A} \in \mathbb{C}^{n \times n}$ mit zugehörigem Eigenwert $\lambda \in \mathbb{C}$, so gilt

$$\lambda = \rho_{\mathbf{A}}(\mathbf{x}) := \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}}.$$

Das Konzept des RAYLEIGH-Quotienten lässt sich auf Matrixpolynome verallgemeinern. Für eine Arbeit, die sich ausgiebig mit nichtlinearen RAYLEIGH-Funktionalen befasst, sei an dieser Stelle auf [75] verwiesen.

4.1.3. Fehler-Definitionen

Da insbesondere die iterativen numerischen Methoden die tatsächliche Lösung eines Eigenwertproblems nur sukzessive approximieren, verbleibt beim approximativen Eigenpaar $(\tilde{\lambda}, \tilde{\mathbf{x}})$ als Rest das Residuum \mathbf{r} .

Definition 4.1.22 (Residuum): Sei $\Psi(\tilde{\lambda})$ Matrixpolynom und $(\tilde{\lambda}, \tilde{\mathbf{x}})$ kein Eigenpaar, so wird \mathbf{r} Residuum genannt und es gilt

$$\mathbf{r} = \Psi(\tilde{\lambda})\tilde{\mathbf{x}} \neq \mathbf{0}.$$

Die Vektornorm von \mathbf{r} kann als Abstand zur (im Rahmen der Maschinengenauigkeit) tatsächlichen Lösung angesehen werden. Weitere Fehlerdefinitionen sind für die Analyse numerischer Verfahren hilfreich.

Definition 4.1.23 (Vorwärtsfehler): Der Vorwärtsfehler ist der Betrag der Differenz aus Ergebnis $\tilde{\lambda}$ eines numerischen Verfahrens und tatsächlicher Lösung λ

$$\eta_V = |\tilde{\lambda} - \lambda|.$$

Der Vorwärtsfehler ist somit insbesondere dann nützlich, wenn die exakte Lösung aufgrund analytischer oder physikalischer Überlegungen vorliegt. Der Rückwärtsfehler hingegen trifft eine Aussage darüber, welches Ausgangsproblem $\tilde{\Psi}(\tilde{\lambda}) = \Psi(\tilde{\lambda}) + \Delta\Psi(\tilde{\lambda})$ unter der Annahme der exakten Lösung vorliegen muss, um $\tilde{\lambda}$ zu erhalten. Dieser Zusammenhang wird in Abbildung 4.1 veranschaulicht. Für die Definition des Rückwärtsfehlers nach [84] wird die Spektralnorm $\|\mathbf{A}\|_2 = \sqrt{\lambda_{\max}(\mathbf{A}^H \mathbf{A})}$ benötigt, die sich als Quadratwurzel aus dem betragsgrößten Eigenwert des Matrixprodukts $\mathbf{A}^H \mathbf{A}$ berechnen lässt.

Definition 4.1.24 (Rückwärtsfehler): Der Rückwärtsfehler ist für polynomiale Eigenwertprobleme definiert durch

$$\eta_R(\tilde{\lambda}, \tilde{\mathbf{x}}) := \min\{\epsilon : \tilde{\Psi}(\tilde{\lambda})\tilde{\mathbf{x}} = \mathbf{0}, \|\Delta \mathbf{A}_i\|_2 \leq \epsilon \|\mathbf{A}_i\|_2, i = 0 \dots \ell\}$$

und kann, wie in [84] gezeigt wurde, durch

$$\eta_R(\tilde{\lambda}, \tilde{\mathbf{x}}) = \frac{\|\Psi(\tilde{\lambda})\tilde{\mathbf{x}}\|_2}{\sum_{i=0}^{\ell} |\tilde{\lambda}|^i \|\mathbf{A}_i\|_2 \|\tilde{\mathbf{x}}\|_2}$$

berechnet werden.

4.1.4. Kondition

Die Kondition bzw. die Konditionszahl eines numerischen Problems ist ein Maß dafür, wie empfindlich die Ausgangsgrößen auf kleine Störungen in den Koeffizientenmatrizen reagieren und somit eine Kennzahl für die Lösbarkeit mittels numerischer Algorithmen. Im Rahmen dieser Arbeit sind die Ausgangsgrößen die Eigenwerte, und die numerischen Algorithmen sind die Methoden zur Eigenwertberechnung. Ein Problem ist schlecht konditioniert, wenn seine Konditionszahl gross ist. Für Konditionszahlen nah bei Eins, wird von guter Kondition gesprochen.

Definition 4.1.25 (Konditionszahl): Für den einfachen, endlichen Eigenwert $\lambda \neq 0$ eines Matrixpolynoms Ψ ist die Konditionszahl als Sensitivität von $\Psi(\lambda)$ gegenüber den Störungen $\Delta\Psi, \Delta\lambda, \Delta\mathbf{x}$ definiert.

$$\kappa_{\Psi}(\lambda) := \limsup_{\epsilon \rightarrow 0} \left\{ \frac{|\Delta\lambda|}{\epsilon|\lambda|} : (\Psi(\lambda + \Delta\lambda) + \Delta\Psi(\lambda + \Delta\lambda))(\mathbf{x} + \Delta\mathbf{x}) = \mathbf{0}, \|\Delta \mathbf{A}_i\|_2 \leq \epsilon \|\mathbf{A}_i\|_2, i = 0 \dots \ell \right\}. \quad (4.3)$$

Wie in [83] gezeigt wurde, kann $\kappa_{\Psi}(\lambda)$ als

$$\kappa_{\Psi}(\lambda) = \frac{\|\mathbf{y}\| \|\mathbf{x}\| \sum_{i=0}^{\ell} |\lambda|^i \|\mathbf{A}_i\|_2}{|\mathbf{y}^H \Psi'(\lambda) \mathbf{x}| |\lambda|} \quad (4.4)$$

berechnet werden. Dabei sind \mathbf{x}, \mathbf{y} rechter und linker Eigenvektor bezüglich des Eigenwerts λ und Ψ' die Ableitung des Matrixpolynoms $\Psi(\lambda)$ bezüglich λ . Vorwärts- und Rückwärtsfehler können mittels der Konditionszahl in Beziehung gesetzt werden. Der Vorwärtsfehler ist höchstens so groß wie die Konditionszahl multipliziert mit dem Rückwärtsfehler,

$$\text{Vorwärtsfehler} \leq \text{Konditionszahl} \times \text{Rückwärtsfehler}. \quad (4.5)$$

Vorwärts- und Rückwärtsfehler sind äquivalent, wenn die Konditionszahl $\kappa_{\Psi}(\lambda) = 1$ ist.

4.1.4.1. Vorkonditionierung von PEPs

Zur Verbesserung der Kondition polynomialer Eigenwertprobleme ist in [48] einerseits eine optimale Skalierung der Koeffizientenmatrizen und andererseits die Spektraltransformation eingeführt. Als Skalierungsmatrix werden Diagonalmatrizen mit

Definition 4.1.26 (Skalierungsmatrix):

$$\mathcal{D}_n := \{\mathbf{D} : \mathbf{D} \in \mathbb{C}^{n \times n} \text{ ist diagonal mit } \det(\mathbf{D}) \neq 0\}$$

zugelassen, so dass ihre Hauptdiagonale keine Nullelemente enthält.

Definition 4.1.27 (Skalierung der Koeffizientenmatrizen): Als Skalierung des Matrixpolynoms $\Psi(\lambda)$ wird das Matrixpolynom $\mathbf{D}_1\Psi(\lambda)\mathbf{D}_2$ mit $\mathbf{D}_1, \mathbf{D}_2 \in \mathcal{D}_n$ bezeichnet [48].

Die Eigenwerte sind invariant bezüglich der Skalierung. Wenn (λ, \mathbf{x}) Eigenpaar des Matrixpolynoms $\Psi(\lambda)$ ist, dann ist $(\lambda, \mathbf{D}_2^{-1}\mathbf{x})$ Eigenpaar des skalierten Matrixpolynoms $\mathbf{D}_1\Psi(\lambda)\mathbf{D}_2$.

Satz 4.1.28 (Optimale Skalierung): Sei λ ein einfacher, endlicher, von Null verschiedener Eigenwert des Matrixpolynoms $\Psi(\lambda)$. $\Psi(\lambda)$ heißt bezüglich des Eigenwerts λ optimal skaliert, wenn die Konditionszahl $\kappa_\Psi(\lambda)$ durch die Skalierung minimal wird

$$\kappa_\Psi(\lambda) = \inf_{\mathbf{D}_1, \mathbf{D}_2 \in \mathcal{D}_n} \kappa_{\mathbf{D}_1\Psi\mathbf{D}_2}(\lambda).$$

Ein auf einem heuristischen Ansatz basierender Algorithmus zur Berechnung von $\mathbf{D}_1, \mathbf{D}_2$ ist in [48] angegeben. Darin werden als Einträge von $\mathbf{D}_1, \mathbf{D}_2$ Zweierpotenzen bestimmt, da diese bei der Multiplikation in der Gleitkommaarithmetik lediglich eine Bitverschiebung und keine zusätzlichen Rundungsfehler verursachen. Als warnender Hinweis wird in [88] gezeigt, dass die Skalierung von Eigenwertproblemen der Kondition auch abträglich sein kann, wenn schon ein relativ gut konditioniertes Eigenwertproblem vorliegt.

Eine Maßnahme zur Verbesserung der Kondition von Linearisierungen polynomialer Eigenwertprobleme nach Satz 4.1.8 ist die Spektraltransformation nach [48]. Dabei wird die Größe $\rho(\Psi, \alpha_S)$ zur Abschätzung der Kondition von polynomialen Eigenwertproblem $\Psi(\lambda)$ und zugehöriger Linearisierung $L(\lambda)$ verwendet. Im Rahmen dieser Arbeit wird jedoch nur das Ergebnis aus [48] benötigt.

Satz 4.1.29 (Spektraltransformation): Sei $\Psi(\lambda)$ ein Matrixpolynom vom Grad ℓ und sei

$$\rho(\Psi, \alpha_S) := \frac{\max_{0 \leq i \leq \ell} \alpha_S^i \|\mathbf{A}_i\|_2}{\min(\|\mathbf{A}_0\|_2, \alpha_S^\ell \|\mathbf{A}_\ell\|_2)} \geq 1$$

für $\alpha_S > 0$ definiert. Dann wird $\rho(\Psi, \alpha_S)$ minimal für $\alpha_{opt} = (\|\mathbf{A}_0\|_2 / \|\mathbf{A}_\ell\|_2)^{\frac{1}{\ell}}$ [48].

Hierbei ist zu beachten, dass die Spektraltransformation $\lambda_{opt} = \alpha_{opt}\lambda$ weder Kondition noch Rückwärtsfehler des originalen PEPs ändert. Sie beeinflusst nur die entsprechenden Größen der zugehörigen Linearisierung $L(\lambda)$ aus Satz 4.1.8 [48].

4.1.4.2. Vorkonditionierung von Standardeigenwertproblemen

Zur Skalierung von Standardeigenwertproblemen eignet sich Algorithmus 4.1. Dabei ist es notwendig, zweimal einen Spektralradius $\rho(\mathbf{A})$ nach Definition 4.1.16 zu berechnen. Allerdings reicht die näherungsweise Bestimmung von $\rho(\mathbf{A})$ für praktische Anwendungen aus, so dass die im folgenden Abschnitt vorgestellte *Potenzmethode* nach Algorithmus 4.2 zum Einsatz kommen kann.

Algorithmus 4.1 Skalierung und Verschiebung des Spektrums von Standard eigenwertproblemen

Eingabe: Matrix \mathbf{A} des Standard eigenwertproblems

Ausgabe: Skalierte Matrix \mathbf{A}_σ sowie Skalierung $\hat{\sigma}$ und Verschiebung σ_S

```

1:  $\hat{\sigma} \leftarrow \rho(\mathbf{A})$ 
2:  $\check{\sigma} \leftarrow \hat{\sigma} - \rho(\hat{\sigma}\mathbf{I} - \mathbf{A})$ 
3: if  $\check{\sigma} < 0$  then
4:    $\sigma_S \leftarrow 1.2\check{\sigma}$ 
5: else
6:    $\sigma_S \leftarrow 0$ 
7: end if
8:  $\mathbf{A}_\sigma = \hat{\sigma}^{-1}(\mathbf{A} - \sigma_S\mathbf{I})$ 

```

4.2. Löser für das lineare Eigenwertproblem

Wie in der Einleitung gezeigt ist, können Eigenwerte als Nullstelle des *charakteristischen Polynoms* p_Ψ aufgefasst werden. Jedoch existiert keine analytische Formel für die Nullstellenberechnung, wenn sein Polynomgrad $\deg p_\Psi > 4$ wird, so dass generell nur iterative Verfahren wie z.B. die NEWTON²⁴-Iteration in Frage kommen. Selbst kleine Fehler in den Polynomkoeffizienten führen unter Umständen zu großen Genauigkeitsverlusten in den Nullstellen²⁵, so dass die Eigenwertberechnung über das Lösen des *charakteristischen Polynoms* nur für Polynomgrade ≤ 4 überhaupt in Erwägung gezogen werden kann. Dennoch kann die Anschauung der Eigenwerte als Nullstellen eines *charakteristischen Polynoms* in theoretischen Fragestellungen durchaus hilfreich sein.

Zwar sind alle Verfahren zur Eigenwertbestimmung iterativ, aber es gibt qualitative Unterschiede in der „Iterativität“. Dem Vorschlag in [46] folgend, werden Verfahren, die in einer festen Anzahl an Gleitkommaoperationen (FLOPS, engl. floating point operations per second) alle Eigenwerte einer Matrix berechnen können und darüber hinaus so gut wie nie fehlschlagen, *direkte Verfahren* genannt. Der Speicheraufwand skaliert hierbei mit $\mathcal{O}(n^2)$ für vollbesetzte Matrizen und mit $\mathcal{O}(n)$ für schwachbesetzte Matrizen. Der zugehörige Rechenaufwand liegt in der Größenordnung von $\mathcal{O}(n^3)$. Daraus resultiert, dass diese Verfahren insbesondere für kleine bis mittlere Eigenwertprobleme²⁶ geeignet sind, bei denen der Aufwand mit verfügbaren Rechenanlagen handhabbar bleibt.

Der Aufbau dieses Abschnitts orientiert sich an [62]. Ergänzende und weiterführende Informationen sind ebenfalls in [62] enthalten. Zunächst werden einige iterative Verfahren betrachtet, die dazu geeignet sind, einzelne Eigenwerte zu berechnen. Danach folgen die *direkten Verfahren*, die prinzipiell dazu geeignet sind, alle Eigenwerte eines linearen Eigenwertproblems zu berechnen.

²⁴Sir Isaac Newton (1642-1726), englischer Naturforscher und Verwaltungsbeamter.

²⁵vgl. Beispiel 5.4 in [53]

²⁶Die Bezeichnungen *klein* und *mittlere* Größe sind transient und auf die zugrunde liegenden Rechnerarchitekturen bezogen: Was heute noch *mittlere* Größe ist, lässt sich in einigen Jahren bereits als *kleine* Größe bezeichnen.

4.2.1. Iterative Löser für einzelne Eigenwerte des Standardeigenwertproblems

Das vielleicht einfachste Verfahren zur Berechnung von einzelnen Eigenpaaren ist die in Algorithmus 4.2 gegebene *Potenzmethode* von VON MISES. Unter wenig strengen Annahmen konvergiert die *Potenzmethode* gegen den betragsgrößten Eigenwert und zugehörigen rechten Eigenvektor und ist daher ein mögliches Verfahren zur Berechnung des *Spektralradius* $\rho(\mathbf{A})$ nach Definition 4.1.16. Der Startvektor \mathbf{v}_0 wird mit zufälligen Einträgen besetzt.

Algorithmus 4.2 Potenzmethode nach VON MISES als fortgesetzte Matrixvektormultiplikation mit einem Startvektor.

Eingabe: Startvektor \mathbf{v}_0 , Matrix \mathbf{A} , Toleranzschwelle ϵ , maximale Anzahl an Iterationen

i_{\max}
Ausgabe: Eigenpaar (λ, \mathbf{x})

- 1: starte mit $\mathbf{z} = \mathbf{v}_0$ dem Startvektor
- 2: **for** $i = 1, 2, \dots, i_{\max}$ **do**
- 3: $\mathbf{v} \leftarrow \mathbf{z} / \|\mathbf{z}\|$
- 4: $\mathbf{z} \leftarrow \mathbf{A}\mathbf{v}$
- 5: $\theta \leftarrow \mathbf{v}^H \mathbf{z}$
- 6: **if** $\|\mathbf{z} - \theta \mathbf{v}\| \leq \epsilon |\theta|$ **then**
- 7: Stoppe
- 8: **end if**
- 9: **end for**
- 10: Setze $\lambda \leftarrow \theta$ und $\mathbf{x} \leftarrow \mathbf{v}$

Die Eigenschaften der *Potenzmethode* im Überblick:

- Nur der betragsgrößte Eigenwert λ_1 wird berechnet.
- Betragsgrößter Eigenwert λ_1 muss einfache algebraische Vielfachheit haben, d.h. $|\frac{\lambda_2}{\lambda_1}| < 1$, wenn λ_2 der zweitgrößte Eigenwert ist.
- Der Startvektor \mathbf{v}_0 darf nicht orthogonal zum Eigenvektor \mathbf{x}_1 des betragsgrößten Eigenwerts sein $\mathbf{x}_1^H \mathbf{v}_0 \neq 0$.

Für Fälle mit konjugiert-komplexen betragsgrößten Eigenwerten sei auf die ausführliche Diskussion der *Potenzmethode* in [90] verwiesen. Das Konvergenzverhalten der *Potenzmethode* lässt sich durch die Verwendung von TSCHEBYSCHOW'schen²⁷ Polynomen beschleunigen. Details sind in [24] sowie ausführlich in [53, Kapitel 8.2] angegeben. Zur Berechnung des Eigenpaars mit dem betragsmäßig kleinsten Eigenwert der Matrix \mathbf{A} eignet sich die *Potenzmethode* aus Algorithmus 4.2, wenn zur Iteration an Stelle der Matrix \mathbf{A} ihre Inverse \mathbf{A}^{-1} verwendet wird, da die Eigenwerte der Inversen \mathbf{A}^{-1} gleich der inversen Eigenwerte von \mathbf{A} sind. Zur Berechnung eines Eigenpaars nahe σ_S eignet sich die verschobene *inverse Potenzmethode* oder verschobene *inverse Iteration* aus Algorithmus 4.3. Die zugrunde liegende Idee ist die Anwendung der *Potenzmethode* auf die Inverse der verschobenen Matrix $(\mathbf{A} - \sigma_S \mathbf{I})^{-1}$, wobei σ_S die Verschiebung ist. Die Inverse muss nicht explizit vorliegen, es kann auch in jedem Iterationsschritt ein entsprechendes lineares Gleichungssystem gelöst werden. Der Startvektor \mathbf{v}_0 wird mit zufälligen Einträgen besetzt.

²⁷Pafnuti Lwowitsch Tschebyschow (auch als Tschebyscheff / Tschebyschew / Tschebyschew und im Englischen als Chebyshev transkribiert) (1821-1894), russischer Mathematiker.

Algorithmus 4.3 Inverse Potenzmethode nach WIELANDT mit Verschiebung σ_S , in deren Nähe ein Eigenpaar berechnet wird.

Eingabe: Startvektor \mathbf{v}_0 , Matrix \mathbf{A} , Verschiebung σ_S , Toleranzschwelle ϵ , maximale Anzahl an Iterationen i_{\max}

Ausgabe: Eigenpaar (λ, \mathbf{x})

```

1: Starte mit  $\mathbf{z} = \mathbf{v}_0$  dem Startvektor
2: for  $i = 1, 2, \dots, i_{\max}$  do
3:    $\mathbf{v} \leftarrow \mathbf{z} / \|\mathbf{z}\|$ 
4:    $\mathbf{z} \leftarrow (\mathbf{A} - \sigma_S \mathbf{I})^{-1} \mathbf{v}$ 
5:    $\theta \leftarrow \mathbf{v}^H \mathbf{z}$ 
6:   if  $\|\mathbf{z} - \theta \mathbf{v}\| \leq \epsilon |\theta|$  then
7:     Stoppe
8:   end if
9: end for
10: Setze  $\lambda \leftarrow \sigma_S + 1/\theta$  und  $\mathbf{x} \leftarrow \mathbf{v}/\theta$ 

```

Die Eigenschaften der *inversen Potenzmethode* mit Verschiebung im Überblick:

- Nur der Eigenwert λ_i wird berechnet, der betragsmäßig am nächsten zur Verschiebung σ_S liegt.
- Die Vielfachheit von Eigenwert λ_i muss einfach sein, d.h. $\max_{m \neq i} \left| \frac{\lambda_i - \sigma_S}{\lambda_m - \sigma_S} \right| < 1$.
- Der Startvektor \mathbf{v}_0 darf nicht orthogonal zum gesuchten Eigenvektor \mathbf{x}_i sein $\mathbf{x}_i^H \mathbf{v}_0 \neq 0$.
- Beim Lösen des linearen Gleichungssystems kann Rechenzeit gespart werden, wenn \mathbf{A} faktorisiert vorliegt (z.B. als LU-Zerlegung).
- Da das lineare Gleichungssystem singulärer wird, je näher sich die Verschiebung σ_S an einen Eigenwert θ annähert, wird die Kondition des linearen Gleichungssystems sukzessive schlechter und der numerische Aufwand zur Lösung sukzessive größer.

Eine weitere Verbesserung der *inversen Potenzmethode* mit Verschiebung kann mit Hilfe der *Rayleigh-Quotienten-Iteration* nach Algorithmus 4.4 erreicht werden [69]. Dabei wird die initiale Verschiebung $\sigma_{S,0}$ in den folgenden Iterationen jeweils durch den RAYLEIGH-Quotienten nach Definition 4.1.20 ersetzt, der die bis dahin beste Approximation des Eigenwerts darstellt. Da die Verschiebung nicht fest ist, sollte eine vorliegende LU-Zerlegung zur Lösung des auftretenden Gleichungssystems in jeder Iteration aktualisiert werden. Liegt ein Zielwert σ_S vor, von dem bekannt ist, dass er dicht am gesuchten Eigenwert λ liegt, bietet es sich an, zunächst einige Iterationen der *inversen Iteration* durchzuführen und dann auf die *Rayleigh-Quotienten-Iteration* (RQI) umzuschalten. Analog dazu ist es ebenfalls möglich, die rechte Seite des auftretenden linearen Gleichungssystems für einige *Rayleigh-Quotienten-Iterationen* konstant zu halten, wenn ein Startvektor verfügbar ist, der eine gute Approximation des zu berechnenden Eigenvektors darstellt, wie in Algorithmus 4.5 gezeigt ist.

Die Eigenschaften der *Rayleigh-Quotienten-Iteration* (RQI) im Überblick:

- Nur der Eigenwert λ_i wird berechnet, der betragsmäßig am nächsten zur Verschiebung $\sigma_{S,0}$ liegt.
- Konvergiert, wenn der Eigenwert einfache Vielfachheit hat $\max_{m \neq i} \left| \frac{\lambda_i - \sigma_{S,0}}{\lambda_m - \sigma_{S,0}} \right| < 1$ und der Startvektor nicht orthogonal zum gesuchten Eigenvektor ist $\mathbf{x}_i^H \mathbf{v}_0 \neq 0$.

Algorithmus 4.4 RAYLEIGH–Quotienten–Iteration nach [69]**Eingabe:** Startvektor \mathbf{v}_0 , Matrix \mathbf{A} , Verschiebung σ_S , maximale Anzahl an Iterationen i_{\max} **Ausgabe:** Eigenpaar (λ, \mathbf{x})

```

1: Starte mit dem Startvektor  $\mathbf{v}_0$  und der initialen Verschiebung  $\sigma_{S,0}$ 
2:  $\mathbf{v}_0 \leftarrow \mathbf{v}_0 / \|\mathbf{v}_0\|$ 
3: for  $i = 0, 1, \dots, i_{\max}$  do
4:    $\mathbf{z}_{i+1} \leftarrow (\mathbf{A} - \sigma_{S,i} \mathbf{I})^{-1} \mathbf{v}_i$ 
5:   if  $\mathbf{A} - \sigma_{S,i} \mathbf{I}$  ist singulär then
6:     Löse  $(\mathbf{A} - \sigma_{S,i} \mathbf{I}) \mathbf{x} = \mathbf{0}$ 
7:      $\lambda \leftarrow \sigma_{S,i}$ 
8:     Stoppe
9:   end if
10:   $\mathbf{v}_{i+1} \leftarrow \mathbf{z}_{i+1} / \|\mathbf{z}_{i+1}\|$ 
11:   $\sigma_{S,i+1} \leftarrow \rho_A(\mathbf{v}_{i+1}) := \mathbf{v}_{i+1}^H \mathbf{A} \mathbf{v}_{i+1}$ 
12: end for

```

Algorithmus 4.5 RAYLEIGH–Quotienten–Iteration mit konstanter rechter Seite während der ersten ι_{RQI} Iterationen.**Eingabe:** Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$, Verschiebung σ , Startvektor $\mathbf{v}_0 \in \mathbb{C}^n$, Schranke ι_{RQI} , Toleranz ϵ **Ausgabe:** Eigenpaar (λ, \mathbf{x})

```

1:  $\mathbf{v}_0 \leftarrow \mathbf{v}_0 / \|\mathbf{v}_0\|$ 
2: for  $i = 1, \dots, i_{\max}$  do
3:   if  $i < \iota_{RQI}$  then
4:      $\mathbf{v}_1 \leftarrow (\mathbf{A} - \sigma \mathbf{I})^{-1} \mathbf{v}_0$ 
5:   else
6:      $\mathbf{v}_1 \leftarrow (\mathbf{A} - \sigma \mathbf{I})^{-1} \mathbf{v}_1$ 
7:   end if
8:   if  $\mathbf{A} - \sigma \mathbf{I}$  ist singulär then
9:     Löse  $(\mathbf{A} - \sigma \mathbf{I}) \mathbf{x} = \mathbf{0}$ 
10:     $\lambda \leftarrow \sigma$ 
11:    Stoppe
12:   end if
13:   $\mathbf{v}_1 \leftarrow \mathbf{v}_1 / \|\mathbf{v}_1\|$ 
14:   $\sigma \leftarrow \mathbf{v}_1^H \mathbf{A} \mathbf{v}_1$ 
15: end for

```

- LU-Zerlegung von \mathbf{A} muss oft aktualisiert werden.
- Für die exakte Lösung des linearen Gleichungssystems ergibt sich bei nicht-hermiteschen Matrizen quadratische Konvergenz [69, 62].
- Da das lineare Gleichungssystem singulärer wird, je näher sich die Verschiebung σ an einen Eigenwert θ annähert, wird die Kondition des linearen Gleichungssystems sukzessive schlechter und der numerische Aufwand zur Lösung sukzessive größer.
- Die RQI schlägt fehl, wenn $\|\mathbf{v}_1\| = 0$ auftritt.

4.2.2. Direkte Löser für das Standardeigenwertproblem

Für die vollständige Berechnung kleiner und mittlerer Standardeigenwertprobleme ist der QR-Algorithmus 4.6 der Stand der Technik und in vielen numerischen Bibliotheken²⁸ verfügbar. Gesucht sind die Eigenwerte von komplexen Matrizen, wozu der Satz von SCHUR²⁹ über die Trigonalisierbarkeit von Matrizen hilfreich ist.

Satz 4.2.1 (Schur-Zerlegung): *Jeder Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ lässt sich eine unitäre Matrix $\mathbf{Q} \in \mathbb{C}^{n \times n}$ zuordnen, so dass $\mathbf{Q}^H \mathbf{A} \mathbf{Q} = \mathbf{R}$ eine obere Dreiecksmatrix ist.*

Die Eigenwerte der oberen Dreiecksmatrix \mathbf{R} entsprechen den Hauptdiagonalelementen. Die QR-Zerlegung ist definiert als:

Definition 4.2.2 (QR-Zerlegung): *Die QR-Zerlegung einer Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ ist definiert als*

$$\mathbf{A} = \mathbf{Q}\mathbf{R},$$

wobei $\mathbf{Q} \in \mathbb{C}^{n \times n}$ unitär mit $\mathbf{Q}^H \mathbf{Q} = \mathbf{I}$ ist und $\mathbf{R} \in \mathbb{C}^{n \times n}$ eine obere Dreiecksmatrix ist.

Damit ist die QR-Zerlegung bis auf unitäre Diagonaltransformationen eindeutig bestimmt [81]. Die QR-Zerlegung kann über sukzessive Anwendung von HOUSEHOLDER³⁰-Transformationen oder GIVENS-Rotationen (siehe Anhang A.2) berechnet werden. Innerhalb der QR-Zerlegung, die in Algorithmus 4.6 gegeben ist, entstehen die Approximationen der Eigenwerte von \mathbf{A} aus einer Folge unitärer Ähnlichkeitstransformationen

$$\mathbf{A}_{l+1} := \mathbf{R}_l \mathbf{Q}_l = \mathbf{Q}_l^H \mathbf{A}_l \mathbf{Q}_l = (\mathbf{Q}_{l-1} \mathbf{Q}_l)^H \mathbf{A}_l (\mathbf{Q}_{l-1} \mathbf{Q}_l) = \hat{\mathbf{Q}}_l^H \mathbf{A}_0 \hat{\mathbf{Q}}_l$$

letztendlich auf der Diagonalen von \mathbf{A}_{l+1} , die eine obere Dreiecksmatrix ist. Dabei werden Vektoren $\mathbf{q}_1, \dots, \mathbf{q}_l$ berechnet, so dass $\mathbf{A} \mathbf{Q}_l = \mathbf{Q}_l \mathbf{R}_l$ mit $\mathbf{Q}_l^H \mathbf{Q}_l = \mathbf{I}_l$, und \mathbf{R}_l ist eine $l \times l$ obere Dreiecksmatrix mit den gesuchten Eigenwerten auf der Diagonalen. Die Eigenvektoren können aus \mathbf{Q}_l und \mathbf{R}_l berechnet werden. Einen besonderen Fall für die QR-Zerlegung stellen HESSENBERG-Matrizen³¹ dar (siehe Anhang A.2). Der numerische Aufwand für die QR-Zerlegung ist im Allgemeinen $\mathcal{O}(n^3)$, jedoch nur $\mathcal{O}(n^2)$ für obere HESSENBERG-Matrizen, da dann alle Einträge unterhalb der ersten Nebendiagonalen verschwinden. Daher ist es für die QR-Zerlegung ab einer gewissen Matrixdimension ratsam, die Ausgangsmatrix \mathbf{A} durch unitäre Transformation zunächst auf die Gestalt einer oberen HESSENBERG-Matrix zu bringen,

²⁸In dieser Arbeit wird die MATLAB[126]-Routine `eig` verwendet.

²⁹Issai Schur (1875-1941), weißrussischer Mathematiker.

³⁰Alston Scott Householder (1904-1993), US-amerikanischer Mathematiker.

³¹Karl Adolf Hessenberg (1904-1959), deutscher Elektrotechnik-Ingenieur und Mathematiker.

Algorithmus 4.6 QR-Zerlegung

Eingabe: Matrix \mathbf{A} **Ausgabe:** unitäre \mathbf{Q} und obere Dreiecksmatrix \mathbf{R}

- 1: starte mit $\mathbf{A}_0 \leftarrow \mathbf{A}$
 - 2: **for** $l = 0, 1, \dots$ **do**
 - 3: Berechne die QR-Zerlegung $\mathbf{A}_l = \mathbf{Q}_l \mathbf{R}_l$.
 - 4: $\mathbf{A}_{l+1} \leftarrow \mathbf{R}_l \mathbf{Q}_l$
 - 5: **end for**
-

wobei die unitären Transformationen die Eigenwerte nicht verändern. Vorteilhaft ist, dass während der spaltenweise voranschreitenden QR-Zerlegung die übrige Gestalt der oberen HESSENBERG-Matrix nicht zerstört wird.

Die ursprüngliche Form der QR-Zerlegung nach Algorithmus 4.6 konvergiert oft nur wenig zufriedenstellend, jedoch lässt sich die Konvergenz durch das Einführen einer gut gewählten Verschiebung weiter beschleunigen. Dazu sei \mathbf{H} eine obere HESSENBERG-Matrix. Wenn der Eintrag $h_{n,n-1}$ eliminiert wird, ist ein Eigenwert identisch zum Eintrag $h_{n,n}$. Dann ist das Problem für die unbekanntes Eigenwerte um eine Dimension reduziert auf $\mathbf{H} = (h_{il})$ mit $i, l = 1, \dots, n-1$. Dies kann durch Verwendung der WILKINSON-Verschiebung σ_l folgendermaßen erreicht werden: Für den 2×2 Matrixblock in der rechten unteren Ecke von \mathbf{H}

$$\begin{bmatrix} h_{n-1,n-1} & h_{n-1,n} \\ h_{n,n-1} & h_{n,n} \end{bmatrix}$$

lassen sich über das charakteristische Polynom analytisch die beiden Eigenwerte bestimmen. Der Eigenwert, der $h_{n,n}$ am nächsten ist, wird als Verschiebung verwendet. Daraus resultiert schließlich Algorithmus 4.7.

Algorithmus 4.7 QR-Zerlegung mit Verschiebung

Eingabe: Matrix \mathbf{A} **Ausgabe:** unitäre \mathbf{Q} und obere Dreiecksmatrix \mathbf{R}

- 1: starte mit $\mathbf{A}_0 = \mathbf{A}$
 - 2: **for** $l = 0, 1, \dots$ **do**
 - 3: Wähle Verschiebung σ_l .
 - 4: Berechne die QR-Zerlegung $\mathbf{A}_l - \sigma_l \mathbf{I} = \mathbf{Q}_l \mathbf{R}_l$.
 - 5: Berechne die QR-Zerlegung $\mathbf{A}_l = \mathbf{Q}_l \mathbf{R}_l$.
 - 6: $\mathbf{A}_{l+1} \leftarrow \mathbf{R}_l \mathbf{Q}_l + \sigma_l \mathbf{I}$
 - 7: **end for**
-

4.2.3. Direkte Löser für das allgemeine Eigenwertproblem

Da aus der Linearisierung des polynomialen Eigenwertproblems entsprechend Satz 4.1.8 bzw. (4.1) und (4.2) allgemeine Eigenwertprobleme nach Bemerkung 4.1.5 entstehen, sollen hierfür Lösungsmöglichkeiten benannt werden. Zur Lösung des allgemeinen Eigenwertproblems $\mathbf{A}\mathbf{x} = \lambda\mathbf{B}\mathbf{x}$ ist die *Potenzmethode* aus Algorithmus 4.2 nicht anwendbar. Jedoch ist die *inverse Potenzmethode* (mit und ohne Verschiebung) aus Algorithmus 4.3 ebenso wie die *Rayleigh-Quotienten-Iteration* nach Algorithmus 4.4 anwendbar, wenn jeweils die Identitätsmatrix \mathbf{I} durch \mathbf{B} ersetzt wird, so dass das allgemeine Eigenwertproblem $(\mathbf{A} - \lambda\mathbf{B})\mathbf{x} = 0$ gelöst wird.

4.2.3.1. QZ–Zerlegung

Im Folgenden wird die Vorgehensweise bei der QZ–Zerlegung³² skizziert, die (mit Ausnahme von Spezialfällen) in der Lage ist, in einer festen Anzahl von Iterationen das allgemeine Eigenwertproblem vollständig zu lösen. Da in dieser Arbeit letzten Endes die MATLAB[126]–Routine `eig` verwendet wird, bleibt es bei einer prinzipiellen Skizzierung, die [46] folgt. Ausführliche Darstellungen sind beispielsweise in [60, Kapitel 7.7] und in [67, Kapitel 2] enthalten. Mit Hilfe der QZ–Zerlegung lassen sich alle Eigenwerte des allgemeinen Eigenwertproblems $\mathbf{Ax} = \lambda\mathbf{Bx}$ bestimmen, so dass die QZ–Zerlegung für allgemeine Eigenwertprobleme das Analogon der QR–Zerlegung bei Standardeigenwertproblemen darstellt. Die prinzipielle Vorgehensweise für reelle Matrizen \mathbf{A}, \mathbf{B} lautet:

1. Zunächst wird das Matrixbüschel \mathbf{A}, \mathbf{B} einer Transformation unterzogen, so dass \mathbf{A} eine obere HESSENBERG–Matrix und \mathbf{B} eine obere Dreiecksmatrix ist. Im nächsten Schritt muss auch \mathbf{A} auf obere Dreiecksform gebracht werden.
2. Dazu werden die Auswirkungen einzelner QR–Schritte auf das Produkt \mathbf{AB}^{-1} dargestellt durch unitäre Äquivalenztransformationen \mathbf{Q} und \mathbf{Z} auf den Matrixbüschel \mathbf{A}, \mathbf{B} . Dies ist das Herzstück der QZ–Zerlegung: \mathbf{A} wird sukzessive auf obere Dreiecksform gebracht, während gleichzeitig darauf geachtet wird, dass die obere Dreiecksform von \mathbf{B} nicht zerstört wird. Nach erfolgter Konvergenz resultieren orthogonale \mathbf{Q} und \mathbf{Z} , so dass \mathbf{QAZ} und \mathbf{QBZ} obere Dreiecksform haben.
3. Eigenwerte werden durch Quotientenbildung der Diagonaleinträge von \mathbf{QAZ} und \mathbf{QBZ} berechnet. Eigenvektoren können aus dem triangulierten System berechnet und mittels \mathbf{Z} auf die Eigenvektoren des Ausgangsproblems zurück transformiert werden.

Die QZ–Zerlegung benötigt für vollbesetzte Matrizen etwa $30n^3$ FLOPS zur Berechnung der Eigenwerte. Weitere $16n^3$ FLOPS sind notwendig zur Berechnung der Eigenvektoren.

4.3. Grundlagen des Jacobi–Davidson–Verfahrens für Standardeigenwertprobleme

Für viele Fragestellungen aus Naturwissenschaft und Technik ist jedoch weniger das komplette Spektrum an Eigenwerten interessant. Es reichen oft einige 10 bis 100 Eigenwerte. Darüber hinaus führen die Eigenwertprobleme, die insbesondere durch Diskretisierung mit FIT und FLAME entstehen, auf schwachbesetzte Matrizen hoher Dimension, so dass die Berechnung aller Eigenwerte so unnötig wie unpraktikabel ist. Für diese Art von Problemstellung gibt es Algorithmen zur Eigenwertberechnung, die auf Projektionsmatrizen aufbauen, mit denen die großen Matrizen zunächst auf Matrizen kleinerer bis mittlerer Größe reduziert werden. Diese kleineren Matrizen können dann effizient mit den direkten Verfahren aus Abschnitt 4.2 gelöst werden. Die reduzierte Matrix wird dann iterativ erweitert, mit dem Ziel, dass einige Eigenwerte der reduzierten Matrix gute Approximationen der gesuchten Eigenwerte der großen Ausgangsmatrix sind. Diese iterativen Unterraumverfahren sind für die Berechnung einer begrenzten Anzahl von Eigenwerten geeignet – die Berechnung aller Eigenwerte führte zu erheblichen Aufwand und ist nicht praktikabel.

³²Die QZ–Zerlegung wird auch verallgemeinerte SCHUR–Zerlegung genannt.

Abgrenzung zu Krylow–Unterraumverfahren

Das in dieser Arbeit verwendete Verfahren aus der Klasse der iterativen Unterraumverfahren ist das JACOBI–DAVIDSON–Verfahren [80]. Generell sind Eigenwertlöser aus der Klasse der KRYLOW³³–Unterraumverfahren³⁴ für die in dieser Arbeit behandelten Eigenwertprobleme geeignet – wenn auch nicht in gleichem Maße, wie das JACOBI–DAVIDSON–Verfahren. Die KRYLOW–Unterräume sind zunächst aufwendiger zu erzeugen und zusätzlicher Aufwand während der Iteration ist notwendig, um ihre spezielle Struktur zu erhalten, was im Allgemeinen die exakte Lösung eines linearen Gleichungssystems pro Iterationsschritt erfordert [46]. Das JACOBI–DAVIDSON–Verfahren hingegen verwendet direkt die Koeffizientenmatrizen des Matrixpolynoms $\Psi(\lambda)$, so dass keine Inversen oder exakte Lösungen benötigt werden, sondern die auftretenden linearen Gleichungssysteme näherungsweise gelöst werden dürfen. Desweiteren werden keine besonderen Anforderungen an die Koeffizientenmatrizen gestellt³⁵, so dass diese im Gegensatz zu den KRYLOW–Verfahren nicht positiv-semidefinit sein müssen, sondern auch singulär sein dürfen. Darüber hinaus ist eine Verallgemeinerung auf polynomiale Eigenwertprobleme leicht möglich und es existiert eine gewisse Freiheit für die Erweiterung des Unterraums [57].

4.3.1. Einführung

Zum JACOBI–DAVIDSON–Verfahren von SLEIJPEN und VAN DER VORST [56, 77, 78, 79] gibt es mit dem JACOBI–DAVIDSON–Gateway [63] eine umfangreiche Software-, Literatur- und Personensammlung. Nach einer allgemein gehaltenen Einführung liegt der Schwerpunkt auf Varianten des JACOBI–DAVIDSON–Verfahrens für nicht-hermitesche Standard eigenwertprobleme, um die Gleichungen (3.42) und (3.38) lösen zu können. Das JACOBI–DAVIDSON–Verfahren vereinigt grundlegende Ideen der Methoden von DAVIDSON und JACOBI, die in Anhang A.3 ausführlicher erläutert sind. Für die iterativen Unterraumverfahren sind RITZ–Werte³⁶ und RITZ–Vektoren von zentraler Bedeutung. Die Definition erfolgt hier für Matrixpolynome, das Standard eigenwertproblem ist entsprechend Bemerkung 4.1.6 enthalten.

Definition 4.3.1 (Ritz–Wert, Ritz–Vektor): Sei $\mathbf{V}_m \in \mathbb{C}^{n \times m}$ mit $1 < m \ll n$ Orthonormalbasis des m -dimensionalen Unterraums von $\mathbb{C}^{n \times n}$ und die m Spalten von \mathbf{V} sind $\mathbf{v}_1, \dots, \mathbf{v}_m$. Ψ ist ein Matrixpolynom vom Grad ℓ . Das Paar (θ, \mathbf{u}) mit $\theta \in \mathbb{C}$, $\mathbf{0} \neq \mathbf{u} \in \text{im}(\mathbf{V})$ heißt RITZ–Paar, wenn es die GALERKIN³⁷–Bedingung

$$\Psi(\theta)\mathbf{u} \perp \text{im}(\mathbf{V}_m)$$

erfüllt. Diese Bedingung führt auf das reduzierte $m \times m$ -Eigenwertproblem

$$\mathbf{V}_m^H \Psi(\theta) \mathbf{V}_m \mathbf{s} = 0 \quad \text{mit} \quad \mathbf{u} = \mathbf{V}_m \mathbf{s}.$$

Die m Eigenwerte $\theta_i^{(m)}$ werden RITZ–Werte und die m Eigenvektoren $\mathbf{u}_i = \mathbf{V}_m \mathbf{s}_i^{(m)}$ werden RITZ–Vektoren genannt.

³³Alexei Nikolajewitsch Krylow (1863-1945), russischer Schiffsbauingenieur und Mathematiker.

³⁴Ein KRYLOW–Unterraum der Dimension m zur Matrix \mathbf{A} wird durch die Vektoren $\mathbf{q}, \mathbf{A}\mathbf{q}, \mathbf{A}^2\mathbf{q}, \dots, \mathbf{A}^{m-1}\mathbf{q}$ aufgespannt.

³⁵Wenn besondere Eigenschaften der Matrizen wie Symmetrie bekannt sind, so lassen sich diese dennoch berücksichtigen, was zu effizienteren Algorithmen für diese speziellen Matrizen führt.

³⁶Walter Ritz (1878-1909), schweizerischer Mathematiker und Physiker.

³⁷Boris Grigorjewitsch Galjorkin (auch Galerkin transkribiert) (1871-1945), russischer Ingenieur und Mathematiker.

Der hochgestellte Index symbolisiert dabei, dass das jeweilige RITZ–Paar unter Verwendung eines m -dimensionalen Unterraums berechnet ist. Demnach wird mit Definition 4.3.1 sichergestellt, dass das Residuum $\mathbf{r} = \Psi(\theta_i^{(m)})\mathbf{u}_i^{(m)}$ orthogonal zum Bildbereich des jeweiligen Unterraums \mathbf{V}_m ist. Die RITZ–Paare stellen Approximationen der Eigenpaare dar und Ziel ist es, den zugrunde liegenden Unterraum \mathbf{V}_m sukzessive so zu erweitern, so dass die RITZ–Paare schließlich die Eigenpaare so gut approximieren, dass sie aufgrund der hinreichend kleinen Norm von \mathbf{r} als Lösung akzeptiert werden können.

Für das Standardeigenwertproblem $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ sei der RITZ–Vektor $\mathbf{u}_i^{(m)}$ eine Approximation des Eigenvektors \mathbf{x} zum Eigenwert λ . Der Idee von JACOBI³⁸ [66] folgend, lässt sich ein orthogonaler Korrekturvektor \mathbf{t} berechnen, so dass entsprechend $\mathbf{x} = \mathbf{u}_i^{(m)} + \mathbf{t}$

$$(\mathbf{A} - \lambda\mathbf{I})(\mathbf{u}_i^{(m)} + \mathbf{t}) = 0 \quad (4.6)$$

gilt. Da $\mathbf{t} \perp \mathbf{u}_i^{(m)}$ gelten soll, kann das Problem auf den zu $\mathbf{u}_i^{(m)}$ orthogonalen Unterraum beschränkt werden. Die Restriktion von \mathbf{A} auf diesen Unterraum ist durch

$$(\mathbf{I} - \mathbf{u}_i^{(m)}\mathbf{u}_i^{(m)H})\mathbf{A}(\mathbf{I} - \mathbf{u}_i^{(m)}\mathbf{u}_i^{(m)H})$$

gegeben und der Korrekturvektor \mathbf{t} erfüllt die Gleichung [46]

$$(\mathbf{I} - \mathbf{u}_i^{(m)}\mathbf{u}_i^{(m)H})(\mathbf{A} - \lambda\mathbf{I})(\mathbf{I} - \mathbf{u}_i^{(m)}\mathbf{u}_i^{(m)H})\mathbf{t} = -(\mathbf{A} - \theta_i^{(m)}\mathbf{I})\mathbf{u}_i^{(m)}.$$

In der Praxis ist λ zunächst nicht exakt bekannt, so dass λ naheliegenderweise durch die verfügbare Approximation, nämlich den RITZ–Wert $\theta_i^{(m)}$, ersetzt wird. Dies führt auf die Korrekturgleichung des JACOBI–DAVIDSON–Verfahrens für einen Korrekturvektor $\mathbf{t}_i^{(m)} \perp \mathbf{u}_i^{(m)}$:

$$(\mathbf{I} - \mathbf{u}_i^{(m)}\mathbf{u}_i^{(m)H})(\mathbf{A} - \theta_i^{(m)}\mathbf{I})(\mathbf{I} - \mathbf{u}_i^{(m)}\mathbf{u}_i^{(m)H})\mathbf{t}_i^{(m)} = -\mathbf{r}_i^{(m)} = -(\mathbf{A} - \theta_i^{(m)}\mathbf{I})\mathbf{u}_i^{(m)}. \quad (4.7)$$

Die exakte Lösung $\mathbf{t}_i^{(m)}$ oder die Näherungslösung $\tilde{\mathbf{t}}_i^{(m)}$ von (4.7) wird zur Erweiterung des Unterraums \mathbf{V}_m verwendet. Die exakte Lösung von (4.7) führt bei symmetrischen und hermiteschen Eigenwertproblemen maximal zu kubischer Konvergenz und bei unsymmetrischen Eigenwertproblemen immer noch zu bestenfalls quadratischer Konvergenz des JACOBI–DAVIDSON–Verfahrens. Aufgrund dieser starken Konvergenzeigenschaften ist eine näherungsweise Lösung dieser Gleichung meistens ausreichend, sie wird in Abschnitt 4.3.7 diskutiert. In Algorithmus 4.8 ist der prinzipielle Ablauf des JACOBI–DAVIDSON–Verfahrens für Matrixpolynome angegeben. Das Standardeigenwertproblem ist entsprechend Bemerkung 4.1.6 enthalten. In jeder Iteration wird ein neues RITZ–Paar (θ, \mathbf{u}) bestimmt (Zeile 6). Die Iteration wird abgebrochen, wenn die Vektornorm des Residuums unter eine zuvor festgelegte Toleranzschwelle fällt (Zeilen 8–11). Die Lösung der Korrekturgleichung (Zeile 12) wird verwendet, um den Unterraum zu erweitern (Zeile 4), in den das Ausgangsproblem projiziert wird (Zeile 5). Varianten des JACOBI–DAVIDSON–Verfahrens, die eine Verkleinerung des Unterraums \mathbf{V}_m ermöglichen, werden in den folgenden Abschnitten diskutiert.

4.3.2. Verwandtschaft mit dem Newton–Verfahren

Die enge Verwandtschaft des JACOBI–DAVIDSON–Verfahrens zum NEWTON–Verfahren kann allgemein für Matrixpolynome Ψ gezeigt werden [75, 77]. Wie oben beschrieben, kann das

³⁸Siehe Anhang A.3

Algorithmus 4.8 Schema des JACOBI–DAVIDSON–Verfahrens nach [78, 80]

Eingabe: Matrixpolynom $\Psi(\lambda)$, Zielwert τ , Startvektor \mathbf{v}_0 , Toleranz ϵ_{tol} **Ausgabe:** Eigenpaar (λ, \mathbf{x})

```

1:  $\mathbf{t} \leftarrow \mathbf{v}_0$ 
2:  $\theta \leftarrow \tau$ 
3: while 1 do
4:   Erweitere und orthonormalisiere  $\mathbf{V}_m$  um  $\mathbf{t}$ 
5:   Projiziere  $\Psi$  auf den aktuellen Unterraum  $\mathbf{V}_m$ :  $\mathbf{M} \leftarrow \mathbf{V}_m^H \Psi(\theta) \mathbf{V}_m$ 
6:   Extrahiere RITZ–Paar  $(\theta, \mathbf{u})$  mit  $\theta$  am nächsten zu  $\tau$  aus projiziertem System  $\mathbf{M}$ 
7:   Berechne Residuum  $\mathbf{r} \leftarrow \Psi(\theta) \mathbf{u}$ 
8:   if  $\|\mathbf{r}\| < \epsilon_{\text{tol}}$  then
9:      $\lambda \leftarrow \theta$ ,  $\mathbf{x} \leftarrow \mathbf{u}$ 
10:   Stoppe
11: end if
12:   Löse Korrekturgleichung nach  $\mathbf{t}$ 
13: end while

```

RITZ–Paar (θ, \mathbf{u}) mit der Korrektur $(\Delta\theta, \mathbf{t})$ zum Eigenpaar $(\lambda = \theta + \Delta\theta, \mathbf{x} = \mathbf{u} + \mathbf{t})$ ergänzt werden, so dass

$$\Psi(\theta + \Delta\theta)(\mathbf{u} + \mathbf{t}) = \Psi(\lambda)\mathbf{x} = \mathbf{0} \quad (4.8)$$

gilt. Für die Betrachtung als NEWTON–Schritt ist es günstig, sich zunächst eine Funktion von λ und \mathbf{x} sowie deren Ableitung zu definieren

$$F(\mathbf{x}, \lambda) := \begin{bmatrix} \Psi(\lambda)\mathbf{x} \\ \mathbf{w}^H \mathbf{x} - 1 \end{bmatrix} \quad \partial F(\mathbf{x}, \lambda) := \begin{bmatrix} \Psi(\lambda) & \Psi'(\lambda)\mathbf{x} \\ \mathbf{w}^H & 0 \end{bmatrix}. \quad (4.9)$$

Dabei ist der untere Eintrag der Normierung bezüglich \mathbf{w} geschuldet. Die Ableitung ist durch die JACOBI–Matrix gegeben. Ein Schritt des NEWTON–Verfahrens für das RITZ–Paar (θ, \mathbf{u}) nach Definition 4.3.1 mit der Korrektur $(\Delta\theta, \mathbf{t})$ ist durch

$$\partial F(\mathbf{u}, \theta) \begin{bmatrix} \mathbf{t} \\ \Delta\theta \end{bmatrix} = -F(\mathbf{u}, \theta) \Leftrightarrow \begin{bmatrix} \Psi(\theta) & \Psi'(\theta)\mathbf{u} \\ 2\mathbf{u}^H & 0 \end{bmatrix} \begin{bmatrix} \mathbf{t} \\ \Delta\theta \end{bmatrix} = - \begin{bmatrix} \Psi(\theta)\mathbf{u} \\ 0 \end{bmatrix} = - \begin{bmatrix} \mathbf{r} \\ 0 \end{bmatrix} \quad (4.10)$$

gegeben, wobei hier zur Normierung $\mathbf{w} = \mathbf{u}$ gesetzt wird. Die Iteration des NEWTON–Verfahrens ergibt sich dann zu $\mathbf{u}_{i+1} = \mathbf{u}_i + \mathbf{t}_i$ und $\theta_{i+1} = \theta_i + \Delta\theta_i$. Zur Herleitung der Korrekturgleichung des JACOBI–DAVIDSON–Verfahrens wird die erste Gleichung aus (4.10) von links mit \mathbf{u}^H multipliziert, so dass

$$\mathbf{u}^H \Psi(\theta)\mathbf{t} + \Delta\theta \mathbf{u}^H \Psi'(\theta)\mathbf{u} = -\mathbf{u}^H \mathbf{r} = 0, \quad \text{da } \mathbf{u} \perp \mathbf{r} \quad (4.11)$$

gilt. Diese Gleichung lässt sich nach $\Delta\theta$ auflösen und das Ergebnis wird in die erste Gleichung von (4.10) eingesetzt. Es folgt

$$\left(\mathbf{I} - \frac{\Psi'(\theta)\mathbf{u}\mathbf{u}^H}{\mathbf{u}^H \Psi'(\theta)\mathbf{u}} \right) \Psi(\theta)\mathbf{t} = -\Psi(\theta)\mathbf{u}. \quad (4.12)$$

Die zweite Gleichung des Systems (4.10) fordert $\mathbf{t} \perp \mathbf{u}$, was durch die Orthogonalprojektion

$$\left(\mathbf{I} - \frac{\mathbf{u}\mathbf{u}^H}{\mathbf{u}^H \mathbf{u}} \right) \mathbf{t} = \mathbf{t} \quad (4.13)$$

ausgedrückt wird. Einsetzen von (4.13) in (4.12) liefert die Korrekturgleichung (4.7) des JACOBI–DAVIDSON–Verfahrens in der Schreibweise für Matrixpolynome $\Psi(\lambda)$

$$\left(\mathbf{I} - \frac{\Psi'(\theta)\mathbf{u}\mathbf{u}^H}{\mathbf{u}^H\Psi'(\theta)\mathbf{u}} \right) \Psi(\theta) \left(\mathbf{I} - \frac{\mathbf{u}\mathbf{u}^H}{\mathbf{u}^H\mathbf{u}} \right) \mathbf{t} = -\Psi(\theta)\mathbf{u} = -\mathbf{r}. \quad (4.14)$$

Da für die Korrekturgleichung (4.14) im JACOBI–DAVIDSON–Verfahren auch eine näherungsweise Lösung in Betracht kommt, wird dieser Fall *inexakter* NEWTON–Schritt genannt [55].

4.3.3. Deflation

Mit dem JACOBI–DAVIDSON–Verfahren nach Algorithmus 4.8 ist es nicht möglich, mehr als ein einziges Eigenpaar zu berechnen. Um diesen Nachteil zu beheben, gibt es die Strategie der *Deflation*, die im Folgenden anhand des Standardeigenwertproblems diskutiert wird. Dazu wird die Matrix \mathbf{A}_d betrachtet, die sich zu

$$\mathbf{A}_d := (\mathbf{I} - \mathbf{Q}\mathbf{Q}^H)\mathbf{A}(\mathbf{I} - \mathbf{Q}\mathbf{Q}^H) \quad (4.15)$$

ergibt. Dabei wird die Matrix $\mathbf{Q} = [\mathbf{q}_1, \dots, \mathbf{q}_l] \in \mathbb{C}^{n \times l}$ spaltenweise aus den konvergierten RITZ–Vektoren gebildet, welche l Eigenvektorapproximationen von \mathbf{A} darstellen. Durch Anwendung der Korrekturgleichung des JACOBI–DAVIDSON–Verfahrens auf \mathbf{A}_d , wird der zu berechnende Korrekturvektor \mathbf{t} orthogonal zu den bisher gefundenen Eigenvektorapproximationen in \mathbf{Q} gehalten, so dass bereits berechnete Eigenvektorapproximationen kein zweites mal auftreten. Voraussetzung dafür sind hinreichend kleine Residuen von \mathbf{q}_l . Diese Strategie wird *explizite Deflation* genannt und empfiehlt sich laut [56], da sie verlässlich und numerisch stabil ist. Einsetzen von (4.15) in (4.7) liefert die modifizierte Korrekturgleichung

$$(\mathbf{I} - \mathbf{u}\mathbf{u}^H)(\mathbf{A}_d - \theta\mathbf{I})(\mathbf{I} - \mathbf{u}\mathbf{u}^H)\mathbf{t} = -\mathbf{r}$$

bzw. mit der Definition $\tilde{\mathbf{Q}} := [\mathbf{Q}, \mathbf{u}]$ ergibt sich

$$(\mathbf{I} - \tilde{\mathbf{Q}}\tilde{\mathbf{Q}}^H)(\mathbf{A} - \theta\mathbf{I})(\mathbf{I} - \tilde{\mathbf{Q}}\tilde{\mathbf{Q}}^H)\mathbf{t} = -\mathbf{r}.$$

Mit dieser Erweiterung ist das JACOBI–DAVIDSON–Verfahren aus Algorithmus 4.8 in der Lage mehr als ein Eigenpaar zu berechnen.

4.3.4. Neustarts

Im JACOBI–DAVIDSON–Verfahren nach Algorithmus 4.8 wird in jedem Iterationsschritt der Unterraum \mathbf{V}_m um den Korrekturvektor \mathbf{t} erweitert, so dass die Anzahl der Spalten m von \mathbf{V}_m mit jedem Iterationsschritt anwächst. Somit wird die Extraktion eines neuen RITZ–Paares durch Lösen des niederdimensionalen Eigenwertproblems sukzessive aufwendiger. Als Neustart wird die Verkleinerung des Unterraums \mathbf{V}_m auf eine vorgegebene Dimension m_{\min} bezeichnet. Dem Unterraum \mathbf{V}_m liegt keine spezielle Struktur zugrunde, die während der Iteration erhalten werden muss, so dass es naheliegend wäre, für einen Neustart einfach den letzten berechneten RITZ–Vektor zu verwenden. Dies ist allerdings nicht die effizienteste Strategie für einen Neustart, da dabei die möglicherweise brauchbaren Komponenten, die in den anderen RITZ–Vektoren enthalten sein können, vernachlässigt werden. Da im Allgemeinen alle

Vektoren des Unterraums \mathbf{V}_m (mehr oder weniger) brauchbare Komponenten des gesuchten Eigenvektors enthalten, würde ein Neustart mit einem einzigen Vektor den Konvergenzverlauf verlangsamen. Daher ist es in den meisten Fällen besser, den Neustart mit $m_{\min} > 1$ Vektoren durchzuführen. Als gute Strategie hat sich die Wahl der m_{\min} RITZ-Vektoren herausgestellt, die zu den m_{\min} RITZ-Werten gehören, die dem Zielwert τ am nächsten liegen [46].

4.3.5. Berechnung äußerer Eigenwerte nicht-hermitescher Matrizen

Äußere Eigenwerte in der komplexen Ebene lassen sich dadurch charakterisieren, dass sie extremal sind, also minimal oder maximal bezüglich des Betrags, des Realteils oder des Imaginärteils. Nicht-hermitesche Matrizen haben im Allgemeinen komplexe Eigenwerte und ihre Eigenvektoren bilden keine Orthogonalbasis. Mit dem JACOBI-DAVISON-Algorithmus 4.9 nach [46, 56, 86] ist es möglich, l_{\max} äußere Eigenwerte des Standard eigenwertproblems zu berechnen. Dabei kann die Vorgehensweise aus Algorithmus 4.8 übernommen werden, wenn gleich die Ausgestaltung einzelner Schritte problemangepasst erfolgt. Die orthogonale Unterraumprojektion nach Definition 4.3.1 gilt weiterhin und das Residuum \mathbf{r} soll weiterhin orthogonal zum approximierenden Unterraum sein. Die Extraktion der jeweiligen RITZ-Paare muss selbstverständlich mit einer Methode erfolgen, die für nicht-hermitesche Eigenwertprobleme geeignet ist. Die maximale Dimension $m = m_{\max}$ des Unterraums \mathbf{V}_m wird begrenzt, um nicht unnötig Performanz zu verlieren, da für das projizierte Eigenwertproblem die direkten Methoden aus Abschnitt 4.2 verwendet werden, deren asymptotische Komplexität typischerweise $\mathcal{O}(m^3)$ beträgt. Darüber hinaus lassen sich mit Algorithmus 4.9 im Gegensatz zu Algorithmus 4.8 mehrere Eigenwerte berechnen.

Der Unterraum \mathbf{V}_m in Algorithmus 4.9 wird aus Gründen der numerischen Stabilität weiterhin orthonormal angesetzt. Anstelle der Berechnung der Eigenvektoren des projizierten Eigenwertproblems wird eine QR-Zerlegung nach Definition 4.2.2 durchgeführt, so dass $\mathbf{A}\mathbf{Q}_m = \mathbf{Q}_m\mathbf{R}_m$ mit $\mathbf{A} \in \mathbb{C}^{n \times n}$, orthonormalem $\mathbf{Q} \in \mathbb{C}^{n \times m}$ und der oberen Dreiecksmatrix $\mathbf{R}_m \in \mathbb{C}^{m \times m}$. Aufgrund der durchgängig aufrecht erhaltenen QR-Zerlegung des projizierten Eigenwertproblems wird Algorithmus 4.9 im Folgenden auch mit JDQR bezeichnet. Algorithmus 4.9 liefert l_{\max} Eigenwerte nahe dem Zielwert τ in der komplexen Ebene. Da es keine inhärente Ordnung in der komplexen Ebene gibt, erfolgt die Sortierung nach dem Betrag. Eingabeparameter sind der Startvektor \mathbf{v}_0 , die Toleranzschwelle ϵ und die Anzahl gesuchter Eigenwerte l_{\max} . Die Dimension des orthogonalen Unterraums wird zwischen m_{\min} und m_{\max} gehalten. Nach erfolgreicher Terminierung liefert Algorithmus 4.9 l_{\max} Eigenwerte in der Nähe von τ und die zu \mathbf{A} gehörige partielle SCHUR-Form $\mathbf{A}\mathbf{Q} = \mathbf{Q}\mathbf{R}$ mit orthogonalem $\mathbf{Q} \in \mathbb{C}^{n \times l_{\max}}$ und der oberen Dreiecksmatrix $\mathbf{R} \in \mathbb{C}^{l_{\max} \times l_{\max}}$, die auf ihrer Diagonalen die Eigenwerte enthält. Die Eigenvektoren können danach aus der partiellen SCHUR-Zerlegung gewonnen werden. Zur Diskussion der Abschnitte von Algorithmus 4.9 (JDQR) mit Referenzen auf die jeweiligen Zeilennummern:

- (1) Initialisierung.
- (3) Sofern $m > 0$ wird der Korrekturvektor \mathbf{t} zum Unterraum \mathbf{V}_m orthogonalisiert. Hierbei kann zur besseren numerischen Stabilität auch Algorithmus A.1 verwendet werden.
- (7)–(10) Berechnung der letzten Zeile und Spalte der vollbesetzten, niederdimensionalen Matrix $\mathbf{M} = \mathbf{V}_m^H \mathbf{A} \mathbf{V}_m = \mathbf{V}_m^H \mathbf{V}_m^A$. \mathbf{V}_m besteht aus den Spalten \mathbf{v}_i , so dass

Algorithmus 4.9 JACOBI–DAVIDSON–Verfahren nach [56, 80] zur Berechnung l_{\max} äußerer Eigenwerte des Standardeigenwertproblems (JDQR)

Eingabe: Matrix \mathbf{A} , Startvektor \mathbf{v}_0 , Toleranz ϵ , Zielwert τ , Anzahl gesuchter Eigenpaare

l_{\max}

Ausgabe: Partielle Schur-Zerlegung \mathbf{Q}, \mathbf{R}

```

1:  $\mathbf{t} \leftarrow \mathbf{v}_0, l \leftarrow 0, m \leftarrow 0, \mathbf{Q} \leftarrow [], \mathbf{R} \leftarrow []$ 
2: while  $l < l_{\max}$  do                                     % gesuchte Anzahl an Eigenpaaren noch nicht gefunden
3:   for  $i = 1, \dots, m$  do
4:      $\mathbf{t} \leftarrow \mathbf{v}_i^H \mathbf{t} \mathbf{v}_i$ 
5:   end for
6:    $m \leftarrow m + 1, \mathbf{v}_m \leftarrow \mathbf{t} / \|\mathbf{t}\|, \mathbf{v}_m^A \leftarrow \mathbf{A} \mathbf{v}_m$ 
7:   for  $i = 1, \dots, m - 1$  do
8:      $\mathbf{M}_{i,m} \leftarrow \mathbf{v}_i^H \mathbf{v}_m^A, \mathbf{M}_{m,i} \leftarrow \mathbf{v}_m^H \mathbf{v}_i^A$ 
9:   end for
10:   $\mathbf{M}_{m,m} \leftarrow \mathbf{v}_m^H \mathbf{v}_m^A$ 
11:  Berechne SCHUR–Zerlegung  $\mathbf{M} = \mathbf{S} \mathbf{T} \mathbf{S}^H$ 
12:  Sortiere dabei bezüglich  $\tau$ , so dass  $|\mathbf{T}_{i,i} - \tau| \leq |\mathbf{T}_{i+1,i+1} - \tau|$ 
13:   $\mathbf{u} \leftarrow \mathbf{V} \mathbf{s}, \mathbf{u}^A \leftarrow \mathbf{A} \mathbf{V} \mathbf{s}_1, \theta \leftarrow \mathbf{T}_{1,1}, \mathbf{r} \leftarrow \mathbf{u}^A - \theta \mathbf{u}, \tilde{\mathbf{a}} \leftarrow \mathbf{Q}^H \mathbf{r}, \tilde{\mathbf{r}} \leftarrow \mathbf{r} - \mathbf{Q} \tilde{\mathbf{a}}$ 
14:  while  $\|\tilde{\mathbf{r}}\| \leq \epsilon$  do                                     % Konvergenz?
15:     $\mathbf{R} \leftarrow \begin{bmatrix} \mathbf{R} & \tilde{\mathbf{a}} \\ \mathbf{0} & \theta \end{bmatrix}, \mathbf{Q} \leftarrow [\mathbf{Q}, \mathbf{u}], l \leftarrow l + 1$ 
16:    if  $l = l_{\max}$  then
17:      Stoppe
18:    end if
19:     $m \leftarrow m - 1$ 
20:    for  $i = 1, \dots, m$  do
21:       $\mathbf{v}_i \leftarrow \mathbf{V} \mathbf{s}_{i+1}, \mathbf{v}_i^A \leftarrow \mathbf{A} \mathbf{V} \mathbf{s}_{i+1}, \mathbf{s}_i \leftarrow \mathbf{e}_i$ 
22:    end for
23:    Setze  $\mathbf{M}$  auf rechten unteren  $m \times m$  Block von  $\mathbf{T}$  und
24:     $\mathbf{u} \leftarrow \mathbf{v}_1, \theta \leftarrow \mathbf{M}_{1,1}, \mathbf{r} \leftarrow \mathbf{v}_1^A - \theta \mathbf{u}, \tilde{\mathbf{a}} \leftarrow \mathbf{Q}^H \mathbf{r}, \tilde{\mathbf{r}} \leftarrow \mathbf{r} - \mathbf{Q} \tilde{\mathbf{a}}$ 
25:  end while
26:  if  $m \geq m_{\max}$  then
27:    for  $i = 2, \dots, m_{\min}$  do
28:       $\mathbf{v}_i \leftarrow \mathbf{V} \mathbf{s}_i, \mathbf{v}_i^A \leftarrow \mathbf{A} \mathbf{V} \mathbf{s}_i$ 
29:    end for
30:    Setze  $\mathbf{M}$  auf oberen  $m_{\min} \times m_{\min}$  Block von  $\mathbf{T}$ 
31:     $\mathbf{v}_1 \leftarrow \mathbf{u}, \mathbf{v}_1^A \leftarrow \mathbf{u}^A, m \leftarrow m_{\min}$ 
32:  end if
33:   $\tilde{\mathbf{Q}} \leftarrow [\mathbf{Q}, \mathbf{u}]$ 
34:  Berechne Näherungslösung für  $\mathbf{t} \perp \tilde{\mathbf{Q}}$  aus
     $(\mathbf{I} - \tilde{\mathbf{Q}} \tilde{\mathbf{Q}}^H)(\mathbf{A} - \theta \mathbf{I})(\mathbf{I} - \tilde{\mathbf{Q}} \tilde{\mathbf{Q}}^H) \mathbf{t} = -\mathbf{r}$ 

```

$\mathbf{V}_m = [\mathbf{v}_1, \dots, \mathbf{v}_m]$ gilt. Analog dazu besteht \mathbf{V}_m^A aus den Spalten \mathbf{v}_i^A , so dass $\mathbf{V}_m^A = [\mathbf{v}_1^A, \dots, \mathbf{v}_m^A]$ gilt.

- (11) Die SCHUR-Zerlegung nach Satz 4.2.1 der vollbesetzten Matrix $\mathbf{M} \in \mathbb{C}^{m \times m}$ ist so durchzuführen, dass in jedem Iterationsschritt eine Sortierung gewährleistet wird, bei welcher der Eintrag $\mathbf{R}_{1,1}$ am nächsten am Zielwert τ liegt. Wenn $m \geq m_{\max}$, steht eine Reduktion des Unterraums an und die m_{\max} dem Zielwert τ am nächsten liegenden RITZ-Werte müssen an die ersten m_{\max} Positionen der Diagonalen von \mathbf{R} sortiert werden. $\mathbf{S} \in \mathbb{C}^{m \times m}$ besteht aus den Spalten \mathbf{s}_i .
- (13) Mit dem Stoppkriterium wird sichergestellt, dass ein normierter RITZ-Vektor akzeptiert wird, wenn die Norm des Residuums unterhalb der Toleranzschwelle ϵ liegt.
- (19) Nachdem ein konvergiertes RITZ-Paar als Lösung akzeptiert ist, geht die Suche nach dem nächsten RITZ-Paar weiter. Dazu wird der Unterraum entsprechend aktualisiert.
- (25) Wenn die Dimension des Unterraums m_{\max} erreicht, wird dieser wieder auf die Dimension m_{\min} reduziert. Dazu werden die m_{\min} RITZ-Vektoren verwendet, deren RITZ-Wert dem Zielwert τ am nächsten liegen.
- (33) Die bereits konvergierten RITZ-Vektoren werden in \mathbf{Q} zusammengefasst und $\tilde{\mathbf{Q}}$ ist \mathbf{Q} erweitert um die aktuelle Approximation \mathbf{u} . Dann kann die Korrekturgleichung gelöst werden, deren Ergebnis $\mathbf{t} \perp \tilde{\mathbf{Q}}$ ist.

Dass Eigenwerte von Algorithmus 4.9 nicht gefunden werden, kann theoretisch vorkommen, wenn \mathbf{V}_m orthogonal zu einem gesuchten Eigenvektor ist. Aufgrund der während der Iteration auftretenden Rundungsfehler und Startvektoren mit zufällig verteilten Einträgen, ist dies in der Praxis jedoch sehr unwahrscheinlich.

4.3.6. Berechnung innerer Eigenwerte nicht-hermitescher Matrizen

Die RITZ-Werte in Algorithmus 4.9 (JDQR) konvergieren typischerweise gegen äußere Eigenwerte. Eigenwerte im Innern des Spektrums lassen sich zwar mit Algorithmus 4.9 finden, aber die Konvergenzeigenschaften sind oft nicht zufriedenstellend. Ein Ausweg bietet sich mit den *harmonischen* RITZ-Werten an [70, 79], welche die inversen RITZ-Werte von \mathbf{A}^{-1} sind. Hierzu werden zwei verschiedene orthogonale Unterräume verwendet: Der Suchraum \mathbf{V}_m und der Testraum $\mathbf{W}_m = \mathbf{A}\mathbf{V}_m$.

Definition 4.3.2 (Harmonische Ritz-Werte): Gegeben sei eine Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$. Sei $\mathbf{V}_m \in \mathbb{C}^{n \times m}$ mit $1 < m \ll n$ Orthonormalbasis des m -dimensionalen Unterraums von $\mathbb{C}^{n \times n}$ und die m Spalten von \mathbf{V}_m sind $\mathbf{v}_1, \dots, \mathbf{v}_m$. Der Wert $\tilde{\theta}_m \in \mathbb{C}$ ist ein harmonischer RITZ-Wert von \mathbf{A} bezüglich des Unterraums $\mathbf{W}_m := \mathbf{A}\mathbf{V}_m$, wenn

$$\mathbf{A}\tilde{\mathbf{u}}_m - \tilde{\theta}_m\tilde{\mathbf{u}}_m \perp \mathbf{A}\mathbf{V}_m \quad \text{für} \quad \tilde{\mathbf{u}}_m \in \mathbf{V}_m, \tilde{\mathbf{u}}_m \neq \mathbf{0} \quad (4.16)$$

erfüllt ist.

Die m Spalten von \mathbf{W}_m sind $\mathbf{w}_1, \dots, \mathbf{w}_m$ und $\tilde{\mathbf{M}}_m := (\mathbf{W}_m^H \mathbf{V}_m)^{-1} \mathbf{W}_m^H \mathbf{A} \mathbf{V}_m$, so dass (4.16) äquivalent zu dem reduzierten $m \times m$ -Eigenwertproblem

$$\tilde{\mathbf{M}}_m \mathbf{s} = \tilde{\theta}_m \mathbf{s} \quad \text{für} \quad \mathbf{s} \in \mathbb{C}^m, \mathbf{s} \neq \mathbf{0}, \tilde{\mathbf{u}}_m = \mathbf{V}_m \mathbf{s}$$

ist. Die Eigenwerte der $m \times m$ -Matrix $\widetilde{\mathbf{M}}_m$ sind die harmonischen RITZ–Werte von \mathbf{A} .

Elegant an dieser Formulierung ist die Tatsache, dass die Inverse von \mathbf{A} nicht benötigt wird und auch kein lineares Gleichungssystem gelöst werden muss. In der Praxis muss das reduzierte allgemeine Eigenwertproblem der Form

$$\mathbf{W}_m^H \mathbf{A} \mathbf{V}_m \mathbf{s} - \theta \mathbf{W}_m^H \mathbf{V}_m \mathbf{s} = 0 \quad (4.17)$$

gelöst werden, während die Verwendung der Standard–RITZ–Werte nach Definition 4.3.1 in Algorithmus 4.9 nur auf ein niederdimensionales Standardeigenwertproblem führen.

In Algorithmus 4.10 sind die entsprechenden Erweiterungen von Algorithmus 4.9 enthalten, die zur Berechnung harmonischer RITZ–Werte notwendig sind [46, 56, 86]. Die Standard–RITZ–Werte konvergieren typischerweise gegen äußere Eigenwerte von \mathbf{A} . Analog dazu konvergieren die inneren harmonischen RITZ–Werte der verschobenen Matrix $\mathbf{A} - \tau \mathbf{I}$ typischerweise gegen Eigenwerte $\lambda \neq \tau$ in der Nähe der Verschiebung τ . Der Suchraum \mathbf{V}_m ist für verschobene und unverschobene Matrizen identisch, was zur Berechnung der harmonischen RITZ–Werte benutzt werden kann. Aus Gründen der numerischen Stabilität sind \mathbf{V}_m und \mathbf{W}_m orthonormal und es gilt $(\mathbf{A} - \tau \mathbf{I}) \mathbf{V}_m = \mathbf{W}_m \mathbf{M}_m^A$ mit der oberen Dreiecksmatrix \mathbf{M}_m^A . Die Berechnung der Eigenwerte von (4.17) erfolgt über eine partielle verallgemeinerte SCHUR–Form, analog dem QZ–Verfahren aus Abschnitt 4.2.3.1. Daher wird Algorithmus 4.10 im Folgenden auch mit JDQZ referenziert.

Zur Diskussion der Abschnitte von Algorithmus 4.10 (JDQZ) mit Referenzen auf die jeweiligen Zeilennummern:

- (1) Initialisierung.
- (3) Der Korrekturvektor \mathbf{t} wird mittels Algorithmus A.1 zum Suchraum \mathbf{V}_m bestehend aus den Vektoren \mathbf{v}_i orthogonalisiert. $\mathbf{V}^A := \mathbf{A} \mathbf{V} - \tau \mathbf{V}$ und \mathbf{v}_m^A ist der Vektor um den \mathbf{V}^A erweitert wird.
- (7) Der Erweiterungsvektor für \mathbf{W} berechnet sich aus der Orthogonalisierung mittels Algorithmus A.1 von \mathbf{v}_m^A bezüglich der berechneten SCHUR–Vektoren \mathbf{q}_i , die in der Matrix \mathbf{Q} zusammengefasst sind.
- (10)–(15) Aktualisierung der projizierten $m \times m$ -Matrizen $\mathbf{M} = \mathbf{W}^H \mathbf{V}$ und $\mathbf{M}^A = \mathbf{W}^H \mathbf{V}^A$.
- (18) Berechnung der QZ–Zerlegung (bzw. der verallgemeinerten SCHUR–Form nach Satz 4.2.1) des Matrixbüschels $(\mathbf{M}^A, \mathbf{M})$: $\mathbf{M}^A \mathbf{S}^R = \mathbf{S}^L \mathbf{T}^A$ sowie $\mathbf{M} \mathbf{S}^R = \mathbf{S}^L \mathbf{T}$, wobei \mathbf{S}^R und \mathbf{S}^L unitär und \mathbf{T}^A und \mathbf{T} obere Dreiecksmatrizen sind. Die QZ–Zerlegung ist so durchzuführen, dass in jedem Iterationsschritt eine Sortierung gewährleistet wird, bei welcher der Eintrag $|\mathbf{T}_{1,1}^A / \mathbf{T}_{1,1}|$ minimal wird. Wenn $m \geq m_{\max}$ gilt, steht eine Reduktion des Unterraums an und die m_{\max} kleinsten harmonischen RITZ–Werte von $|\mathbf{T}_{1,1}^A / \mathbf{T}_{1,1}|$ müssen an die ersten m_{\max} Positionen der Diagonalen von \mathbf{T}^A und \mathbf{T} sortiert werden. $\mathbf{S}^R \in \mathbb{C}^{m \times m}$ besteht aus den Spalten \mathbf{s}_i^R und \mathbf{S}^L entsprechend aus den Spalten \mathbf{s}_i^L .
- (19) An dieser Stelle wird der RAYLEIGH–Quotient ϑ (statt des harmonischen RITZ–Werts θ) aus den harmonischen RITZ–Vektor \mathbf{u} berechnet. Grund dafür ist die Forderung nach $\mathbf{r} = (\mathbf{A} - \tau \mathbf{I}) \mathbf{u} - \vartheta \mathbf{u} \perp \mathbf{u}$ anstatt $\perp \mathbf{W}$. $\overline{\mathbf{T}}_{1,1}$ ist konjugiert komplex zu $\mathbf{T}_{1,1}$.

Algorithmus 4.10 JACOBI-DAVIDSON-Verfahren nach [56] zur Berechnung l_{\max} innerer Eigenwerte des Standardeigenwertproblems (JDQZ)

Eingabe: Matrix \mathbf{A} , Startvektor \mathbf{v}_0 , Toleranz ϵ , Zielwert τ , Anzahl gesuchter Eigenpaare l_{\max}

Ausgabe: \mathbf{Q} , \mathbf{R}

```

1:  $\mathbf{t} \leftarrow \mathbf{v}_0$ ,  $l \leftarrow 0$ ,  $m \leftarrow 0$ ,  $\mathbf{Q} \leftarrow []$ ,  $\mathbf{R} \leftarrow []$ 
2: while  $l < l_{\max}$  do                                     % gesuchte Anzahl an Eigenpaaren noch nicht gefunden
3:   for  $i = 1, \dots, m$  do
4:      $\mathbf{t} \leftarrow \mathbf{t} - (\mathbf{v}_i^H \mathbf{t}) \mathbf{v}_i$ 
5:   end for
6:    $m \leftarrow m + 1$ ,  $\mathbf{v}_m \leftarrow \mathbf{t} / \|\mathbf{t}\|$ ,  $\mathbf{v}_m^A \leftarrow \mathbf{A} \mathbf{v}_m - \tau \mathbf{v}_m$ ,  $\mathbf{w} \leftarrow \mathbf{v}_m^A$ 
7:   for  $i = 1, \dots, l$  do
8:      $\mathbf{w} \leftarrow \mathbf{w} - (\mathbf{q}_i^H \mathbf{w}) \mathbf{q}_i$ 
9:   end for
10:  for  $i = 1, \dots, m - 1$  do
11:     $\mathbf{M}_{i,m}^A \leftarrow \mathbf{w}^H \mathbf{w}$ ,  $\mathbf{w} \leftarrow \mathbf{w} - \mathbf{M}_{i,m}^A \mathbf{w}_i$ 
12:  end for
13:   $\mathbf{M}_{m,m}^A \leftarrow \|\mathbf{w}\|$ ,  $\mathbf{w}_m \leftarrow \mathbf{w} / \|\mathbf{w}\|$ 
14:  for  $i = 1, \dots, m - 1$  do
15:     $\mathbf{M}_{i,m} \leftarrow \mathbf{w}_i^H \mathbf{v}_m$ ,  $\mathbf{M}_{m,i} \leftarrow \mathbf{w}_m^H \mathbf{v}_i$ 
16:  end for
17:   $\mathbf{M}_{m,m} \leftarrow \mathbf{w}_m^H \mathbf{v}_m$ 
18:  Berechnung der verallgemeinerten SCHUR-Zerlegung der Form  $\mathbf{M}^A \mathbf{S}^R = \mathbf{S}^L \mathbf{T}^A$ ,
     $\mathbf{M} \mathbf{S}^R = \mathbf{S}^L \mathbf{T}$ . Sortiere dabei aufsteigend, so dass  $|\mathbf{T}_{i,i}^A / \mathbf{T}_{i,i}| \leq |\mathbf{T}_{i+1,i+1}^A / \mathbf{T}_{i+1,i+1}|$ 
19:   $\mathbf{u} \leftarrow \mathbf{V} \mathbf{s}_1^R$ ,  $\mathbf{u}^A \leftarrow \mathbf{V}^A \mathbf{s}_1^R$ ,  $\vartheta \leftarrow \overline{\mathbf{T}_{1,1}} \mathbf{T}_{1,1}^A$ ,  $\mathbf{r} \leftarrow \mathbf{u}^A - \vartheta \mathbf{u}$ ,  $\tilde{\mathbf{a}} \leftarrow \mathbf{Q}^H \mathbf{r}$ ,  $\tilde{\mathbf{r}} \leftarrow \mathbf{r} - \mathbf{Q} \tilde{\mathbf{a}}$ 
20:  while  $\|\tilde{\mathbf{r}}\| \leq \epsilon$  do                                     % Konvergenz?
21:     $\mathbf{R} \leftarrow \begin{bmatrix} \mathbf{R} & \tilde{\mathbf{a}} \\ \mathbf{0} & \vartheta + \tau \end{bmatrix}$ ,  $\mathbf{Q} \leftarrow [\mathbf{Q}, \mathbf{u}]$ ,  $l \leftarrow l + 1$ 
22:    if  $l = l_{\max}$  then
23:      Stoppe
24:    end if
25:     $m \leftarrow m - 1$ 
26:    for  $i = 1, \dots, m$  do
27:       $\mathbf{v}_i \leftarrow \mathbf{V} \mathbf{s}_{i+1}^R$ ,  $\mathbf{v}_i^A \leftarrow \mathbf{V}^A \mathbf{s}_{i+1}^R$ ,  $\mathbf{w}_i \leftarrow \mathbf{W} \mathbf{s}_{i+1}^L$ ,  $\mathbf{s}_i^R \leftarrow \mathbf{s}_i^L \leftarrow \mathbf{e}_i$ 
28:    end for
29:    Setze  $\mathbf{M}$  bzw.  $\mathbf{M}^A$  auf unteren  $m \times m$  Block von  $\mathbf{T}$  bzw.  $\mathbf{T}^A$ 
30:     $\mathbf{u} \leftarrow \mathbf{v}_1$ ,  $\mathbf{u}^A \leftarrow \mathbf{v}_1^A$ ,  $\vartheta \leftarrow \overline{\mathbf{M}_{1,1}} \mathbf{M}_{1,1}^A$ ,  $\mathbf{r} \leftarrow \mathbf{u}^A - \vartheta \mathbf{u}$ ,  $\tilde{\mathbf{a}} \leftarrow \mathbf{Q}^H \mathbf{r}$ ,  $\tilde{\mathbf{r}} \leftarrow \mathbf{r} - \mathbf{Q} \tilde{\mathbf{a}}$ 
31:  end while
32:  if  $m \geq m_{\max}$  then
33:    for  $i = 2, \dots, m_{\min}$  do
34:       $\mathbf{v}_i \leftarrow \mathbf{V} \mathbf{s}_i^R$ ,  $\mathbf{v}_i^A \leftarrow \mathbf{V}^A \mathbf{s}_i^R$ ,  $\mathbf{w}_i \leftarrow \mathbf{W} \mathbf{s}_i^L$ 
35:    end for
36:    Setze  $\mathbf{M} \leftarrow \mathbf{T}(1 : m_{\min}, 1 : m_{\min})$  und  $\mathbf{M}^A \leftarrow \mathbf{T}^A(1 : m_{\min}, 1 : m_{\min})$ 
37:     $\mathbf{v}_1 \leftarrow \mathbf{u}$ ,  $\mathbf{v}_1^A \leftarrow \mathbf{u}^A$ ,  $\mathbf{w}_1 \leftarrow \mathbf{W} \mathbf{s}_1^L$ ,  $m \leftarrow m_{\min}$ 
38:  end if
39:   $\theta \leftarrow \vartheta + \tau$ ,  $\tilde{\mathbf{Q}} \leftarrow [\mathbf{Q}, \mathbf{u}]$ 
40:  Berechne Näherungslösung für  $\mathbf{t} \perp \tilde{\mathbf{Q}}$  aus  $(\mathbf{I} - \tilde{\mathbf{Q}} \tilde{\mathbf{Q}}^H)(\mathbf{A} - \theta \mathbf{I})(\mathbf{I} - \tilde{\mathbf{Q}} \tilde{\mathbf{Q}}^H) \mathbf{t} = -\mathbf{r}$ 
41: end while

```

- (20) Mit dem Stoppkriterium wird sichergestellt, dass ein normierter RITZ–Vektor akzeptiert wird, wenn die Norm des Residuums unterhalb der Toleranzschwelle ϵ liegt.
- (25) Neustart nach erfolgter Akzeptanz. Dazu wird der Unterraum entsprechend aktualisiert.
- (32) Wenn die Dimension des Unterraums m_{\max} erreicht, wird dieser wieder auf die Dimension m_{\min} reduziert.
- (39) Die bereits konvergierten RITZ–Vektoren werden in \mathbf{Q} zusammengefasst. $\tilde{\mathbf{Q}}$ ist \mathbf{Q} erweitert um die aktuelle Approximation \mathbf{u} . Somit kann die Korrekturgleichung gelöst werden, deren Ergebnis $\mathbf{t} \perp \tilde{\mathbf{Q}}$ ist.

4.3.7. Lösung der Korrekturgleichung

Die Korrekturgleichung im JACOBI–DAVIDSON–Verfahren (4.7) dient zur Bestimmung eines Korrekturvektors \mathbf{t} , mit dem der Unterraum erweitert wird, um sukzessive eine bessere Approximation eines Eigenpaars durch ein RITZ–Paar zu erreichen. Dabei soll $\mathbf{t} \perp \mathbf{u}$ gelten, wenn \mathbf{u} der RITZ–Vektor der zuletzt berechneten Approximation darstellt. Generell gibt es mehrere Möglichkeiten, die Korrekturgleichung zu modifizieren, ohne dass die Konvergenz verloren geht. In [57] werden Möglichkeiten gezeigt, mit denen ein KRYLOW–Unterraum während der Iteration erhalten werden kann. In [54] werden verschiedene Vereinfachungen vorgestellt, die im dort betrachteten Beispiel dennoch zu Konvergenz führen.

Der in [79] enthaltene Vorschlag zur Lösung der Korrekturgleichung (4.7) wurde in [65] auf polynomiale Eigenwertprobleme des Matrixpolynoms $\Psi(\lambda)$ erweitert. Er sieht Umformungen vor, welche den vollbesetzten Term $(\mathbf{I} - \mathbf{u}\mathbf{u}^H)$ in (4.7) vermeiden. So lässt sich (4.7) mit $\mathbf{p} = \Psi'(\theta)\mathbf{u}$ schreiben als

$$\mathbf{t} = \underbrace{-\Psi(\theta)^{-1}\mathbf{r}}_{\mathbf{t}_1} + \epsilon_O \underbrace{\Psi(\theta)^{-1}\mathbf{p}}_{\mathbf{t}_2}. \quad (4.18)$$

Die ϵ_O -gewichtete Summe der Lösungen \mathbf{t}_1 und \mathbf{t}_2 der linearen Gleichungssysteme $\Psi(\theta)\mathbf{t}_1 = -\mathbf{r}$ und $\Psi(\theta)\mathbf{t}_2 = \mathbf{p}$ ergibt dabei die Lösung von (4.7), ohne dass vollbesetzte Matrizen auftreten. Für den Koeffizienten ϵ_O gilt entsprechend [65]

$$\epsilon_O = \frac{\mathbf{u}^H \Psi(\theta)^{-1} \mathbf{r}}{\mathbf{u}^H \Psi(\theta)^{-1} \mathbf{p}}. \quad (4.19)$$

Darüber hinaus darf $\Psi(\theta)^{-1}$ durch einen geeigneten Vorkonditionierer ersetzt werden, der eine Näherung darstellt, aber leichter invertierbar ist als $\Psi(\theta)$.

Zur Lösung der linearen Gleichungssysteme (4.7) und (4.18) gibt es verschiedene Verfahren. Jegliche Verfahren müssen für komplex unsymmetrische lineare Gleichungssysteme geeignet sein und müssen darüber hinaus mit der unter Umständen schlechten Kondition der zu lösenden Systeme zurecht kommen. Im Folgenden werden die verwendeten Löser kurz benannt.

4.4. Löser für lineare Gleichungssysteme

Direkte Löser für dünn-besetzte Matrizen führen zu komplizierteren Algorithmen als direkte Löser für dicht-besetzte Matrizen. Hauptgrund ist die Notwendigkeit, den auftretenden Fill-In bei der LU-Faktorisierung zu begrenzen. Daher bestehen direkte Löser für lineare Gleichungssysteme dünn-besetzter Matrizen typischerweise aus vier Schritten [46]:

1. Im Umsortierungsschritt werden Zeilen und Spalten vertauscht, so dass die Besetzung der Matrixfaktoren so wenig dichter wie möglich wird.
2. Im Analyseschritt wird die Besetzungsstruktur der Faktoren bestimmt und passende Datenstrukturen erzeugt.
3. In der numerischen Faktorisierung werden die Faktoren \mathbf{L} und \mathbf{U} berechnet.
4. Im Lösungsschritt wird durch Vorwärts- und Rückwärtssubstitution die Lösung berechnet.

Für jeden Schritt gibt es eine Vielzahl an Algorithmen, die der Spezialliteratur zu entnehmen sind. Für die genaue Lösung von linearen Gleichungssystemen sollte zuerst die Verwendung von direkten Löser in Betracht gezogen werden. Wenn mehr als ein Gleichungssystem gelöst werden muss, amortisiert sich der hohe numerische Aufwand für eine dünn-besetzte LU-Zerlegung gegebenenfalls schnell.

Sofern keine hochgradig genaue Lösung benötigt wird oder direkte Löser aufgrund von Speicher- oder Rechenzeitbeschränkungen zu aufwendig sind, kann die Verwendung iterativer Gleichungssystemlöser in Betracht gezogen werden. Eine Übersicht über iterative Gleichungssystemlöser ist in [47] enthalten. Die linearen Gleichungssysteme, die bei der Berechnung von Eigenwertproblemen auftreten, haben aufgrund der involvierten spektralen Verschiebung σ (siehe Algorithmen 4.3 und 4.4) typischerweise eine indefinite Systemmatrix, so dass sie positive und negative Eigenwerte besitzt. Konvergenzprobleme sind dabei insbesondere bei spektralen Verschiebungen zu erwarten, die zu inneren Eigenwerten gehören [46]. Darüber hinaus führen spektrale Verschiebungen, die nah an einem Eigenwert liegen, zu einer annähernd singulären Systemmatrix. Diese quasi-singulären Gleichungssysteme bereiten den iterativen Verfahren große Schwierigkeiten. Daher ist gemeinhin bekannt, dass iterative Verfahren mit effizienten Vorkonditionierern kombiniert werden müssen, um gute Konvergenzeigenschaften aufzuweisen. Jedoch ist die Konstruktion von Vorkonditionierern für indefinite Matrizen selbst schon eine schwierige Aufgabe [46].

4.4.1. Direkte Löser

Als direkter Löser für die linearen Gleichungssysteme kommt die in der *Math Kernel Library* von INTEL [124] enthaltene Implementierung von PARDISO (Parallel Sparse Direct Solver) [74] zum Einsatz. In PARDISO gibt es spezielle Prozeduren für struktur-symmetrisch besetzte Matrizen, zu denen die Matrizen der FIT-Formulierungen aus Kapitel 3 zählen.

4.4.2. Iterative Löser

Zur Lösung der Korrekturgleichung (4.7) im JACOBI–DAVIDSON–Verfahren können iterative Löser für lineare Gleichungssysteme zum Einsatz kommen. Iterative Löser kommen zunächst ohne die Faktorisierung der Matrizen aus, so dass der Hauptaufwand in Matrixvektormultiplikationen steckt. Zur Auswahl eines iterativen Gleichungssystemlösers bieten sich beispielsweise die Empfehlungen im Schema [47, Anhang D] an. Für die unsymmetrischen komplexen Matrizen fällt die Wahl auf das BiCGStab–Verfahren (engl. BiConjugate Gradient Stabilized), da Experimente mit den denkbaren Alternativen CGS (engl. conjugate gradient squared) und GMRes (engl. generalized minimal residual) zu jeweils schlechterem Konvergenzverhalten führten. Herleitung und Diskussion des BiCGStab–Verfahrens befinden sich in [71, 73]. Eine Erweiterung stellt das BiCGStab(ℓ)-Verfahren aus [76] dar. Das Konvergenzverhalten lässt sich oft durch die Verwendung von Vorkonditionierern verbessern, die im folgenden Abschnitt beschrieben werden.

4.4.3. Vorkonditionierung

Iterative Löser für lineare Gleichungssysteme konvergieren insbesondere dann sehr schnell, wenn die Systemmatrix \mathbf{A} näherungsweise eine Einheitsmatrix \mathbf{I} ist. Da dies im Allgemeinen nicht der Fall ist, brauchen die iterativen Löser für lineare Gleichungssysteme oft sehr viele Iterationen bis die gesuchte Lösung in der geforderten Genauigkeit erzielt wird. Eine Möglichkeit, das Konvergenzverhalten iterativer Löser für lineare Gleichungssysteme zu verbessern, ist die sogenannte Vorkonditionierung. Dazu ist eine Matrix \mathbf{M} erforderlich, welche die Systemmatrix $\mathbf{A} \approx \mathbf{M}$ hinreichend gut approximiert und darüber hinaus leicht invertierbar sein muss. Vorkonditionierung durch Multiplikation von links ergibt näherungsweise $\mathbf{M}^{-1}\mathbf{A} \approx \mathbf{I}$. Die Vorkonditionierung verursacht, dass in jedem Iterationsschritt (mindestens) ein lineares Gleichungssystem gelöst wird, da \mathbf{M}^{-1} aus Speicherplatzgründen nicht in Matrixform aufgestellt wird. Das vorkonditionierte System lautet dann

$$\mathbf{M}^{-1}\mathbf{A}\mathbf{x} = \mathbf{M}^{-1}\mathbf{b}. \quad (4.20)$$

Der Aufwand an Rechenzeitressourcen, der zur Erstellung und Verwendung eines Vorkonditionierers notwendig ist, darf selbstverständlich nicht größer sein, als der langsame Konvergenzverlauf des jeweiligen iterativen Lösers für lineare Gleichungssysteme ohne Vorkonditionierer. Darüber hinaus sollte der Speicherbedarf des Vorkonditionierers kontrolliert werden können.

4.4.3.1. LU

Für die besonderen Voraussetzungen, dass es sich bei der zugrunde liegenden Problemstellung um die Lösung relativ kleiner linearer Gleichungssysteme handelt bzw. dass ein hinreichend großer Arbeitsspeicher vorhanden ist und im Voraus bekannt ist, dass sehr viele lineare Gleichungssysteme des gleichen Typs gelöst werden müssen, kann eine vollständige LU–Zerlegung³⁹ der Form $\mathbf{PAQ} = \mathbf{LU}$ ihre Berechtigung als Vorkonditionierer haben. Dabei ist \mathbf{L} eine untere und \mathbf{U} eine obere Dreiecksmatrix. \mathbf{P} und \mathbf{Q} sind Permutationsmatrizen. Zur

³⁹ von engl. lower upper (triangular matrix)

Berechnung der LU-Zerlegung existiert das Paket UMFPACK [51], das für SMP-Rechner⁴⁰ gut parallelisiert vorliegt.

4.4.3.2. ILU

Da der Speicherbedarf der LU-Zerlegung bei steigender Dimension der Systemmatrix \mathbf{A} unverhältnismäßig stark ansteigt, lohnt sich der Blick auf Alternativen, die beispielsweise mit der Klasse der unvollständigen LU-Zerlegungen ILU (engl. incomplete LU) existiert. Eine ausführliche Herleitung sowie ein Vergleich verschiedener ILU-Varianten befindet sich in [73].

Da bei der herkömmlichen LU-Zerlegung alle Einträge berücksichtigt werden, kann es passieren, dass Rechenzeit und Speicherplatz für Einträge verwendet wird, die sehr klein sind und somit für die Approximationseigenschaften der Zerlegung von untergeordneter Bedeutung sind. In dieser Arbeit wird die Methode ILUT verwendet, deren Speicherbedarf sich mit dem Schwellwert t_{ILU} steuern lässt, in dem Einträge von \mathbf{A} nicht berücksichtigt werden, wenn sie kleiner als t_{ILU} sind.

4.4.4. Umindizierung der FIT-Matrizen für ILU-Zerlegungen

Die exakte Zerlegung von schwachbesetzten Matrizen bringt im Allgemeinen Faktoren hervor, die nicht annähernd so schwach besetzt sind, wie die Ausgangsmatrix. Um die Anzahl der durch die Faktorisierung zusätzlich entstehenden Einträge gering zu halten, haben sich Umsortierungsalgorithmen bewährt, mit denen Zeilen und Spalten einer Matrix vor der Faktorisierung vertauscht werden [73]. Während in den direkten Verfahren zur Lösung linearer Gleichungssysteme entsprechende Algorithmen integriert sind, ist dies bei der ILU im Allgemeinen nicht der Fall. Die beiden in MATLAB 2009b[126] verfügbaren Verfahren, die sich für die FIT-Formulierungen aus Kapitel 3 als günstig erwiesen haben, sind

- Reverse CUTHILL-MCKEE ordering (`symrcm`) [49] und
- Column approximate minimum degree permutation (`colamd`) [52].

Um die Anzahl der bei der ILUT zusätzlich auftretenden Einträge (engl. *Fill-In*) in \mathbf{L} und \mathbf{U} zu begrenzen, sollte die Systemmatrix \mathbf{A} immer einer Umsortierung unterzogen werden. Der *Fill-In* berechnet sich als Verhältnis der Anzahl der von Null verschiedenen Einträge von $\mathbf{L} + \mathbf{U}$ zur Anzahl der von Null verschiedenen Einträge der Systemmatrix \mathbf{A} . Abbildung 4.2 zeigt die zusätzlichen Einträge in Abhängigkeit der Matrixdimension, der Umindizierung und dem Schwellwert t_{ILU} . Abbildung 4.2a liegt die reelle FIT-Systemmatrix nach Gleichung (2.93) zugrunde, während Abbildung 4.2b auf einer komplexen FIT-Systemmatrix nach Gleichung (3.42) basiert. Es zeigt sich, dass die Umindizierung mit `symrcm` stets einen geringeren *Fill-In* als die Umindizierung mit `colamd` zur Folge hat.

⁴⁰In Rechnern mit *shared memory processors* (SMP) haben mehr als ein Prozessor Zugriff auf einen gemeinsamen Arbeitsspeicher, so dass der Datenaustausch durch Zugriffe auf unterschiedliche Bereiche im Arbeitsspeicher organisiert werden kann.

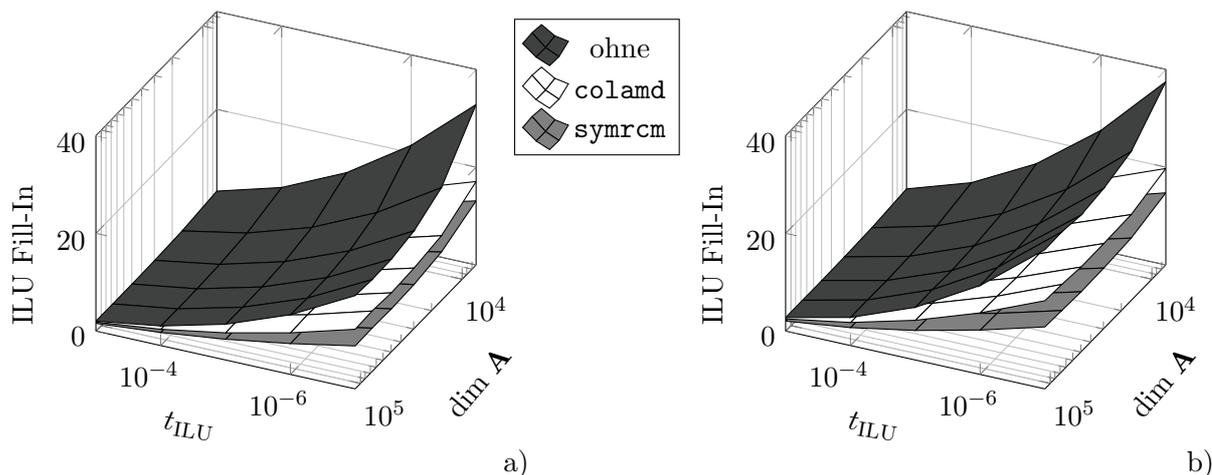


Abbildung 4.2.: Anzahl der zusätzlichen Einträge $\text{nnz}(\mathbf{L} + \mathbf{U})/\text{nnz}(\mathbf{A})$ in den Faktoren der ILU–Zerlegung bei Umindizierung der FIT–Systemmatrix des reellen Eigenwertproblems (2.93) (a) und des komplexen Eigenwertproblems (3.42) (b).

4.5. Erweiterung des Selektionsprozesses im Jacobi–Davidson–Verfahren

Für viele Eigenwertprobleme, die aus Wissenschaft und Technik resultieren, sind oft a priori Informationen über das zu berechnende Ergebnis bekannt. Diese Informationen helfen, eine Schätzung für den Zielwert τ oder einen passenden Startvektor für den Eigenwertlöser zu generieren. Da diese Schätzungen im Allgemeinen ungenau sind, ist auch unbekannt, ob die Schätzungen nah am richtigen Eigenpaar liegen oder nicht. Das kann dazu führen, dass in der Umgebung von τ mehrere Eigenpaare berechnet werden müssen, an denen kein Interesse besteht, bevor die gesuchten Eigenpaare schließlich gefunden werden. Dies kostet sowohl Rechenzeit als auch Speicherplatz, was besonders bei sehr großen Eigenwertproblemen ins Gewicht fällt. Andererseits sind bei den gesuchten Eigenvektoren, welche eine elektromagnetische Feldverteilung im Raum repräsentieren, oft einige Eigenschaften bekannt, die sie von den unerwünschten Eigenvektoren unterscheiden. Das kann beispielsweise eine bestimmte Feldverteilung in gewissen Regionen oder auch ein Maximum in einer bestimmten Region sein. Die Idee ist, die RITZ–Vektoren, die als Eigenvektor–Approximationen während der Iterationen des JACOBI–DAVIDSON–Verfahrens auftreten, daraufhin zu untersuchen, ob sie die Eigenschaften der gesuchten Eigenvektoren bereits zu einem gewissen Grad erfüllen. In Abschnitt 4.5.1 werden zunächst zwei Selektionskriterien eingeführt, mit denen sich die gesuchten von den unerwünschten Eigenvektoren unterscheiden lassen. Anschließend folgt in Abschnitt 4.5.2 die Darstellung der konkreten Implementierung des erweiterten Selektionsprozesses für den JDQR (Algorithmus 4.9) und den JDQZ (Algorithmus 4.10).

4.5.1. Selektionskriterien zur gezielten Auswahl gesuchter Eigenvektoren

Im Folgenden werden zwei Selektionskriterien betrachtet. Das erste Selektionskriterium findet seine Anwendung bei der Berechnung von Wellenleitermoden. Dabei soll typischerweise der fundamentale Wellenleitermode und gegebenenfalls einige benachbarte höhere Wellenlei-

termen berechnet werden. Beispielsweise können die verwendeten Randbedingungen die Ursache dafür sein, dass der fundamentale Wellenleitermode nicht mehr durch eine extremale Sortierbedingung während der Eigenwertberechnung identifiziert werden kann, so dass die Verwendung des erweiterten Selektionskriteriums notwendig wird. Ein weiteres Selektionskriterium wird vorgestellt, das seine Anwendung bei der Berechnung von Eigenmoden höherer Ordnung von dreidimensionalen Strukturen findet. Da es sich um Eigenmoden höherer Ordnung handelt, sind benachbarte Eigenmoden nicht durch die Modellbildung bedingt, sondern struktur-immanent.

4.5.1.1. Physikalisch motivierte Selektion bei Berechnung von Wellenleitermoden

Grundannahme ist, dass bei der Berechnung von Wellenleitermoden diejenigen Lösungen gesucht sind, die innerhalb des Wellenleiters geführt werden. Bei optischen Wellenleitern basiert die Führung der Wellenausbreitung auf einem Brechzahlunterschied zwischen dem inneren Bereich, dem Kerngebiet, und einem äußeren Bereich, dem sogenannten Mantel. Wenn die Welle im Kerngebiet geführt wird, erstreckt sich auch der Großteil des Leistungsflusses über das Kerngebiet und klingt zum äußeren Mantel hin ab. Aus diesem Vorwissen lässt sich bereits das Selektionskriterium für die gesuchten Eigenpaare ableiten. Dazu werden innerhalb jeder Iteration aus jedem RITZ-Vektor \mathbf{u}_i die normierten POYNTING-Vektoren in der Querschnittsebene A_{gesamt} des Wellenleiters berechnet. Dann wird die Komponente in Ausbreitungsrichtung (die z -Komponente) der POYNTING-Vektoren jeweils über das zuvor festgelegte Kerngebiet A_1 integriert, um die Information zu erhalten, welcher Anteil der Leistung der jeweiligen Wellenleitermodenapproximation im festgelegten Kerngebiet geführt wird, so dass sich als physikalisch motiviertes Selektionskriterium

$$\frac{\int_{A_1} S_z \, dA}{\int_{A_{\text{gesamt}}} S_z \, dA} = \frac{P_{\text{Kern}}}{P_{\text{gesamt}}} > \iota_l \quad (4.21)$$

ergibt. Sofern der Leistungsfluss im festgelegten Kerngebiet ein gewisses vordefiniertes Limit ι_l erreicht, wird der zugehörige RITZ-Vektor als aktuelle Approximation des Eigenvektors für die weitere Iteration im JACOBI-DAVIDSON-Verfahren verwendet. Dabei ist zu beachten, dass es sich stets um Approximationen der Eigenvektoren handelt, deren Residuum während der Iteration noch groß ist. Ein Anwendungsbeispiel wird in Abschnitt 5.3 gezeigt.

4.5.1.2. Selektion räumlicher Eigenmoden höherer Ordnung

Zur gezielten Selektion räumlicher Eigenmoden höherer Ordnung wird davon ausgegangen, dass sich das Verhalten der Feldverteilung mindestens in einem Untergebiet des Rechengebiets approximieren lässt. Dazu können analytische Modelle oder numerische Ergebnisse vorangegangener Simulationen herangezogen werden. Die Ausgestaltung des Selektionskriteriums für den erweiterten Selektionsprozess basiert letztlich auf dem Skalarprodukt eines Gewichtungsvektors \mathbf{f} , der die a priori vorhandenen Informationen enthält, mit denen während den JACOBI-DAVIDSON-Iterationen auftretenden RITZ-Vektoren \mathbf{u}_i

$$\mathbf{f}^H \mathbf{u}_i > \iota_l. \quad (4.22)$$

Dabei muss der Gewichtungsvektor \mathbf{f} nicht vollbesetzt sein. Ein vielversprechender RITZ-Vektor \mathbf{u}_i für die Verwendung in den nachfolgenden JACOBI-DAVIDSON-Iterationen zeichnet

sich dadurch aus, dass das Skalarprodukt in (4.22) größer als ein zuvor definierter Schwellwert ι_l ist. Ein Anwendungsbeispiel wird in Abschnitt 5.4.2 gezeigt.

4.5.2. Erweiterter Selektionsprozess

Im Folgenden wird ein heuristischer Ansatz präsentiert, der sich dazu eignet, einige wenige Eigenpaare mit a priori bekannten Eigenschaften gezielt zu berechnen. Voraussetzung für rasche Konvergenz ist, dass die a priori vorliegenden Informationen bereits in den Startvektor \mathbf{v}_0 und in den Zielwert τ eingeflossen sind. Dazu wird innerhalb des JACOBI–DAVIDSON–Verfahrens der Selektionsprozess zur Bestimmung des aktuellen RITZ–Paares um die eingangs beschriebenen Selektionskriterien erweitert. Der herkömmliche Selektionsprozess erfolgt in Zeile 11 des JDQR–Algorithmus 4.9 und in Zeile 18 des JDQZ–Algorithmus 4.10, indem jeweils derjenige RITZ–Wert selektiert wird, der den betragsmäßig geringsten Abstand zum Zielwert τ hat. Die Idee ist nun, in der Sortierung diejenigen RITZ–Werte zu bevorzugen, deren RITZ–Vektoren bestimmte Eigenschaften erfüllen, die aus der zugrunde liegenden Problemstellung herrühren und mit den eingangs beschriebenen Selektionskriterien erfasst werden können. In Algorithmus 4.11 wird dazu zunächst im Gegensatz zu JDQR–Algorithmus 4.9 und JDQZ–Algorithmus 4.10 das projizierte $m \times m$ Eigenwertproblem der Matrix \mathbf{M}_m vollständig gelöst. Daher sind in jedem Iterationsschritt m RITZ–Vektoren verfügbar, auf die der erweiterte Selektionsprozess angewendet wird.

Algorithmus 4.11 Erweiterter Selektionsprozess als modifizierte Sortierung der RITZ–Paare im JACOBI–DAVIDSON–Verfahren

Eingabe: \mathbf{M}_m als in \mathbf{V}_m projiziertes Standardeigenwertproblem mit dem Zielwert τ und der Schranke ι_l für das Selektionskriterium mit Gewichtungsvektor \mathbf{f}

Ausgabe: m RITZ–Werte \mathbf{D}_θ und m RITZ–Vektoren \mathbf{U} mit modifizierter Sortierung

- 1: Bestimme m Eigenvektoren \mathbf{S} und Eigenwerte \mathbf{D}_θ von \mathbf{M}_m .
 - 2: Sortiere die RITZ–Werte auf $\text{diag}(\mathbf{D}_\theta)$ bezüglich τ . Neue Indizierung ist I .
 - 3: Berechne RITZ–Vektoren $\mathbf{U}_\theta = \mathbf{V}_m \mathbf{S}$.
 - 4: $\mathbf{v}_P \leftarrow \mathbf{f}^H \mathbf{U}_\theta(:, I)$ Auswertung des Selektionskriteriums für alle sortierten RITZ–Vektoren.
 - 5: $K \leftarrow (\mathbf{v}_P > \iota_l)$ Speichere Indices bei denen das Selektionskriterium erfüllt ist.
 - 6: **if** $K \neq \emptyset$ **then**
 - 7: $I \leftarrow \{I(K), I(\{1, 2, \dots, m\} \setminus K)\}$ Setze die Indices $I(K)$ an die Spitze der Sortierung I .
 - 8: **end if**
 - 9: Sortiere in der Reihenfolge I : $\mathbf{U}_\theta \leftarrow \mathbf{U}_\theta(:, I)$ und $\mathbf{D}_\theta \leftarrow \mathbf{D}_\theta(I, I)$.
-

Eine Auflistung der wesentlichen Schritte von Algorithmus 4.11:

- (1) Das niederdimensionale $m \times m$ Standardeigenwertproblem von \mathbf{M}_m wird gelöst. Dies geschieht mit Standardbibliotheken.
- (2) Die RITZ–Werte, enthalten auf der Diagonalen von \mathbf{D}_θ , werden bezüglich des Zielwerts τ sortiert, und die Indices der neuen Reihenfolge werden in I abgelegt.
- (3) Berechne die RITZ–Vektoren.
- (4) Für die sortierten RITZ–Vektoren $\mathbf{U}_\theta(:, I)$ wird das Selektionskriterium ausgewertet, das beispielsweise die Konzentration von großen Feldkomponenten in bestimmten räumlichen Bereichen berücksichtigen kann.

- (5) Die Auswertung des Selektionskriteriums gibt Aufschluss darüber, bei welchen RITZ–Paaren es über der vorgegebenen Schwelle von ι_l liegt. Die Indices werden in K gespeichert.
- (6) Sofern K eine nicht-leere Menge ist, werden diejenigen Indices von I an die Spitze der Reihenfolge sortiert, welche das Selektionskriterium erfüllen.
- (9) Die RITZ–Paare werden bezüglich der neuen Sortierreihenfolge angeordnet.

Algorithmus 4.11 greift nur dann in die Sortierung ein, wenn es RITZ–Vektoren gibt, die das Selektionskriterium erfüllen. Das heißt im Umkehrschluss, dass mit der Wahl des Startvektors \mathbf{v}_0 dafür Sorge getragen werden muss, dass überhaupt RITZ–Vektoren auftreten, die das Selektionskriterium erfüllen. In der Praxis zeigt sich jedoch, dass dies stets der Fall ist, wenn sich bereits der Startvektor am Gewichtungsvektor \mathbf{f} des Selektionskriteriums orientiert. Es sei ausdrücklich darauf hingewiesen, dass das Selektionskriterium im Zusammenspiel mit der Schranke ι_l nicht zu streng (z.B. im Sinne einer exakt festgelegten Feldverteilung) auszulegen ist, da die RITZ–Vektoren zunächst nur grobe Approximationen der Eigenwerte darstellen, deren Residuum nach Definition 4.1.22 insbesondere zu Beginn der JACOBI–DAVIDSON–Iteration noch groß ist.

4.5.3. Selektion bei Berechnung äußerer Eigenwerte

In Algorithmus 4.11 wird das auf den Unterraum \mathbf{V}_m projizierte Standard eigenwertproblem für \mathbf{M}_m vollständig gelöst. In Algorithmus 4.9 wird hingegen lediglich eine SCHUR–Zerlegung entsprechend Satz 4.2.1 durchgeführt, was den numerischen Gesamtaufwand verringert. Dabei wird ausgenutzt, dass der erste SCHUR–Vektor \mathbf{q}_1 (das ist die erste Spalte von \mathbf{Q}) auch RITZ–Vektor ist, der mit dem ersten Eintrag der oberen Dreiecksmatrix \mathbf{R} das Eigenwertproblem $\mathbf{M}\mathbf{q}_1 = \mathbf{R}_{1,1}\mathbf{q}_1$ erfüllt. Die SCHUR–Zerlegung konvergiert mit kleiner werdendem strikten oberen Dreiecksanteil von \mathbf{R} gegen die Eigenwertzerlegung von \mathbf{M} , bei der \mathbf{R} schließlich Diagonalmatrix ist. Sofern der strikte obere Dreiecksanteil von \mathbf{R} Einträge enthält, die im Vergleich zur Diagonalen von \mathbf{R} (auf der die RITZ–Werte stehen) klein sind, lassen sich die RITZ–Vektoren durch die SCHUR–Vektoren approximieren und Algorithmus 4.11 in leicht modifizierter Form als erweiterter Selektionsprozess innerhalb von Algorithmus 4.9 verwenden. Es resultiert Algorithmus 4.12, der sich in folgenden Punkten von Algorithmus 4.11 unterscheidet:

- (1) Für die niederdimensionale $m \times m$ Matrix \mathbf{M}_m wird die SCHUR–Zerlegung berechnet statt eine Eigenwertzerlegung durchzuführen.
- (2) Die RITZ–Werte stehen auf der Diagonalen von \mathbf{R} . Sie werden bezüglich des Zielwerts τ sortiert, und die Indices der neuen Reihenfolge werden in I abgelegt.
- (3) Das Selektionskriterium mit Gewichtungsvektor \mathbf{f} wird für die SCHUR– statt für die RITZ–Vektoren ausgewertet.
- (8) Die SCHUR–Zerlegung wird neu sortiert. Der im JDQR–Algorithmus 4.9 für die weitere Iteration selektierte RITZ–Wert ist $\theta = \mathbf{R}_{1,1}$.

Die Sortierung der SCHUR–Zerlegung im letzten Schritt erfolgt mit den in [56] zitierten Methoden.

Algorithmus 4.12 Modifizierte Sortierung der SCHUR–Vektoren im JACOBI–DAVIDSON–Verfahren nach Algorithmus 4.9

Eingabe: \mathbf{M}_m als in \mathbf{V}_m projiziertes Standard eigenwertproblem mit Zielwert τ und Schranke ι_l für das Selektionskriterium mit Gewichtungsvektor \mathbf{f}

Ausgabe: SCHUR–Zerlegung (\mathbf{Q}, \mathbf{R}) mit modifizierter Sortierung

- 1: $[\mathbf{Q}, \mathbf{R}] \leftarrow \text{schur}(\mathbf{M}_m)$
 - 2: Sortiere die RITZ–Werte auf $\text{diag}(\mathbf{R})$ bezüglich τ und lege neue Indizierung in I ab.
 - 3: $\mathbf{v}_P \leftarrow \mathbf{f}^H \mathbf{V}_m \mathbf{Q}(:, I)$ Auswertung des Selektionskriteriums für die sortierten SCHUR–Vektoren.
 - 4: $K \leftarrow (\mathbf{v}_P > \iota_l)$ Speichere Indices in K bei denen das Selektionskriterium erfüllt ist.
 - 5: **if** $K \neq \emptyset$ **then**
 - 6: $I \leftarrow \{I(K), I(\{1, 2, \dots, m\} \setminus K)\}$ Setze die Indices $I(K)$ an die Spitze der Sortierung I .
 - 7: **end if**
 - 8: Sortiere die SCHUR–Zerlegung (\mathbf{Q}, \mathbf{R}) in der Reihenfolge I .
-

4.5.4. Selektion bei Berechnung innerer Eigenwerte

Bei der Erweiterung des Selektionsprozesses für Algorithmus 4.10 treten ähnliche Herausforderungen wie im vorigen Abschnitt auf. Da die Auswahl des jeweils aktuellen harmonischen RITZ–Werts aus einer verallgemeinerten SCHUR–Zerlegung erfolgt, tritt mit der ersten Spalte von \mathbf{S}^R entsprechend Zeile 19 in Algorithmus 4.10 nur ein einziger verallgemeinerter SCHUR–Vektor auf, der zugleich auch harmonischer RITZ–Vektor ist. Zur Erweiterung des Selektionsprozesses sind jedoch alle harmonischen RITZ–Vektoren erforderlich. Der in Algorithmus 4.13 gezeigte Ansatz basiert auf der Idee, die verallgemeinerte SCHUR–Zerlegung in Algorithmus 4.10 beizubehalten und zusätzlich die harmonischen RITZ–Paare aus der vollständigen Lösung des zugrunde liegenden niederdimensionalen allgemeinen Eigenwertproblems $(\mathbf{M} - \Theta \mathbf{M}^A) \mathbf{S} = 0$ zu bestimmen. Algorithmus 4.13 im Überblick:

- (1) Die verallgemeinerte SCHUR–Form wird bestimmt.
- (2) Die Eigenwerte, die als paarweise Quotienten der Einträge der Diagonalelemente von \mathbf{T}^A und \mathbf{T} der verallgemeinerten SCHUR–Form auftreten, werden aufsteigend sortiert.
- (3) Die RITZ–Paare werden aus der Lösung des niederdimensionalen allgemeinen Eigenwertproblems $(\mathbf{M} - \Theta \mathbf{M}^A) \mathbf{S} = 0$ bestimmt.
- (5) Das Selektionskriterium wird mit den RITZ–Paaren ausgewertet.
- (7) Sofern RITZ–Paare das Selektionskriterium erfüllen, werden diese in der Sortierung bevorzugt behandelt. Dazu wird die in Zeile 2 gefundene Sortierung I modifiziert.
- (10) Die verallgemeinerte SCHUR–Zerlegung wird bezüglich der neuen Reihenfolge sortiert.

Die Sortierung der verallgemeinerten SCHUR–Zerlegung im letzten Schritt erfolgt mit den in [56] zitierten Methoden.

Algorithmus 4.13 Modifizierte Sortierung im JACOBI–DAVIDSON–Verfahren nach Algorithmus 4.10

Eingabe: Allgemeines Eigenwertproblem mit \mathbf{M} und \mathbf{M}^A im Suchraum \mathbf{V}_m und Schranke ι für das Selektionskriterium mit Gewichtungsvektor \mathbf{f}

Ausgabe: Verallgemeinerte SCHUR-Form $(\mathbf{S}^R, \mathbf{S}^L, \mathbf{T}, \mathbf{T}^A)$ mit modifizierter Sortierung

- 1: Berechne die verallgemeinerte SCHUR–Zerlegung der Formen $\mathbf{M}^A \mathbf{S}^R = \mathbf{S}^L \mathbf{T}^A$, $\mathbf{M} \mathbf{S}^R = \mathbf{S}^L \mathbf{T}$
 - 2: Sortiere die Eigenwerte $\mathbf{T}_{i,i}^A / \mathbf{T}_{i,i}$ betragsmäßig aufsteigend. Lege neue Reihenfolge in I ab.
 - 3: $[\mathbf{S}, \Theta] \leftarrow \text{eig}(\mathbf{M}, \mathbf{M}^A)$ Löse allgemeines Eigenwertproblem. Eigenvektoren in \mathbf{S} und RITZ–Werte auf $\text{diag}(\Theta)$.
 - 4: Sortiere die RITZ–Werte in $\text{diag}(\Theta)$ aufsteigend, so dass jeweils $\Theta_{i,i} = \mathbf{T}_{i,i}^A / \mathbf{T}_{i,i}$ gilt.
 - 5: $\mathbf{v}_P \leftarrow \mathbf{f}^H \mathbf{V}_m \mathbf{S}(:, I)$ Auswertung des Selektionskriteriums für die sortierten RITZ–Vektoren.
 - 6: $K \leftarrow \{\mathbf{v}_P > \iota\}$ Speichere Indices in K bei denen das Selektionskriterium erfüllt ist.
 - 7: **if** $K \neq \emptyset$ **then**
 - 8: $I \leftarrow \{I(K), I(\{1, 2, \dots, m\} \setminus K)\}$ Setze die Indices $I(K)$ an die Spitze der Sortierung I .
 - 9: **end if**
 - 10: Sortiere die verallgemeinerte SCHUR–Zerlegung der Formen $\mathbf{M}^A \mathbf{S}^R = \mathbf{S}^L \mathbf{T}^A$, $\mathbf{M} \mathbf{S}^R = \mathbf{S}^L \mathbf{T}$ entsprechend der Reihenfolge in I .
-

4.6. Löser für polynomiale Eigenwertprobleme

In den Formulierungen in Kapitel 3 tauchen mit den Gleichungen (3.17) und (3.24) polynomiale Eigenwertprobleme auf. Eine weitere Verallgemeinerung der QZ –Zerlegung aus Abschnitt 4.2.3.1 für polynomiale Eigenwertprobleme ist nicht bekannt. Somit gibt es keinen Algorithmus, der als direkter Algorithmus bezeichnet werden könnte, der ein polynomiales Eigenwertproblem durch die Transformation auf obere Dreiecksform lösen kann. Mit der Linearisierung polynomialer Eigenwertprobleme, mit dem NEWTON–Verfahren und mit dem JACOBI–DAVIDSON–Verfahren für polynomiale Eigenwertprobleme nach [78] werden im Folgenden drei verschiedene Methoden zur Berechnung polynomialer Eigenwertprobleme eingeführt. In einem weiteren Abschnitt wird auf die Implementierung eines Löser für die Formulierung aus Gleichung (3.24) eingegangen.

4.6.1. Linearisierung polynomialer Eigenwertprobleme

Wie in Satz 4.1.8 gezeigt wird, lässt sich jedes polynomiale Eigenwertproblem $\Psi(\lambda)$ mittels Linearisierung $L(\lambda)$ als allgemeines Eigenwertproblem darstellen. In Abschnitt 4.2.3 wird gezeigt, dass die Darstellung der Linearisierung nach Gleichung (4.2) für die Anwendung der QZ –Zerlegung günstig ist, da ein allgemeines Eigenwertproblem $\mathbf{A}\mathbf{x} = \lambda\mathbf{B}\mathbf{x}$ entsteht, bei dem \mathbf{A} schon nah an der Form einer oberen HESSENBERG–Matrix und \mathbf{B} nah an der Form einer oberen Dreiecksmatrix ist. Die Koeffizientenmatrizen des allgemeinen Eigenwertproblems, das aus der Linearisierung eines polynomialen Eigenwertproblems $\Psi(\lambda)$ entsteht, haben die Dimension ℓn , wobei ℓ der Polynomgrad und n die Dimension der Koeffizientenmatrizen von $\Psi(\lambda)$ sind. Für den praktischen Gebrauch ist damit ersichtlich, dass der Linearisierung polynomialer Eigenwertprobleme enge Grenzen gesetzt sind, wenn das zugrunde liegende Matrixpolynom schon von großer Dimension ist. Für die Lösung kleiner polynomialer

Eigenwertprobleme eignet sich das Verfahren der Linearisierung jedoch gut und kann für die Lösung des auf einen orthogonalen Unterraum projizierte, niederdimensionale Eigenwertprobleme angewendet werden, wie es beispielsweise im JACOBI–DAVIDSON–Verfahren für polynomiale Eigenwertprobleme auftritt. Zu beachten ist hierbei, dass die aus der Linearisierung $L(\lambda)$ berechneten Lösungen nicht notwendigerweise in jedem Fall Lösungen des polynomialen Eigenwertproblems $\Psi(\lambda)$ sind [46].

4.6.2. Newton–Verfahren

Das NEWTON–Verfahren für nichtlineare Eigenwertprobleme kann als Verallgemeinerung der Methode der inversen Iteration angesehen werden [75]. Voraussetzung ist, dass eine Ableitung des nichtlinearen Eigenwertproblems nach dem Eigenwert zugänglich ist, was jedoch im Fall der polynomialen Eigenwertprobleme möglich ist.

Die NEWTON–Iteration basiert auf Funktion \mathbf{F} , die von der Eigenwertnäherung \mathbf{u} und der Eigenwertschätzung τ abhängig ist. Die Ableitung von \mathbf{F} ist durch die JACOBI–Matrix gegeben, so dass \mathbf{F} und \mathbf{J} folgende Gestalt annehmen

$$\mathbf{F} \left(\begin{bmatrix} \mathbf{u} \\ \tau \end{bmatrix} \right) = \begin{bmatrix} \Psi(\tau)\mathbf{u} \\ \mathbf{w}^H\mathbf{u} - 1 \end{bmatrix}, \quad \mathbf{J} \left(\begin{bmatrix} \mathbf{u} \\ \tau \end{bmatrix} \right) = \begin{bmatrix} \Psi'(\tau) & \Psi'(\tau)\mathbf{u} \\ \mathbf{w}^H & 0 \end{bmatrix}. \quad (4.23)$$

Die Funktion \mathbf{F} besteht aus dem Matrixpolynom $\Psi(\lambda)$ und einem Normierungsvektor \mathbf{w} , der sicherstellt, dass $\mathbf{w}^H\mathbf{u} = 1$ während der Iteration gilt. $\Psi'(\lambda)$ bezeichnet die Ableitung des Matrixpolynoms $\Psi(\lambda)$ bezüglich λ . Beginnend mit den Startwerten (\mathbf{u}_0, τ_0) , die in der unmittelbaren Umgebung des gesuchten Eigenpaars liegen müssen, ergeben sich die fortfolgenden Schritte der NEWTON–Iteration

$$\mathbf{J} \left(\begin{bmatrix} \mathbf{u}_i \\ \tau_i \end{bmatrix} \right) \begin{bmatrix} \Delta\mathbf{u}_{i+1} \\ \Delta\tau_{i+1} \end{bmatrix} = -\mathbf{F} \left(\begin{bmatrix} \mathbf{u}_i \\ \tau_i \end{bmatrix} \right), \quad (4.24)$$

wobei der NEWTON–Schritt durch

$$\mathbf{u}_{i+1} = \mathbf{u}_i + \Delta\mathbf{u}_{i+1}, \quad \tau_{i+1} = \tau_i + \Delta\tau_{i+1} \quad (4.25)$$

gegeben ist. In jedem Iterationsschritt von (4.24) muss ein lineares Gleichungssystem exakt gelöst werden, wobei in diesem Fall keine Dimensionsreduktion durch eine Unterraumprojektion vorgesehen ist. Die NEWTON–Iteration wird abgebrochen, wenn die Norm des Residuums $\|\mathbf{r}\| = \|\Psi(\tau_{i+1})\mathbf{u}_{i+1}\|$ unter die Toleranzschwelle ϵ fällt.

4.6.3. Jacobi–Davidson–Verfahren für polynomiale Eigenwertprobleme

Die prinzipielle Struktur des JACOBI–DAVIDSON–Verfahrens aus Algorithmus 4.8 dient wiederum als Ausgangspunkt für das JACOBI–DAVIDSON–Verfahren für polynomiale Eigenwertprobleme (JDPEP). Allerdings lassen sich aufgrund der fehlenden SCHUR–Form die Erweiterungen aus den Algorithmen 4.9 und 4.10 nicht verwenden. In Algorithmus 4.14 ist das JACOBI–DAVIDSON–Verfahren für polynomiale Eigenwertprobleme angegeben [46]. Es dient dazu, zu einem vorgegebenen Zielwert τ und einem Startvektor \mathbf{v}_0 ein einziges Eigenpaar (λ, \mathbf{x}) des Matrixpolynoms Ψ zu berechnen.

Zur Diskussion der Abschnitte von Algorithmus 4.14:

Algorithmus 4.14 JACOBI-DAVIDSON-Verfahren zur Berechnung eines Eigenpaars eines polynomialen Eigenwertproblems $\Psi(\lambda)$ vom Grad ℓ nach [78]

Eingabe: Matrixpolynom $\Psi(\lambda) = \sum_{i=0}^{\ell} \lambda^i \mathbf{A}_i$, Startvektor \mathbf{v}_0 , Toleranz ϵ , Zielwert τ

Ausgabe: Eigenpaar (λ, \mathbf{x})

```

1: Wähle eine orthogonale  $n \times m$  Basis  $\mathbf{V}_m$ , die  $\mathbf{v}_0$  enthält, als Unterraum.
2: for  $i = 0, \dots, \ell$  do
3:    $\mathbf{W}_i = \mathbf{A}_i \mathbf{V}$ ,  $\mathbf{M}_i = \mathbf{V}^H \mathbf{W}_i$ 
4: end for
5: while 1 do                                     % Konvergenz noch nicht erreicht
6:   Bestimme RITZ-Paare  $(\theta, \mathbf{s})$  von  $(\sum_{i=0}^{\ell} \theta^i \mathbf{M}_i) \mathbf{s} = 0$            % Löse kleines allg. EWP
7:   Wähle RITZ-Paar  $(\theta, \mathbf{s})$  mit  $\theta$  nahe  $\tau$  und normiere  $\|\mathbf{s}\| = 1$ .
8:    $\mathbf{u} = \mathbf{V} \mathbf{s}$ ,  $\mathbf{p} = \Psi'(\theta) \mathbf{u}$ ,  $\mathbf{r} = \Psi(\theta) \mathbf{u}$ .
9:   if  $(\|\mathbf{r}\| < \epsilon)$  then
10:     $\lambda = \theta$ ,  $\mathbf{x} = \mathbf{u}$ , Stoppe
11:   end if
12:   Löse  $(\mathbf{I} - \frac{\mathbf{p} \mathbf{u}^H}{\mathbf{u}^H \mathbf{p}}) \Psi(\theta) (\mathbf{I} - \mathbf{u} \mathbf{u}^H) \mathbf{t} = -\mathbf{r}$  für  $\mathbf{t} \perp \mathbf{u}$  % ggf. iterativ und näherungsweise
13:   Orthonormiere  $\mathbf{t} \perp \mathbf{V}$ ,  $\mathbf{v} = \mathbf{t} / \|\mathbf{t}\|$ 
14:   for  $i = 0, \dots, \ell$  do
15:      $\mathbf{w}_i = \mathbf{A}_i \mathbf{v}$ 
16:      $\mathbf{M}_i = \begin{pmatrix} \mathbf{M}_i & \mathbf{V}^H \mathbf{w}_i \\ \mathbf{v}^H \mathbf{W}_i & \mathbf{v}^H \mathbf{w}_i \end{pmatrix}$ ,  $\mathbf{W}_i = [\mathbf{W}_i, \mathbf{w}_i]$ 
17:   end for
18:    $\mathbf{V} = [\mathbf{V}, \mathbf{v}]$ 
19: end while

```

- (2) Projektion der Koeffizientenmatrizen von Ψ in den Unterraum \mathbf{V}_m .
- (6) Die Lösung des niederdimensionalen allgemeinen Eigenwertproblems kann gut mit der Linearisierung aus Abschnitt 4.6.1 und einem Löser für kleine allgemeine Eigenwertprobleme erfolgen.
- (7) Bei der Selektion der RITZ–Paare ist sicherzustellen, dass das Einsetzen der RITZ–Paare in Ψ auch tatsächlich ein Residuum mit kleiner Norm liefert, Gegenbeispiele z.B. in [46].
- (8) Ψ' ist die Ableitung des Matrixpolynoms bezüglich θ .
- (9) Wenn die Approximation eines RITZ–Paares hinreichend gut bezüglich der Toleranzschwelle ϵ ausfällt, kann abgebrochen werden.
- (12) Die Korrekturgleichung des JACOBI–DAVIDSON–Verfahrens muss mindestens näherungsweise gelöst werden, um eine neue Suchrichtung \mathbf{t} für den Unterraum \mathbf{V}_m zu erhalten.
- (13) Die neue Suchrichtung wird dem Unterraum hinzugefügt und die projizierten Matrizen werden aktualisiert.

Da diese prinzipielle Vorgehensweise kaum auf die benötigten Lösungsansätze für spezifische polynomiale Eigenwertprobleme eingeht, wird im nachfolgenden Abschnitt auf die Modifikation von Algorithmus 4.14 für das PEP aus Gleichung (3.24) eingegangen.

4.6.4. Implementierungsdetails des Jacobi–Davidson–Verfahren für polynomiale Eigenwertprobleme

Algorithmus 4.14 gibt die wesentliche Vorgehensweise für das JACOBI–DAVIDSON–Verfahren für polynomiale Eigenwertprobleme wieder, ohne auf Spezifika einzugehen, die bei der Anwendung auf eine konkrete Problemstellung aus Implementierungssicht auftreten. Algorithmus 4.15 berücksichtigt diese Aspekte und ist darauf ausgerichtet, das polynomiale Eigenwertproblem der FIT nach Gleichung (3.24) für verlustbehaftete Elementarzellen zu lösen. Als Eingabeparameter werden dabei unter anderem die Koeffizientenmatrizen \mathbf{A}_0 bis \mathbf{A}_2 vorausgesetzt, die nach Satz 4.1.28 mit der Methode aus [48] skaliert worden sind. Zur Diskussion der Abschnitte von Algorithmus 4.15:

- (5) Projektion der Koeffizientenmatrizen von in den Unterraum \mathbf{V}_m .
- (7) Die Transformation des Eigenwerts nach Satz 4.1.29 dient dazu, die Kondition der Linearisierung $L(\lambda)$ zu verbessern [48].
- (8) Linearisierung des polynomialen Eigenwertproblems nach Satz 4.1.8.
- (9) Das Lösen des niederdimensionalen allgemeinen Eigenwertproblems erfolgt mit Standard–Verfahren.
- (10) Überprüfung, ob die mit der Linearisierung berechneten Lösungen, das polynomiale Eigenwertproblem erfüllen, Gegenbeispiele z.B. in [46].
- (16) Für die weitere Iteration wird der RITZ–Wert selektiert, der τ am nächsten liegt.

Algorithmus 4.15 JACOBI–DAVIDSON–Verfahren zur Berechnung eines einzigen Eigenpaars des quadratischen Eigenwertproblems nach Gleichung (3.24)

Eingabe: Skalierte Koeffizientenmatrizen $\mathbf{A}_0, \mathbf{A}_1, \mathbf{A}_2$, Startvektor \mathbf{v}_0 , Toleranz ϵ , Zielwert τ , maximale Iterationszahl it_{\max}

Ausgabe: Eigenpaar (λ, \mathbf{x})

```

1:  $m_{\min} \leftarrow 6, m_{\max} \leftarrow 17$            % Setze minimale und maximale Unterraumdimension.
2:  $it \leftarrow 1, n \leftarrow \dim(\mathbf{A}_0)$ 
3: Initialisiere eine orthonormale  $n \times m$  Basis  $\mathbf{V}_{m_{\min}}$ , die  $\mathbf{v}_0$  enthält, als Unterraum.
4: Initialisiere LU–Vorkonditionierer.
5:  $\mathbf{M}_0 \leftarrow \mathbf{V}_m^H \mathbf{A}_0 \mathbf{V}_m, \mathbf{M}_1 \leftarrow \mathbf{V}_m^H \mathbf{A}_1 \mathbf{V}_m, \mathbf{M}_2 \leftarrow \mathbf{V}_m^H \mathbf{A}_2 \mathbf{V}_m$ 
6: while  $it \leq it_{\max}$  do           % Konvergenz / maximale Iterationszahl noch nicht erreicht
7:    $\mathbf{M}_1 \leftarrow \alpha_{opt} \mathbf{M}_2$  und  $\mathbf{M}_2 \leftarrow \alpha_{opt}^2 \mathbf{M}_2$  mit  $\alpha_{opt} \leftarrow \sqrt{\|\mathbf{M}_0\|_2 / \|\mathbf{M}_2\|_2}$  (Satz 4.1.29).
8:   Bilde Linearisierung  $L(\lambda)$  nach (4.2) mit  $\mathbf{A} \leftarrow \begin{pmatrix} \mathbf{M}_1 & \mathbf{M}_0 \\ -\mathbf{I} & \mathbf{0} \end{pmatrix}$  und  $\mathbf{B} \leftarrow \begin{pmatrix} -\mathbf{M}_2 & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} \end{pmatrix}$ .
9:   Löse  $[\mathbf{S}, \Theta] \leftarrow \text{eig}(\mathbf{A}, \mathbf{B})$  mit Eigenvektoren  $\mathbf{S}$  und Eigenwerten  $\Theta$ .
10:  for  $i = 1, \dots, 2m$  do           % Überprüfung, ob Lösungen PEP erfüllen.
11:     $r_i \leftarrow \|(\Theta_i^2 \mathbf{M}_2 + \Theta_i \mathbf{M}_1 + \mathbf{M}_0) \mathbf{S}_{m+1, \dots, 2m; i}\|$ 
12:    Ggf. Aussortierung fehlerhafter Lösungen, Reduktion von  $\mathbf{S}$  und  $\Theta$ .
13:  end for
14:   $\Theta \leftarrow \alpha_{opt} \Theta$ 
15:   $I \leftarrow \text{sort}(|\theta - \tau|)$            % Selektiere RITZ–Wert.
16:   $\theta \leftarrow \Theta(I_1)$ 
17:   $\mathbf{s} \leftarrow \mathbf{S}_{m+1, \dots, 2m; I_1}, \mathbf{s} \leftarrow \mathbf{s} / \|\mathbf{s}\|$ 
18:   $\mathbf{u} \leftarrow \mathbf{V}_m \mathbf{s}, \mathbf{u} \leftarrow \mathbf{u} / \|\mathbf{u}\|$            % Bilde RITZ–Vektor.
19:   $\mathbf{A}_\Psi \leftarrow \theta \mathbf{A}_2 + \mathbf{A}_1 + \mathbf{A}_0 / \theta$ 
20:   $\mathbf{r} \leftarrow \mathbf{A}_\Psi \mathbf{u}$            % Bilde Residuum.
21:  if  $(\|\mathbf{r}\| < \epsilon)$  then           % Konvergenz erreicht?
22:     $\lambda \leftarrow \theta, \mathbf{x} \leftarrow \mathbf{u}, \text{Stoppe}$ 
23:  end if
24:  if  $it < 5$  then           % Auswertung an der Ableitung anfangs mit  $\tau$  statt  $\theta$ .
25:     $\mathbf{p} \leftarrow (\mathbf{A}_2 - \mathbf{A}_0 / \tau^2) \mathbf{u}$ 
26:  else
27:     $\mathbf{p} \leftarrow (\mathbf{A}_2 - \mathbf{A}_0 / \theta^2) \mathbf{u}$ 
28:  end if
29:  Löse Gleichung (4.18) als  $\mathbf{A}_\Psi[\mathbf{v}_1, \mathbf{v}_2] = [\mathbf{r}, \mathbf{p}]$  mit BiCGStab und LU–Vorkonditionierung
  in max. 10 Schritten.
30:   $\mathbf{t} \leftarrow -\mathbf{v}_1 + \frac{\mathbf{u}^H \mathbf{v}_1}{\mathbf{u}^H \mathbf{v}_2} \mathbf{v}_2$ 
31:  if  $m == m_{\max}$  then
32:    Reduziere Unterraumdimension von  $\mathbf{V}_m$  auf  $m_{\min}$ ,  $m \leftarrow m_{\min}$ .
33:    Verwende dazu die Eigenvektoren aus  $\mathbf{S}_{m+1, \dots, 2m; I_{2,3, \dots, m_{\min}-1}}$ .
34:  end if
35:  Orthonormiere  $\mathbf{t} \perp \mathbf{V}_m, \mathbf{v} \leftarrow \mathbf{t} / \|\mathbf{t}\|$ .
36:  Erweitere Unterraum  $\mathbf{V}_m \leftarrow [\mathbf{V}_m \mathbf{v}]$ .
37:   $\mathbf{M}_0 \leftarrow \mathbf{V}_m^H \mathbf{A}_0 \mathbf{V}_m, \mathbf{M}_1 \leftarrow \mathbf{V}_m^H \mathbf{A}_1 \mathbf{V}_m, \mathbf{M}_2 \leftarrow \mathbf{V}_m^H \mathbf{A}_2 \mathbf{V}_m$ 
38:   $m \leftarrow m + 1, it \leftarrow it + 1$ 
39: end while

```

- (21) Wenn die Approximation eines RITZ–Paars hinreichend gut bezüglich der Toleranzschwelle ϵ ausfällt, kann abgebrochen werden.
- (24) Zu Beginn von Algorithmus 4.15 ist der vorgegebene Zielwert τ näher am gesuchten Eigenwert λ als die berechneten RITZ–Werte θ , so dass die Ableitung des Matrixpolynoms an τ ausgewertet wird [64]. Mit dieser Maßnahme lassen sich in der Anfangsphase der Iteration bessere Korrekturvektoren \mathbf{t} erhalten, indem die Norm des Residuums noch groß ist.
- (29) Die Korrekturgleichung des JACOBI–DAVIDSON–Verfahrens wird in Form von Gleichung (4.18) mit dem BiCGStab–Verfahren näherungsweise gelöst. Dabei kommt eine LU–Vorkonditionierung zum Einsatz. Aus den Teillösungen \mathbf{v}_1 und \mathbf{v}_2 kann die Erweiterung \mathbf{t} für den Unterraum \mathbf{V}_m berechnet werden.
- (31) Sofern die Unterraumdimension den maximal vorgegebenen Wert überschreitet, wird der Unterraum auf die Dimension m_{\min} reduziert. Dazu werden die m_{\min} verbliebenen Eigenvektoren in \mathbf{S} verwendet, deren RITZ–Wert dem Zielwert τ am nächsten sind.
- (35) Die neue Suchrichtung \mathbf{t} wird orthonormalisiert und dem Unterraum hinzugefügt. Die projizierten Matrizen werden aktualisiert.

Ein JACOBI–DAVIDSON–Verfahren für polynomiale Eigenwertprobleme unter der Berücksichtigung harmonischer und verfeinerter RITZ–Werte ist in [64] angegeben.

Die Erweiterung des JACOBI–DAVIDSON–Verfahrens auf stärker als polynomial nichtlineare Eigenwertprobleme ist generell möglich, wenngleich Lösungsstrategien für das niederdimensional projizierte Eigenwertproblem erforderlich werden. Die Klasse der lokal konvergenten Methoden ist dazu prinzipiell verwendbar – sie wird im nächsten Abschnitt vorgestellt.

4.7. Löser für nichtlineare Eigenwertprobleme

Die Nichtlinearität beschränkt sich in den FIT–Formulierungen, die in Kapitel 3 vorgestellt sind, auf polynomiale Abhängigkeiten vom Eigenwert. Die FLAME–Formulierung aus Abschnitt 3.5 führt auf ein nichtlineares Eigenwertproblem, dessen Abhängigkeit nicht polynomial vom Eigenwert ist. Wenn die in FLAME angesetzten Basisfunktionen Exponentialfunktionen sind und ebene Wellen darstellen, kann den zeilenweisen Einträgen der Systemmatrix $\mathbf{A}_{\text{FLAME}}$ noch die Abhängigkeit vom Eigenwert angesehen werden, wie in Abschnitt 3.5.1 mit den Ergebnissen aus [36] gezeigt ist. Werden jedoch lokal Zylinderwellen in Form von BESSEL- und HANKEL–Funktionen angesetzt, lässt sich nach dem Kern der Knotenmatrix nach Gleichung (2.122) nicht mehr in jedem Fall symbolisch auflösen, so dass die Abhängigkeit der Systemmatrix vom Eigenwert nicht mehr greifbar ist.

Da Eigenwerte von den Koeffizienten des charakteristischen Polynoms nach Definition 4.1.11 abhängen, hängen sie – zumindest in einer lokalen Umgebung – auch kontinuierlich von den Einträgen der Koeffizientenmatrizen von Ψ ab. Diese kontinuierliche Abhängigkeit ermöglicht es, jedes Eigenwertproblem als Nullstellenproblem zu betrachten, dessen Lösung gerade durch das Eigenpaar (λ, \mathbf{x}) gegeben ist. Abbildung 4.3 veranschaulicht diesen Zusammenhang. Diskutiert wird im Folgenden die Fixpunkt–Iteration für das Eigenwertproblem des an τ linearisierten Matrixpolynoms sowie das Sekanten–Verfahren. Dabei handelt es sich jeweils um

4.7.2. Sekanten–Verfahren

Im nicht-polynomialen nichtlinearen Eigenwertproblem $\Xi(\lambda)\mathbf{x} = \mathbf{0}$ aus Gleichung (3.47) taucht der Eigenwert λ nicht mehr als Vorfaktor auf, jedoch ist der gesuchte Eigenvektor offensichtlich identisch mit dem Kern⁴¹ der bei λ ausgewerteten Systemmatrix $\mathbf{x} \in \ker \Xi(\lambda)$. Zur Lösung des nichtlinearen Eigenwertproblems wird der betragskleinste Eigenwert λ_τ der bei τ ausgewerteten Systemmatrix $\Xi(\tau)$ betrachtet. Abbildung 4.3 zeigt den Verlauf von Real- und Imaginärteil von λ_τ über τ eines nichtlinearen Eigenwertproblems der Form (3.47). Für $\lambda_\tau = 0$ ist τ Eigenwert und es gilt $\Xi(\tau)\mathbf{x} = \mathbf{0}$ und $\lambda = \tau$.

Lässt sich die Ableitung eines allgemein nichtlinearen Eigenwertproblems $\Xi(\lambda)\mathbf{x} = \mathbf{0}$ nicht bilden, so lässt sich die für das NEWTON–Verfahren nach Abschnitt 4.6.2 notwendige Ableitung durch einen Differenzenquotienten approximieren, was zwei Funktionsauswertungen erfordert und auf das Sekantenverfahren in Algorithmus 4.17 führt. Der Ablauf ist dabei vollkommen analog zum Sekantenverfahren für skalare analytische Funktionen, wobei hier die Funktionswerte durch den jeweils betragskleinsten Eigenwert von Ξ_τ in Abhängigkeit von τ berechnet werden. Als Eigenwertlöser kommen prinzipiell alle Methoden für Standardeigenwertprobleme in Frage.

Algorithmus 4.17 Sekanten–Verfahren zur Berechnung eines Eigenpaars des nichtlinearen Eigenwertproblems $\Xi(\lambda)\mathbf{x} = \mathbf{0}$

Eingabe: Nichtlineares Eigenwertproblem $\Xi(\tau)$, Startbereich $[\tau_0, \tau_1]$, Maximalanzahl von Iterationen n_{\max} , Toleranz ϵ

Ausgabe: Eigenpaar (λ, \mathbf{x})

- 1: Generiere Systemmatrix $\Xi_{\tau_1} \leftarrow \Xi(\tau_1)$
 - 2: $[\mathbf{x}, \lambda_{\tau_1}] \leftarrow$ Berechne betragskleinsten Eigenwert und -vektor von Ξ_{τ_1}
 - 3: $\mathbf{v}_\tau \leftarrow [\tau]$, $\mathbf{f}_\tau \leftarrow [\lambda_{\tau_1}]$;
 - 4: Generiere Systemmatrix $\Xi_{\tau_0} \leftarrow \Xi(\tau_0)$
 - 5: $[\mathbf{x}, \lambda_{\tau_0}] \leftarrow$ Berechne betragskleinsten Eigenwert und -vektor von Ξ_{τ_0}
 - 6: $\mathbf{v}_\tau \leftarrow [\mathbf{v}_\tau; \tau_0]$
 - 7: $r \leftarrow \epsilon + 1$, $n_{it} \leftarrow 0$, $\mathbf{r} \leftarrow []$
 - 8: **while** ($n_{it} < n_{\max}$ **and** $r > \epsilon$) **do**
 - 9: $n_{it} \leftarrow n_{it} + 1$
 - 10: $\tau \leftarrow \tau_0 - \lambda_{\tau_0} \frac{(\tau_0 - \tau_1)}{(\lambda_{\tau_0} - \lambda_{\tau_1})}$
 - 11: $\mathbf{v}_\tau \leftarrow [\mathbf{v}_\tau; \tau]$
 - 12: Generiere Systemmatrix $\Xi_\tau \leftarrow \Xi(\tau)$
 - 13: $[\mathbf{x}, \lambda_\tau] \leftarrow$ Berechne betragskleinsten Eigenwert und -vektor von Ξ_τ
 - 14: $r \leftarrow |\lambda_{\tau_0} - \lambda_\tau| / |\lambda_\tau|$
 - 15: $\tau_1 \leftarrow \tau_0$, $\lambda_{\tau_1} \leftarrow \lambda_{\tau_0}$
 - 16: $\tau_0 \leftarrow \tau$, $\lambda_{\tau_0} \leftarrow \lambda_\tau$
 - 17: **end while**
 - 18: $\lambda \leftarrow \mathbf{v}_\tau(\text{end})$
-

⁴¹Siehe Definition 4.1.9.

5. Ergebnisse und Anwendungen

Die Diskretisierung und Matrixerzeugung für alle Modelle der nachfolgenden Beispiele und Anwendungen werden mit CST MICROWAVE STUDIO [123] durchgeführt. Es kommt ebenfalls zur Berechnung der benötigten Streumatrizen zum Einsatz.

Da die im Folgenden betrachteten Strukturen und Methoden von unterschiedlichen Ausgangspunkten zu Eigenwertproblemen führen, ist zunächst eine Nomenklatur der resultierenden Eigenlösungen erforderlich:

1. BLOCH-Mode: Dreidimensionale Eigenlösung der MAXWELL'schen Gleichungen, welche die Periodizität des zugrunde liegenden Strukturgitters erfüllt und durch einen diskreten Eigenvektor repräsentiert wird.
2. Eigenmode: Dreidimensionale, durch einen diskreten Eigenvektor repräsentierte, Eigenlösung der MAXWELL'schen Gleichungen, bei der die zugrunde liegenden Strukturen keine Periodizität aufweisen.
3. Eigenzustände: Bezeichnung für die Eigenlösung der Transfer-Streumatrix \mathcal{T} aus Abschnitt 2.1.5.4.
4. Wellenleitermode: Dreikomponentige, Eigenlösung der MAXWELL'schen Gleichungen in einer orthogonal zur Ausbreitungsrichtung des längshomogenen Wellenleiters stehenden zweidimensionalen Querschnittsfläche nach Abschnitt 2.1.5.1.

5.1. Dispersionseigenschaften von Elementarzellen im Zentimeterbereich

Elektromagnetische *Metamaterialien* waren seit etwa dem Jahrtausendwechsel das Hauptaugenmerk vieler Forschungsvorhaben, und beträchtlicher Aufwand wurde für die Analyse sogenannter links-händiger *Metamaterialien* (LHM) erbracht, bei denen POYNTING-Vektor und Wellenvektor antiparallel sind [99, 101, 12]. Typischerweise ist der Entwurf und die Anwendung dieser synthetischen Materialien vom Konzept der effektiven Materialparameter geprägt, mit denen das frequenzabhängige makroskopische Verhalten für elektromagnetische Wellen modelliert werden. In einem begrenzten Frequenzband lassen sich gleichzeitig die effektiv negative Permeabilität und die effektiv negative Permittivität beobachten, die letztlich für das links-händige Wellenverhalten ursächlich sind [109].

Als periodische Strukturgitter mit kleinen geometrischen Details, die nach Abbildung 2.4 kleiner als die Wellenlänge sind, ergeben LHM ein Multiskalenproblem. Die Analyse durch dreidimensionale Simulation des elektromagnetischen Feldes wird für aus mehreren Elementarzellen bestehenden Strukturgitter schnell sehr aufwendig. Daher wird für die Charakterisierung eine einzelne Elementarzelle mit adäquaten Randbedingungen verwendet. Zur Simulation einzelner Elementarzellen sind Volumengittermethoden, wie die Methode der finiten Integration aus

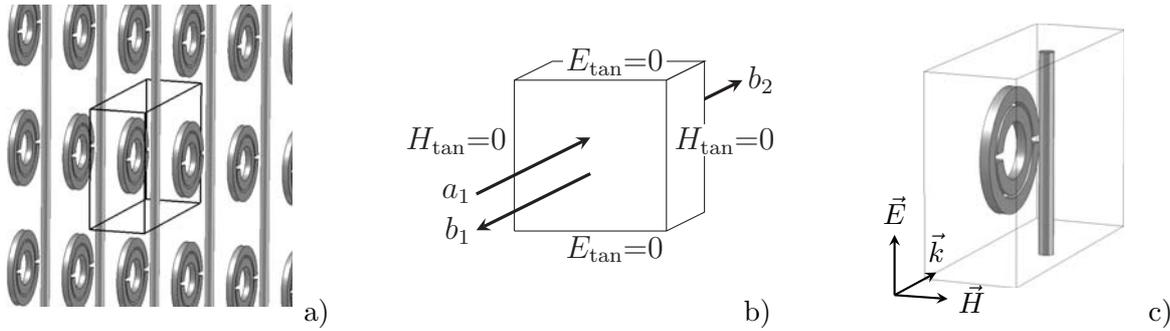


Abbildung 5.1.: Modellierung eines periodischen Strukturgitters a) durch eine einzige Elementarzelle mit Anregung und Randbedingung b). Polarisation der ebenen Welle bezüglich der Elementarzelle zur Anregung der linkshändigen Wellenausbreitung c).

Abschnitt 2.2 gut geeignet und liefern Lösungen für die elektrischen und magnetischen Felder im Rechengitter der Diskretisierung einer Elementarzelle (*mikroskopische Simulationen*). Das Konzept der Beschreibung von Elementarzellen durch effektive Materialparameter, die das äquivalente elektromagnetische Verhalten darstellen, stellt nach wie vor ein nur teilweise gelöstes Problem dar. Darüber hinaus ist die zuverlässige Extraktion effektiver Materialparameter aus den Lösungen der *mikroskopischen Simulationen* schwierig [22]. Sofern verlässliche effektive Materialparameter einer Elementarzelle bekannt sind, können diese in nachfolgenden Simulationen als äquivalente Ersatzmaterialien verwendet werden, was als *makroskopische Simulation* bezeichnet wird und den Rechenaufwand erheblich verringert.

Im Folgenden liegt der Schwerpunkt auf Elementarzellen im Mikrowellenbereich, da bei diesen Frequenzen mit Kupfer und verlustarmen Dielektrika Materialien existieren, die näherungsweise als verlustfrei modelliert werden können.

Aus einem unendlich periodischen Strukturgitter aus Abbildung 5.1a wird eine einzige Elementarzelle betrachtet. Das Modell der Elementarzelle wird in einem Parallelplattenleiter platziert, der transversal mit dem Abschluss paarweiser konservativer Randbedingungen entsprechend Abbildung 5.1b modelliert werden kann. Die Periodizität in longitudinaler Richtung wird durch den jeweiligen Ansatz implementiert. Der Leistungsfluss der senkrecht einfallenden Wellen erfolgt longitudinal durch die Elementarzelle wie in Abbildung 5.1c veranschaulicht ist. Die Berandung durch einen Parallelplattenleiter und die Verwendung seiner Wellenleitermoden stellt an dieser Stelle eine Approximation und Vereinfachung dar. Die Anordnung kann in zwei unterschiedlichen Sichtweisen aufgefasst werden, die auf verschiedene Simulationsansätze führen.

Im ersten Ansatz kann die Elementarzelle als Zweitor mit einfallenden und auslaufenden Wellen betrachtet werden, deren Amplitudenverhältnisse durch die Streumatrix (2.62) charakterisiert werden können. Der fundamentale Wellenleitermode ist dabei der TEM-Mode, der einer auf die Elementarzelle fallende ebenen Welle entspricht. Der sogenannte Streumatrix-Ansatz (SMA) wird häufig zur Charakterisierung unbekannter Materialparameter angewendet und ist im Labor bereits seit den 1970ern eine etablierte Methode, um experimentell Materialparameter einer Probe zu bestimmen [107, 108]. Da dieser Ansatz bei der Charakterisierung effektiver Materialien manchmal fehlschlägt, werden in [96] Verbesserungen für links-händige Materialien eingeführt. Wichtige Stichwörter sind in diesem Zusammenhang die Bestimmung der effektiven Ränder der Elementarzelle, die erzwungene Kontinuität des dispersiven Brechungs-

index und die Eliminierung von Rauschen, das durch Messung oder Simulation bedingt ist, sowie der Einfluss auf die effektive Impedanz des LHM. Die Ergebnisse des SMA können ungenau werden oder der Ansatz kann fehlschlagen, wenn Wellenleitermoden höherer Ordnung angeregt werden¹. Obwohl alle Wellenleitermoden höherer Ordnung im LHM-Frequenzband weit unterhalb ihrer Cutoff-Frequenz liegen, können sie dennoch zum Energieaustausch zwischen einzelnen Elementarzellen beitragen, wenn die Strukturgitterkonstante klein genug ist.

Der zweite Ansatz verwendet eine periodische Randbedingung in Ausbreitungsrichtung nach Abschnitt 2.2.4.2 zur Modellierung der Periodizität. Dieser Ansatz wird z.B. in [115] verwendet und wird im Folgenden als periodisch berandetes Eigenwertproblem (PBE) bezeichnet. Ein Dispersionsdiagramm, wie in Abbildung 2.2 eingeführt, wird dabei aus einer Folge von Simulationen mit variierendem Phasenvorschub φ gewonnen, bei denen jeweils ein Eigenwertproblem gelöst werden muss. Im PBE-Ansatz bereiten Wellenleitermoden höherer Ordnung keine Probleme.

Um die beiden Ansätze SMA und PBE vergleichen zu können, werden zunächst die Felder der PBE-Lösung als Lösung einer Zweitor-Struktur reinterpretiert. Analog dazu lassen sich aus den SMA-Ergebnissen in einem a posteriori Schritt der periodische Fall nach Abschnitt 3.1.1 und das zugehörige Dispersionsdiagramm konstruieren. Das hier gewählte Testbeispiel ist eine modifizierte LHM-Elementarzelle aus [119] und ist in Abbildung 5.1c dargestellt. Die detaillierten Abmessungen beinhaltet Anhang A.4.

5.1.1. Simulationsansätze

Ausgangspunkt ist jeweils die dreidimensionale Simulation der elektromagnetischen Felder in der LHM-Elementarzelle. Die beiden vorgestellten Ansätze werden im Folgenden verglichen, um die Ausbreitungseigenschaften der Wellenleitermoden in einem *unendlichen ausgedehnten Strukturgitter* aus der Lösung von einer einzigen Elementarzelle zu berechnen.

5.1.1.1. Streumatrix-Ansatz (SMA)

Die Vorgehensweise erfolgt nach dem Streumatrix-Ansatz aus Abschnitt 3.1.1. Die BLOCH-Bedingung (3.1) wird verwendet, um a posteriori den periodischen Fall zu konstruieren. Dieser Ansatz kann, wie in Abschnitt 3.1.2 gezeigt worden ist, erweitert werden, um neben dem fundamentalen auch weitere Wellenleitermoden höherer Ordnung berücksichtigen zu können. Die Berechnung des Dispersionsdiagramms erfolgt über die Transfer-Streumatrix \mathcal{T} nach (3.4) und die Bestimmung ihrer Eigenzustände (\mathbf{v}_m, γ_m) nach (3.5) in jedem einzelnen Frequenzpunkt. Die Dimension des Eigenwertproblems entspricht hierbei der doppelten Anzahl der betrachteten Wellenleitermoden, da Wellenamplituden für Vorwärts- und Rückwärtsrichtung auftreten. Im hier betrachteten verlustfreien Fall repräsentieren alle Eigenzustände γ_m mit $|\gamma_m| = 1$ einen ausbreitungsfähigen Eigenzustand, während die Fälle mit $|\gamma_m| \neq 1$ einen evaneszenten Eigenzustand bei der jeweiligen Frequenz darstellen. Die Eigenzustände treten paarweise auf und lassen sich mit (3.7) bezüglich der Richtung ihres Leistungsflusses sortieren. Die Vorwärtsrichtung ist dabei von Tor 1 zu Tor 2 definiert. In der betrachteten LHM-Elementarzelle Abbildung 5.1c ist nur der fundamentale Wellenleitermode im linkshändigen

¹Die Untersuchungen zur Berücksichtigung Moden höherer Ordnung im Streumatrixansatz ist in Kooperation mit Dr.-Ing. Grzegorz Lubkowski und Prof. Dr.-Ing. Thomas Weiland vom Institut Theorie elektromagnetischer Felder an der TU Darmstadt entstanden [142].

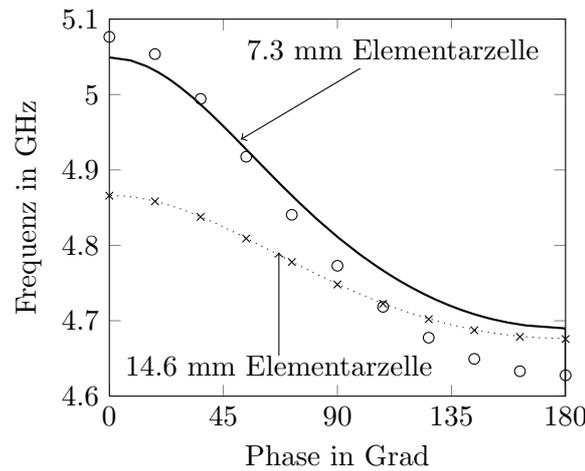


Abbildung 5.2.: Dispersionsdiagramm für eine verkürzte Elementarzelle nach Abbildung A.2 (durchgezogene Linie, \circ) und für die größere Elementarzelle aus Abbildung A.1 (gepunktete Linie, \times). Ergebnisse aus dem periodisch berandeten Eigenwertproblem (PBE) (\times, \circ) und dem Streumatrix-Ansatz (SMA) des fundamentalen Wellenleitermodes (Linien).

Übertragungsband ausbreitungsfähig, was sich durch eine fallende Kurve im Dispersionsdiagramm manifestiert. Die Zuordnung von Eigenzustand γ_m zur gesuchten Phase ist durch $\varphi(\omega) = \angle(\gamma_m)$ gegeben, wie in Abschnitt 3.1.2 gezeigt ist.

5.1.1.2. Periodisch berandetes Eigenwertproblem (PBE)

Im periodisch berandeten Eigenwertproblem werden die transversalen Feldkomponenten zweier gegenüberliegender Randflächen der Elementarzelle in eine feste Phasenbeziehung φ gesetzt. In Kombination mit der Zellgröße in Ausbreitungsrichtung ergibt sich die Wellenzahl der makroskopischen Welle im LHM-Strukturgitter, dem BLOCH-Mode. Wie in Abschnitt 3.2.1 gezeigt worden ist, geben die Eigenwerte der resultierenden komplex-hermiteschen Systemmatrix nach (3.8) oder (3.9) die Frequenzen an, bei denen der BLOCH-Mode als makroskopische Welle ausbreitungsfähig ist. Ein Dispersionsdiagramm nach Abbildung 5.2 wird durch sukzessive Berechnung mehrerer Eigenwertlösungen bei unterschiedlichen Phasenverschiebungen erzeugt [115]. Da die vollständig dreidimensionale Feldlösung einen eingeschwungenen Zustand bei einer festen Frequenz repräsentiert, ist die Feldverteilung an den gegenüberliegenden periodischen Rändern im Allgemeinen durch Superposition vieler Wellenleitermoden darstellbar. Theoretisch gibt es eine unendliche Anzahl von Wellenleitermoden, während im diskretisierten Fall die maximale Anzahl möglicher Wellenleitermoden von der Diskretisierung abhängt. Die Entwicklung der BLOCH-Moden an den periodischen Rändern in Wellenleitermoden gibt Aufschluss darüber, welche und wie viele Wellenleitermoden für genaue Ergebnisse des Streumatrixansatzes SMA notwendig sind.

5.1.2. Numerische Ergebnisse

Die numerischen Simulationen mit FIT werden in CST MICROWAVE STUDIO im Frequenzbereich durchgeführt. Zur Bestimmung der Eigenzustände der \mathcal{T} -Matrix eignet sich die QR-

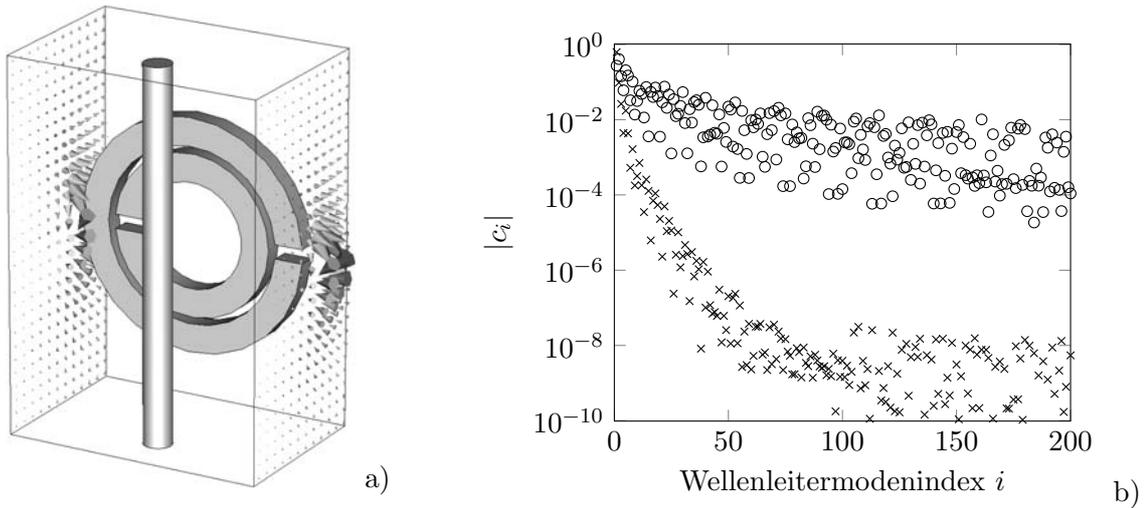


Abbildung 5.3.: a) Elektrische Feldstärke des BLOCH-Modus an den periodischen Rändern, der aus der PBE-Simulation der verkürzten Elementarzelle nach Abbildung A.2 bei einer Phasenverschiebung von 0° erhalten wird. b) Entwicklungskoeffizient c_i für die verkürzte Elementarzelle (\circ) und der originalen Elementarzelle nach Abbildung A.1 (\times). Bei der verkürzten Elementarzelle (\circ) sind die Wellenleitermoden höherer Ordnung erst wenig abgeklungen.

Zerlegung aus Abschnitt 4.2.1. Alle numerischen Experimente werden bei eher grober Diskretisierung durchgeführt und die verglichenen Ergebnisse von SMA und PBE-Ansatz beziehen sich auf das jeweils gleiche numerische Modell, das aus einer Diskretisierung mit FIT resultiert. Der Typ der betrachteten Elementarzelle ist der Literatur entnommen [119]. Es werden zwei verschiedene Strukturgitterkonstanten a_{per} in Ausbreitungsrichtung betrachtet: 7.3 mm und 14.6 mm. Alle weiteren Abmessungen sind Anhang A.4 zu entnehmen.

5.1.2.1. Wellenleitermoden höherer Ordnung

Bei der Analyse der Eigenschaften der spezifischen LHM-Elementarzelle aus Abbildung 5.1c mit dem SMA ist im betrachteten Frequenzbereich von 4.6 bis 5.1 GHz nur der fundamentale TEM-Wellenleitermode ausbreitungsfähig². Der nächst höhere Wellenleitermode ist bei dieser Frequenz evaneszent, da er eine Cutoff-Frequenz von etwa 28 GHz besitzt. Die Evaneszenz der Wellenleitermoden höherer Ordnung verursacht zusammen mit dem nicht vollständig symmetrischen Aufbau der Elementarzelle, dass die Streumatrix unsymmetrisch ist, obwohl die Substruktur der LHM-Elementarzelle als verlustfrei modelliert ist. Das Dispersionsdiagramm in Abbildung 5.2 zeigt, dass die Dispersionskurven aus SMA und PBE-Ansatz für die Elementarzelle mit 14.6 mm Länge in Ausbreitungsrichtung aus Abbildung A.1 gut übereinstimmen. Ebenfalls gut erkennbar ist in den PBE-Ergebnissen, dass sich das Frequenzband der linkshändigen Wellenausbreitung nach Abbildung 5.2 verbreitert, wenn die Länge in Ausbreitungsrichtung auf 7.3 mm reduziert wird (Geometrie nach Abbildung A.2). Die gute Übereinstimmung zwischen PBE-Ansatz und SMA geht dabei verloren, wenn nur der fundamentale Wellenleitermode im SMA berücksichtigt wird. Dies lässt darauf schließen, dass Wellenleitermoden höherer Ordnung berücksichtigt werden sollten. Um einen Einblick in das modale Spektrum zu erhalten, wird für beide Längen der Elementarzelle exemplarisch

²Abmessungen der Apertur in Anhang A.4

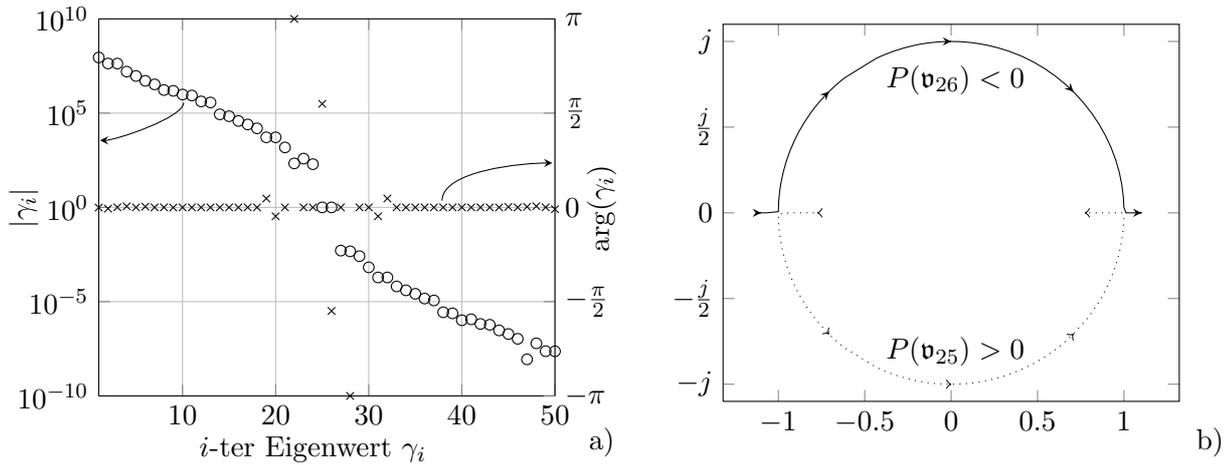


Abbildung 5.4.: a) Die Eigenzustände von \mathcal{T} treten stets paarweise auf, hier ist die Verteilung bei 4.75 GHz gezeigt. Die Paare γ , γ^{-1} repräsentieren Ausbreitung in Vorwärts- und Rückwärtsrichtung, die linke Skala und (○) zeigen die Beträge und die rechte Skala und (×) zeigen die jeweiligen Phasen. Die Ausbreitungsrichtung kann der Richtung des Leistungsflusses nach Gleichung (3.7) entnommen werden. b) Darstellung der ausbreitungsfähigen Eigenzustände als Ortskurve: Im linkshändigen Übertragungsband von 4.6-5.1 GHz liegen die Eigenwerte auf dem Einheitskreis. Die untere Halbebene ($P(\mathbf{v}_{25}) > 0$) bezieht sich auf die Vorwärtsrichtung, die obere Halbebene ($P(\mathbf{v}_{26}) < 0$) auf die Rückwärtsrichtung.

eine PBE-Lösung mit einer Phasenverschiebung von 0° über der Ausdehnung der Elementarzelle in Ausbreitungsrichtung berechnet (vgl. Abbildung 5.3a). Im nächsten Schritt wird die PBE-Lösung an den periodischen Randflächen in Wellenleitermoden entwickelt, wozu das Skalarprodukt (2.100) verwendet wird.

Das Ergebnis der ersten 200 Entwicklungskoeffizienten c_i ist in Abbildung 5.3b gezeigt. Erwartungsgemäß gibt es für die kleinere Elementarzelle nach nach Abbildung A.2 eine ganze Reihe von Wellenleitermoden höherer Ordnung, deren betragsmäßige Amplitude in der Größenordnung von 10^{-2} liegt, was die große Abweichung zwischen SMA und PBE-Ansatz in Abbildung 5.2 erklärt. Für die größere Elementarzelle mit 14.6 mm Ausdehnung in Ausbreitungsrichtung existieren lediglich drei Wellenleitermoden höherer Ordnung deren Amplituden in dieser Größenordnung liegen, was die bessere Übereinstimmung der Dispersionskurven von SMA und PBE-Ansatz in Abbildung 5.2 erklärt.

5.1.2.2. Eigenzustände der T-Matrix

Die LHM-Elementarzelle mit reduzierter Ausdehnung von 7.3 mm in Ausbreitungsrichtung nach nach Abbildung A.2 wird mit dem SMA analysiert, wobei die multimodale Streumatrix für insgesamt 25 Wellenleitermoden bestimmt wird. Dies führt auf eine 50×50 \mathcal{T} -Matrix mit insgesamt 50 Eigenzuständen. Die Eigenzustände der \mathcal{T} -Matrix bei einer Frequenz von 4.75 GHz sind in Abbildung 5.4a dargestellt. Dabei treten die Eigenwerte in Paaren zu γ , γ^{-1} auf, was sich durch die Symmetrie der Betragskurve (○) auszeichnet. Für die ersten und letzten Eigenwerte stimmt diese Paarbildung aufgrund der endlichen numerischen Genauigkeit nur näherungsweise, da der Faktor 10^{16} zwischen den ersten und letzten Eigenwerten liegt.

Im LHM-Frequenzband gibt es je einen ausbreitungsfähigen Eigenzustand der \mathcal{T} -Matrix pro Richtung des Leistungsflusses. So gilt für die Eigenzustände mit der Ordnungsnummer 25 und 26 $|\gamma_{25}| = |\gamma_{26}| = 1$. Die Richtung des Leistungsflusses wird mit Gleichung (3.7) bestimmt: γ_{25} propagiert in Vorwärtsrichtung und γ_{26} in Rückwärtsrichtung. Der Leistungsfluss der übrigen Eigenzustände ist Null, so dass diese Eigenzustände zu evaneszenten BLOCH-Moden gehören. Dies kann ebenfalls in Abbildung 5.4a der Phase (\times) verschiedener γ_i entnommen werden. Nur die ausbreitungsfähigen Eigenzustände γ_{25} und γ_{26} haben eine Phase, die kein Vielfaches von π ist. Die Eigenwerte γ_{19} , γ_{20} und γ_{31} , γ_{32} gehören zu entarteten Eigenzuständen, deren Phase auf Null oder ein Vielfaches von π gedreht werden kann. Abbildung 5.4b zeigt die Ortskurve von γ_{25} und γ_{26} in der komplexen Ebene als eine Funktion der Frequenz. Das Übertragungsband des in Vorwärtsrichtung ausbreitungsfähigen Eigenzustands γ_{25} befindet sich in der unteren Halbebene. Mit ansteigender Frequenz verläuft die Phase von $\gamma_{25}(\omega)$ gegen den Uhrzeigersinn von π nach 2π , wie es für eine links-händiges Übertragungsband erwartet wird. Der entsprechende Eigenwert für Ausbreitung in Rückwärtsrichtung befindet sich in der oberen Halbebene und variiert seine Phase im Uhrzeigersinn von π nach 0. Bei Frequenzen außerhalb des links-händigen Übertragungsbandes gilt für alle Eigenzustände $|\gamma_i| \neq 1$, was mit einer starken Dämpfung einhergeht, so dass kein Leistungsfluss in ausbreitungsfähigen Eigenzuständen stattfindet.

5.1.2.3. Dispersionsdiagramm

Ein Vergleich der Dispersionskurven aus den verschiedenen Ansätzen wird in Abbildung 5.5a gezeigt. Die Ergebnisse der PBE-Simulationen für 101 verschiedene Phasenvorschübe stellen dabei die Referenzlösung dar. Wie bereits in Abbildung 5.2 gezeigt ist, gibt es eine signifikante Differenz der Dispersionskurven von PBE-Simulation und Streumatrix-Ansatz, wenn in Letzterem nur der fundamentale Wellenleitermode enthalten ist. Die Einbeziehung der nächsten 24 Wellenleitermoden höherer Ordnung verbessert die Genauigkeit der SMA-Ergebnisse. Abbildung 5.5b zeigt das Residuum, das als Norm von $|f_{\text{SMA}} - f_{\text{PBE}}|/f_{\text{PBE}}$ über die 101 verschiedenen Werte von φ berechnet wird.

5.1.3. Fazit

In diesem Abschnitt wird das Auftreten von Wellenleitermoden höherer Ordnung in LHM-Elementarzellen analysiert, deren metallische Bestandteile nah an der Hüllfläche der Elementarzelle lokalisiert sind, was das Auftreten von Wellenleitermoden höherer Ordnung begünstigt. Die Vernachlässigung von Wellenleitermoden höherer Ordnung kann sich nachteilig auf die Genauigkeit von Dispersionskurven auswirken, die mit dem Streumatrix-Ansatz berechnet werden. Die BLOCH-Moden als Referenzlösung werden aus einem dreidimensionalen periodisch berandeten Eigenwertproblem gewonnen, das für verschiedene Phasenverschiebungen gelöst wird. Die Entwicklung dieser Felder in Wellenleitermoden dient zu einer Quantifizierung des Einflusses von Wellenleitermoden höherer Ordnung. Für eine Erweiterung des Streumatrix-Ansatzes auf Wellenleitermoden höherer Ordnung wird der Weg über die Transfer-Streumatrix aufgezeigt. Somit ist eine Charakterisierung von Strukturgrittern aus kurzen Elementarzellen anhand des Dispersionsdiagramms immer möglich. Das Konzept der effektiven Materialparameter versucht das elektromagnetische Verhalten als frequenzabhängige (ggf. auch richtungsabhängige) Werte der Permittivität und Permeabilität darzustellen.

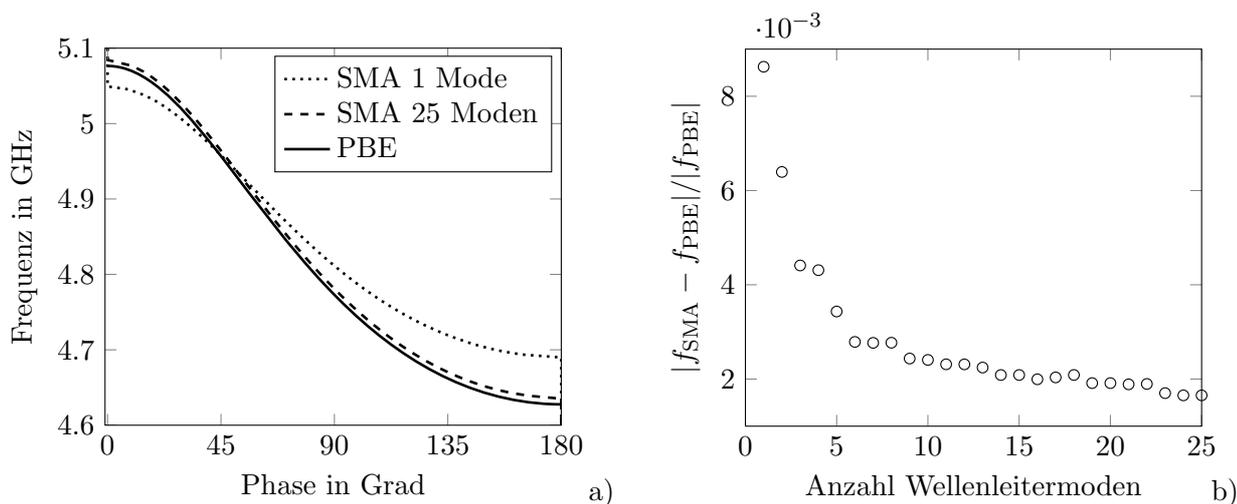


Abbildung 5.5.: a) Dispersionskurven für 0 (gepunktete Linie) und 24 (gestrichelte Linie) berücksichtigte Wellenleitermoden höherer Ordnung im SMA sowie Vergleich zur Referenzlösung des PBE (durchgezogene Linie). b) Relatives Residuum als Abweichung zwischen PBE und SMA für einen bis 25 berücksichtigte Wellenleitermoden.

Eine Homogenisierung als effektives Material ist aber nur dann möglich, wenn die Wellenausbreitung mit einem einzigen Wellenleitermode beschreibbar ist, da kein homogenes Material eine Konversion von Wellenleitermoden verursacht. Dies hat zur Folge, dass das Konzept der effektiven Materialparameter nur in Strukturgittern anwendbar ist, in denen keine Kopplung zu Wellenleitermoden höherer Ordnung erfolgt. Beispiele sind in [147] gegeben. Die Dispersionskurve aus dem dreidimensionalen periodisch berandeten Eigenwertproblem ist aufwendiger zu berechnen, da eine Folge von Eigenwertproblemen gelöst werden muss. Jedoch sind in dieser Lösung alle Wellenleitermoden höherer Ordnung inbegriffen, was beim Streumatrix-Ansatz nicht a priori gewährleistet ist. Die berechneten Ergebnisse demonstrieren das linkshändige Verhalten mit negativer Steigung in der Dispersionskurve. Diese und aufbauende Ergebnisse sind in [128, 129, 130, 133, 135, 142, 147, 148, 149] veröffentlicht worden.

5.2. Dispersionseigenschaften von Elementarzellen im Nanometerbereich

Die ersten Experimente, in denen *Metamaterialien* demonstriert wurden, deren elektromagnetisches Verhalten sich in einem begrenzten Frequenzband durch eine negative effektive Permeabilität beschreiben lässt, wurden im Mikrowellenbereich durchgeführt [109]. Zum Einsatz kommen dabei Elementarzellen im Zentimeterbereich, die deutlich kleiner als die Wellenlänge der einfallenden Welle ist [109]. Vorteilhaft ist an diesen Anordnungen, dass die Vernachlässigung von Verlusten in Dielektrika und Metallen meist keine großen systematischen Fehler verursacht und zur Produktion eines Strukturgitters technologisch nicht übermäßig große Anforderungen aufkommen. Seit Beginn der Forschungen rund um linkshändige *Metamaterialien* ist ein Wettbewerb mit dem Ziel entbrannt, Elementarzellen für immer höhere Frequenzen bzw. kürzere Wellenlängen zu erschaffen, bis schließlich das infrarote Spektrum des Lichts³

³Wellenlängen von etwa 780 nm bis 1 mm bzw. Frequenzen von 300 GHz bis rund 400 THz.

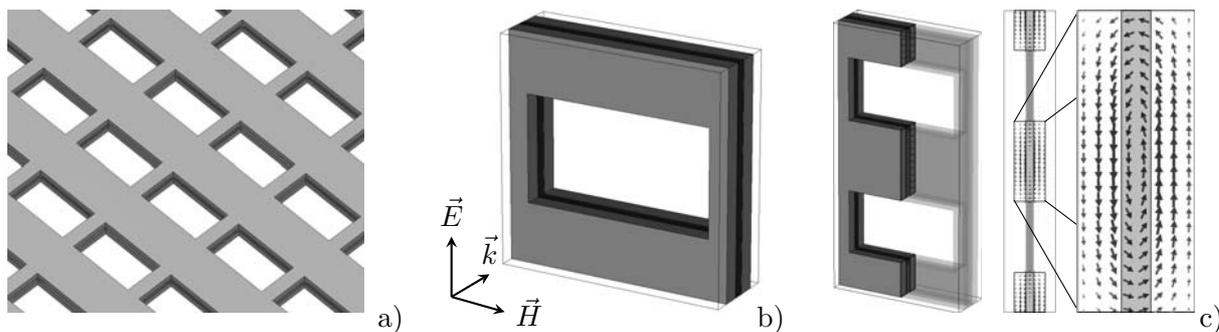


Abbildung 5.6.: a) Einlagiges Strukturgitter bestehend aus der *Fishnet*-Elementarzelle aus [98]. b) *Fishnet*-Elementarzelle mit der Polarisation einer ebenen Welle, bei der sich linkshändiges Verhalten zeigt. c) Vektordarstellung der Verschiebungsstromdichte in der Schnittebene von b). Schnitt durch zwei übereinander angeordnete Elementarzellen und Vektordarstellung der Verschiebungsstromdichte innerhalb der Struktur.

erreicht wurde [94, 12].

5.2.1. Linkshändige Metamaterialien bei optischen Frequenzen

Bei kleineren Wellenlängen muss weiterhin gemäß Abbildung 2.4 gewährleistet sein, dass die zugrunde liegende Elementarzelle des LHM klein gegenüber der Wellenlänge ist. Die Skalierbarkeit des *Split-Ring-Resonators* der linkshändigen Elementarzelle nach Abbildung 5.3 ist jedoch begrenzt, so dass dieses Design bei Frequenzen im infraroten Spektrum nur noch eingeschränkt in Frage kommt [105]. Ein effektiv magnetisches Materialverhalten im infraroten Spektrum lässt sich jedoch mit den Resonanzen dünner Metallstreifen darstellen [94] und die resultierende Elementarzelle wird als *Fishnet*-Elementarzelle bezeichnet. Ein LHM aus *Fishnet*-Elementarzellen ist in Abbildung 5.6a dargestellt. Abbildung 5.6b zeigt die dreilagige Elementarzelle aus [98], bestehend aus zwei Silberschichten⁴, die durch ein Dielektrikum⁵ getrennt sind, sowie die Polarisation der einfallenden ebenen Welle, für die ein linkshändiges Verhalten beobachtet wird. Die zirkulare Verschiebungsstromdichte in Abbildung 5.6c verursacht das effektiv magnetische Verhalten bei einer Wellenlänge von rund $1.435 \mu\text{m}$.

5.2.1.1. Modellierung der Fishnet-Elementarzelle

Im Gegensatz zu Kupfer im Mikrowellenbereich kann Silber im infraroten Bereich nicht mehr als ideal elektrisch leitend betrachtet werden. Stattdessen lassen sich die experimentell ermittelten Eigenschaften von Silber für nicht zu große Frequenzen in Form einer dispersiven Permittivität gut interpolieren. Dazu werden die Koeffizienten $\epsilon_\infty = 5$, $\alpha = 3.22 \times 10^{13}$ und $\omega_p^2 = 1.96 \times 10^{32}$ für das Dispersionsmodell nach DRUDE (2.37) gewählt [104]. In Abbildung 5.7a ist der Verlauf der Dielektrizitätskonstante für Silber über der Frequenz aufgetragen und es zeigt sich, dass der Realteil ϵ'_r stark frequenzabhängig ist. Die Eindringtiefe bestimmt die Länge, nach der eine einfallende Feldamplitude auf $e^{-1} \approx 0.37$ des Anfangswerts abge-

⁴600 nm \times 600 nm \times 45 nm Silber mit einer Apertur von 316 nm \times 100 nm.

⁵30 nm dickes MgF₂ mit relativer Dielektrizitätskonstante $\epsilon_r = 1.9044$.

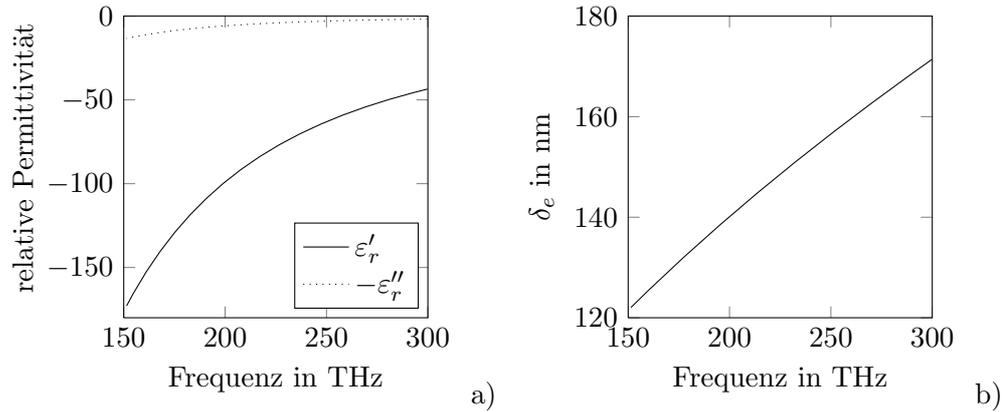


Abbildung 5.7.: a) Komplexe Permittivität von Silber. b) Aus a) berechnete Eindringtiefe δ_e .

klungen ist. Die Eindringtiefe $\delta_e = \sqrt{\frac{2}{\omega\mu\kappa}}$ lässt sich berechnen, wenn der Imaginärteil der Permittivität ϵ''_r als Leitfähigkeit interpretiert wird. Abbildung 5.7b zeigt die Eindringtiefe von Silber. Sie nimmt bei den hier betrachteten Frequenzen Werte von 120 bis 170 nm an – die Dicke der Silberlagen der *Fishnet*-Elementarzelle beträgt jedoch nur 45 nm, so dass sie teilweise transparent für einfallende elektromagnetische Wellen sind. Dies ist grundsätzlich gegenteilig zum Modell des idealen elektrischen Leiters, das im Mikrowellenbereich angewendet wird.

Die Gitterkonstante a_{per} im numerischen Modell der *Fishnet*-Elementarzelle wird willkürlich auf 200 nm gesetzt, um ein in Ausbreitungsrichtung⁶ unendliches Strukturgitter durch periodische Randbedingungen zu modellieren. Die Permeabilität beträgt μ_0 . In transversaler Richtung werden die Randbedingungen der Parallelplattenleitung entsprechend Abbildung 3.1b derart gesetzt, dass die in Abbildung 5.6b gezeigte ebene Welle ausbreitungsfähig ist. Darüber hinaus ist die Elementarzelle voll symmetrisch, so dass die Reziprozität für die Streumatrix ausgenutzt werden kann.

5.2.1.2. Streumatrix-Ansatz

Der Streumatrix-Ansatz ist bis auf die Voraussetzung, dass der verwendete Löser (z.B. CST [123]) mit dispersiven Materialien umgehen können muss, identisch zu Abschnitt 5.1 bzw. den Abschnitten 3.1.1 und 3.1.2.

5.2.1.3. Periodisch berandetes Eigenwertproblem

Für die Formulierung für das periodisch berandete Eigenwertproblem sind einige Vorüberlegungen notwendig. In Abschnitt 5.1 kam die Formulierung für verlustfreie Elementarzellen aus Abschnitt 3.2.1 zum Einsatz, die bei reeller Phasenverschiebung zu reellen Eigenfrequenzen führt. Einerseits ist die *Fishnet*-Elementarzelle nicht verlustfrei und andererseits ist neben den Verlusten auch der Realteil der Permittivität frequenzabhängig. Dies führt zu zwei möglichen Interpretationen, die in den folgenden Ansätzen münden:

⁶Entsprechend Abbildung 5.6b.

1. Die in der Elementarzelle vorhandenen Verluste verursachen, dass die oszillierende Strukturgitterwelle, der BLOCH-Mode, über der Zeit abklingt. Dies kann durch den Ansatz von $\varphi \in \mathbb{R}$ und $\omega \in \mathbb{C}$ modelliert werden, wie in Abschnitt 3.2.2 gezeigt worden ist und führt auf die Güte Q nach Gleichung (2.49). Dazu muss das polynomiale Eigenwertproblem (3.17) gelöst werden, was mit dem Löser aus Abschnitt 4.6.3 möglich ist [136, 140].
2. Der von der Systemanschauung motivierte Blick ist, dass die Amplitude des BLOCH-Modus bei Ausbreitung durch die Elementarzelle über ihrer Länge abklingen muss. Dies impliziert reelle Frequenzen $\omega \in \mathbb{R}$ sowie komplexe Phasenverschiebungen $\varphi \in \mathbb{C}$. Die numerische Formulierung auf Basis von FIT ist in Abschnitt 3.2.3 gegeben und führt auf die Lösung des polynomialen Eigenwertproblems (3.24), wie im Folgenden gezeigt wird [137].

5.2.2. Lösung der Eigenwert-Formulierung

Das polynomiale Eigenwertproblem $\Psi_{\mathbf{A}}$ (3.24) wird mit dem JACOBI-DAVIDSON-Verfahren aus Algorithmus 4.15 gelöst. In der FIT-Eigenwertformulierung (3.24) ist die reelle Frequenz die unabhängige Variable und die dispersiven Materialeigenschaften von Silber erfordern, dass die Materialmittlung (2.84) für jede Frequenz durchgeführt werden muss, wozu die unterschiedlichen dualen Flächenanteile senkrecht zur Materialgrenze benötigt werden. Dann lassen sich die Koeffizientenmatrizen der FIT-Formulierung für das quadratische Eigenwertproblem (3.24) bei gegebener Frequenz erstellen. Die in Abschnitt 4.1.4.1 vorgestellte Skalierung der Koeffizientenmatrizen sowie die Spektraltransformation verbessern die Kondition der Problemstellung. Während die Skalierung der Koeffizientenmatrizen nach Satz 4.1.28 direkt auf die Koeffizientenmatrizen von $\Psi_{\mathbf{A}}$ angewendet wird, wird darauf verzichtet, die Spektraltransformation nach Satz 4.1.29 für $\Psi_{\mathbf{A}}$ durchzuführen, da die Berechnung der involvierten $\|\cdot\|_2$ -Norm für hochdimensionale Koeffizientenmatrizen zu aufwendig ist. Stattdessen kann eine Spektraltransformation nach Satz 4.1.29 in jedem Iterationsschritt auf das projizierte, niederdimensionale Eigenwertproblem $\Psi_{\mathbf{M}} = \mathbf{V}_m^H \Psi_{\mathbf{A}} \mathbf{V}_m$ angewendet werden. Der Zielwert τ muss möglichst gut gewählt werden, um schnelle Konvergenz zu erreichen.

5.2.2.1. Polynomiales Jacobi-Davidson-Verfahren

Wenn kein Startvektor \mathbf{v}_0 bekannt ist, wird auf einen orthonormalen $n \times 6$ Unterraum $\mathbf{V}_{m=6}$ zurückgegriffen, dessen Einträge zunächst Zufallsvektoren sind, die orthonormalisiert werden. Das niederdimensionale Matrixpolynom $\Psi_{\mathbf{M}}$ entsteht durch Projektion von $\Psi_{\mathbf{A}}$ auf den Unterraum \mathbf{V}_m und α_{opt} wird zur Skalierung des Spektralradius entsprechend Satz 4.1.29 bestimmt. Das niederdimensionale Matrixpolynom $\Psi_{\mathbf{M}}$ wird mit der Linearisierung (4.2) in ein allgemeines Eigenwertproblem überführt und mittels QZ-Zerlegung aus Abschnitt 4.2.3 gelöst⁷. Für jedes berechnete RITZ-Paar⁸ wird das Residuum durch Einsetzen in $\Psi_{\mathbf{A}}$ gebildet, da nicht alle aus der Linearisierung gewonnenen RITZ-Paare gleichzeitig Eigenwert-Approximationen von $\Psi_{\mathbf{A}}$ sein müssen [46]. Aus den verbleibenden RITZ-Werten wird derjenige mit kleinstem Abstand zum Zielwert τ selektiert und weitere RITZ-Paare werden zur eventuell notwendigen Reduktion des Unterraums gespeichert. Während der ersten fünf Iterationen wird $\mathbf{p} = \Psi'_{\mathbf{A}}(\tau)\mathbf{u}$

⁷Routine `eig` in MATLAB [126].

⁸Vgl. Definition 4.3.1.

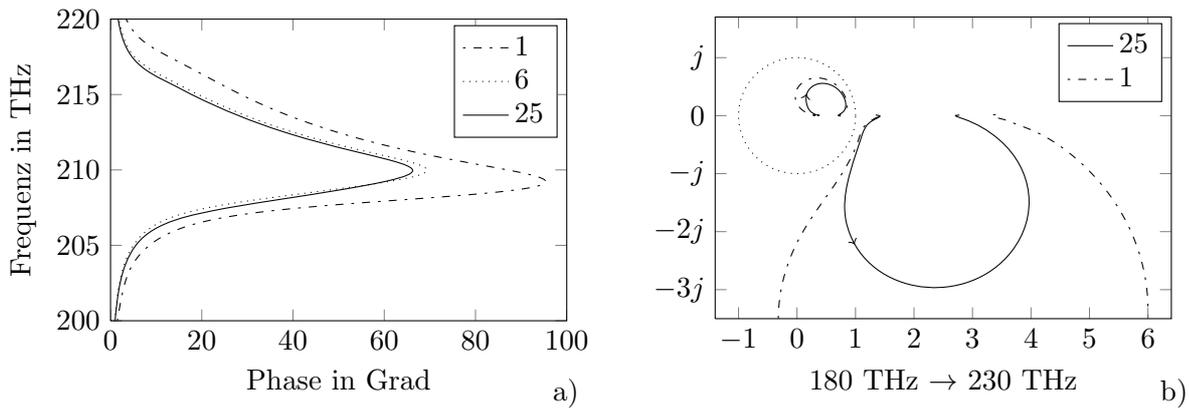


Abbildung 5.8.: a) Dispersionkurven aus dem Streumatrix–Ansatz für einen, sechs und 25 Wellenleitermoden. b) Ortskurven der gepaarten Eigenzustände γ in der komplexen Ebene.

statt $\mathbf{p} = \Psi'_A(\theta)\mathbf{u}$ gebildet, um die Konvergenz bei schlechten Startvektoren zu verbessern (vgl. Abschnitt 4.6.3).

Zur Lösung der zweistufigen Korrekturgleichung (4.18) wird das in Abschnitt 4.4.2 diskutierte BiCGStab(ℓ)-Verfahren aus [76] verwendet. Dabei liegt die Toleranzschwelle für das Residuum der Lösung bei 10^{-10} und es werden maximal 10 Iterationen erlaubt. Zur Vorkonditionierung kommt eine LU–Zerlegung von $\Psi_A(\tau)$ zum Einsatz (vgl. Abschnitt 4.4.3.1), die nur im ersten Iterationsschritt bestimmt werden muss. Die Lösung \mathbf{t} der Korrekturgleichung wird mit Algorithmus A.1 gegen den Unterraum \mathbf{V}_m orthogonalisiert. Ein Neustart kann zur Begrenzung seiner Dimension durchgeführt werden, siehe Abschnitt 4.3.4.

5.2.3. Numerische Ergebnisse

Um die Vergleichbarkeit von Streumatrix–Formulierung und Eigenwert–Formulierung zu gewährleisten, liegt beiden Verfahren das gleiche numerische Modell einer FIT–Diskretisierung aus CST MICROWAVE STUDIO [123] zugrunde. Durch zwei Symmetriebedingungen lässt sich die Anzahl der Freiheitsgrade auf ein Viertel reduzieren, so dass $23 \times 23 \times 15$ Rechengitterlinien zu 20314 Freiheitsgraden führen. Die Berechnung der Streumatrizen geschieht im Zeitbereich und die Berechnung des Dispersionsdiagramms erfolgt wie in den Abschnitten 3.1.1 und 3.1.2 beschrieben.

Abbildung 5.8a zeigt die aus dem Streumatrix–Ansatz berechneten Dispersionsdiagramme unter Berücksichtigung von einem, sechs sowie 25 Wellenleitermoden. Während der Unterschied zwischen einem und sechs Wellenleitermoden recht deutlich ausfällt, sind nur geringe Abweichungen der Dispersionskurven für sechs und 25 berücksichtigte Wellenleitermoden zu erkennen. Die Ortskurven in Abbildung 5.8b zeigen ausbreitungsfähige Eigenzustände γ des Transfer–Streumatrix–Eigenwertproblems (3.5) zwischen 180 THz und 230 THz. Da innerhalb der Elementarzelle eine deutliche Dämpfung auftritt, verlaufen die Kurven nur in der Nähe des Einheitskreises der komplexen Ebene, nicht aber auf dem Einheitskreis wie im verlustfreien Fall aus Abbildung 5.4b. Die in Abbildung 5.4a gezeigte Paarung der Eigenzustände für Ausbreitung in Vorwärts- und Rückwärtsrichtung zeigt sich in Abbildung 5.8b als Spiegelung am Einheitskreis. Zur Berechnung des Dispersionsdiagramms aus dem periodisch berandeten Eigenwertproblem, muss das resultierende, komplexe, polynomiale Eigenwertproblem sukzessive

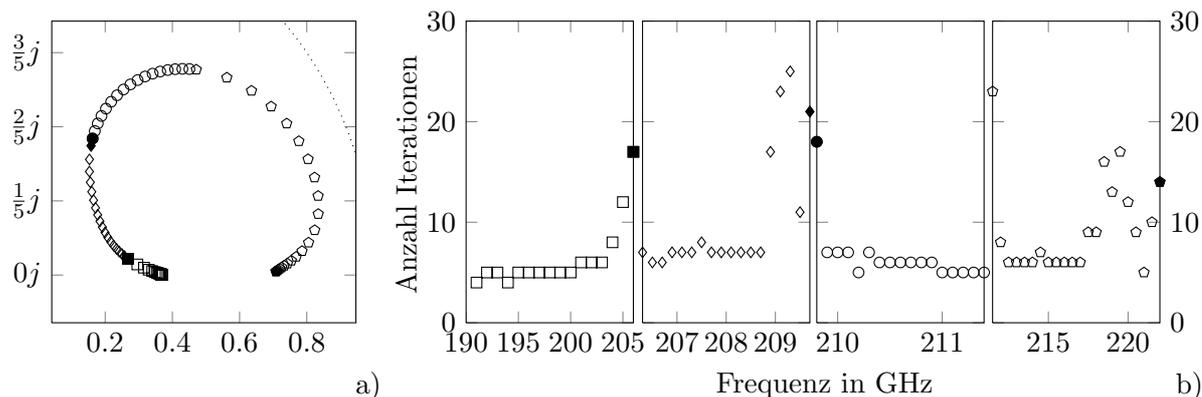


Abbildung 5.9.: a) Ortskurve der Eigenzustände, die aus der Berechnung mit dem polynomialen Eigenwertproblem entsteht. Vier Gruppen von Frequenzpunkten sind unterschiedlich markiert, der erste Frequenzpunkt ist jeweils ausgefüllt. b) Anzahl der benötigten Iterationen an den einzelnen Frequenzpunkten. Die Zuordnung zu a) ist über die Markierung möglich.

für mehrere Frequenzen gelöst werden. Für das polynomiale JACOBI-DAVIDSON-Verfahren werden maximal 200 Iterationen angesetzt, RITZ-Paare mit einem Residuum unter 10^{-9} werden als Eigenpaare akzeptiert, und die Unterraumdimension variiert zwischen sechs und 17.

Der Zielwert τ muss nah am gesuchten Eigenzustand gewählt werden, um rasche Konvergenz zu erhalten. Dazu kann zur Berechnung der Dispersionskurve ein einmal berechneter Eigenwert als Zielwert für die folgende Berechnung verwendet werden. Um darüber hinaus bessere Zielwerte und raschere Konvergenz zu erhalten, werden ab dem fünften berechneten Eigenzustand vorangegangene Werte verwendet, um ein Interpolationspolynom aufzustellen, welches bei der nächsten unberechneten Frequenz zum Erhalten eines neuen Zielwerts ausgewertet wird. Als Startvektor \mathbf{v}_0 bietet sich ebenfalls der zuletzt berechnete Eigenvektor an, was sich positiv auf das Konvergenzverhalten auswirkt. Die Ergebnisse der Dispersionskurve und des Konvergenzverhaltens sind in Abbildung 5.9 veranschaulicht. Die Frequenzpunkte zwischen 190 THz und 220 THz werden in vier Gruppen mit unterschiedlichen Abständen eingeteilt. Pro Frequenzpunkt ist ein Lauf des Löser für das polynomiale Eigenwertproblem erforderlich. Für den ersten Aufruf des Löser in jeder Gruppe wird ein Zielwert τ vorgegeben, während die folgenden Zielwerte τ wie oben beschrieben gewählt werden.

Die resultierende Dispersionskurve ist in Abbildung 5.9a gezeigt. Jeder Gruppe von Frequenzpunkten ist dabei eine eigene Markierung zugeordnet. Die ausgefüllten Markierungen geben dabei den jeweils ersten Wert einer Gruppe an. Abbildung 5.9b zeigt die Anzahl der benötigten Iterationen des Löser für das polynomiale Eigenwertproblem, wobei die dort gewählten Markierungen ihre Entsprechung in den Markierungen von Abbildung 5.9a finden, so dass eine Zuordnung der Frequenzen möglich wird. Die ersten Frequenzen jeder Gruppe⁹ benötigen eine größere Anzahl an Iterationen, da der verfügbare Zielwert τ etwas stärker vom gesuchten Eigenzustand abweicht. Sobald das Interpolationspolynom bessere Zielwerte liefert, sinkt die Anzahl der bis zur Konvergenz benötigten Iterationen deutlich.

In Abbildung 5.10a zeigt sich, dass die Berücksichtigung von 25 Wellenleitermoden im Streumatrix-Ansatz zu einer Dispersionskurve führt, die sehr gut mit dem Ergebnis des polynomialen Eigenwertproblems übereinstimmt. Somit kommen beide Verfahren zum gleichen

⁹Das sind die Markierungen, die den ausgefüllten Markierungen am nächsten sind.

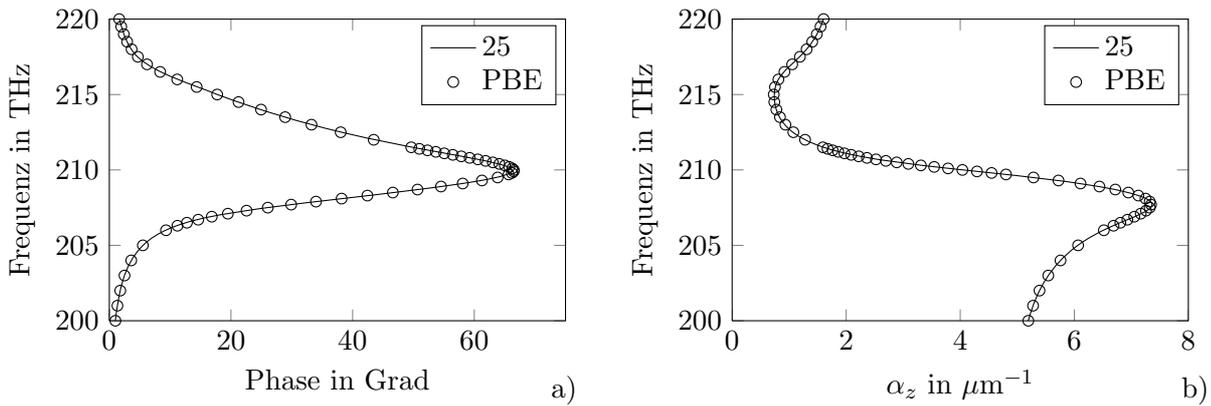


Abbildung 5.10.: Vergleich der Dispersionkurve (a) von Streumatrix-Ansatz mit 25 Wellenleitermoden und Lösungen des periodisch berandeten Eigenwertproblems (PBE) sowie der zugehörigen Dämpfungskonstanten (b).

Ergebnis, obwohl mit Streumatrizen aus Zeitbereichssimulationen und der Formulierung des periodisch berandeten Eigenwertproblems im Frequenzbereich verschiedene Wege zum Ziel verfolgt werden. Die Dispersionskurve zeigt im linkshändigen Frequenzband das erwartete Ergebnis mit negativer Steigung. Abbildung 5.10b zeigt die Frequenzabhängigkeit der Dämpfungskonstante, die berechnet werden kann, da der Phasenvorschub komplex angesetzt ist.

5.2.3.1. Unterschiede zu experimentellen Arbeiten

Während in der Literatur [98] das linkshändige Frequenzband zwischen 205 und 222 THz liegt, ergibt die Simulation ein Frequenzband zwischen 210 und 222 THz. Die Ursachen dafür liegen einerseits darin begründet, dass in der hier gezeigten Simulation ein in Ausbreitungsrichtung unendlich ausgedehntes Strukturgitter modelliert wird, während experimentell nur eine Lage dieser Schicht vermessen wurde. Darüber hinaus verursacht die Herstellung des *Fishnet*-Strukturgitters keine geraden Kanten der Silberlagen (siehe Abbildung 1 in [98]). Ferner wurden in [98] nicht die Parameter für Silber aus [104] verwendet, sondern der Parameter α im DRUDE-Dispersionsmodell (2.37) wurde als freier Parameter verwendet, mit dem die Ergebnisse der Simulation auf die Ergebnisse der Messung angepasst wurden.

5.2.4. Fazit

In diesem Abschnitt wird gezeigt, wie die Berechnung von Dispersionskurven auf verlustbehaftete Elementarzellen im Nanometerbereich ausgedehnt werden kann. Die Herausforderung besteht darin, dass Silber nicht mehr als ideal leitend modelliert werden kann, sondern Modelle für dispersive Dielektrika zur Materialbeschreibung notwendig sind.

Für den Streumatrix-Ansatz ergibt sich als einzigen Unterschied, dass der Löser, mit dem die Streumatrix berechnet wird, passende Dispersionsmodelle beherrscht. Dabei stellt sich analog zum Mikrowellenbereich stets die Frage, wie viele Wellenleitermoden höherer Ordnung in der Streumatrix berücksichtigt werden müssen. Bei verlustbehafteten Elementarzellen kann gut eine Zeitbereichssimulation zur Erzeugung der Streumatrix verwendet werden, da die

ins Rechengebiet eingespeiste Energie durch die vorhandenen Verluste schneller wieder abnimmt als im Fall idealer Leiter. Der Aufwand für die Berechnung des periodisch berandeten Eigenwertproblems für verlustbehaftete Elementarzellen mit Hilfe des vorgestellten Lösers auf Basis des JACOBI–DAVIDSON–Verfahrens ist hoch, da mit einer LU–Zerlegung ein guter Vorkonditionierer zur Lösung der Korrekturgleichung benötigt wird. Um die Dispersionskurve zu erhalten, müssen mehrere polynomiale Eigenwertprobleme sukzessive für verschiedene Frequenzen eines Frequenzbandes berechnet werden. Darüber hinaus ist die Lösung des polynomialen Eigenwertproblems nicht besonders robust, da Zielwerte τ nah am gesuchten Eigenwert vorgegeben werden müssen, um schnelle Konvergenz zu erreichen. Die Schätzung der nächsten Zielwerte ist nach der Lösung von etwa fünf polynomialen Eigenwertproblemen so gut, so dass die benötigte Anzahl an Iterationen deutlich reduziert ist. Die aus dem polynomialen Eigenwertproblem gewonnene Dispersionskurve beinhaltet alle Wellenleitermoden höherer Ordnung bereits in der Formulierung.

Der Streumatrix–Ansatz mit der Berücksichtigung einer hinreichenden Anzahl von Wellenleitermoden höherer Ordnung und die Lösung des polynomialen Eigenwertproblems, resultierend aus der periodisch berandeten Elementarzelle, liefern die gleiche Dispersionskurve in sehr guter Übereinstimmung. Teile der Ergebnisse sind in [131, 132, 134] veröffentlicht worden.

5.3. Berechnung von Wellenleitermoden in photonischen Kristallfasern

Das numerische Modell von Nanostrukturen, die mit transparenten Randbedingungen z.B. in Form der in Abschnitt 2.2.4.4 eingeführten PML versehen sind, ist durch die Formulierung für Wellenleiterquerschnitte entsprechend Abschnitt 3.3 gegeben. Durch den Abschluss des Rechengebiets mit der PML entstehen zusätzliche Eigenlösungen, deren Feldverteilung überwiegend in der PML lokalisiert ist. Ein Ansatz aus der Literatur zur Unterdrückung der unerwünschten Eigenlösungen ist beispielsweise [102], in dem die Approximationen der Eigenwerte nur aus einem geschickt gewähltem Bereich des kompletten Spektrums selektiert werden. Allerdings werden bei diesem Ansatz zusätzlich einige unerwünschte Eigenlösungen berechnet, die a posteriori aussortiert werden müssen. Im Folgenden wird an Anwendungsbeispielen die Leistungsfähigkeit der JACOBI–DAVIDSON–Verfahren mit erweitertem physikalisch motiviertem Selektionskriterium aus Abschnitt 4.5 demonstriert.

Photonische Kristallfasern (PCF, von engl. photonic crystal fibre) bestehen aus einem Glassubstrat, das von einer periodischen Anordnung längshomogener Löcher durchzogen ist [111]. Die Periodizität der Löcher ist im Innern unterbrochen, indem ein oder mehrere Löcher ausgelassen werden, so dass ein ausgefüllter Kernbereich entsteht, in dem nur Glas existiert, sowie ein umgebender Bereich, der von Löchern durchzogen ist. Beispiele sind in Abbildungen 5.11a und 5.11d gegeben. Anschaulich funktioniert die Führung der elektromagnetischen Welle wie in der herkömmlichen dielektrischen optischen Faser, in der es einen Kontrast des Brechungsindex zwischen Kern- und Mantelmaterial gibt. Wie alle optische Fasern können *photonische Kristallfasern* je nach Frequenz höchst multi-modal sein, wobei die Ausbreitungskonstanten jedes Wellenleitermodes nur sehr schwach variieren. Die Berechnung der Ausbreitungskonstanten mit dem Ansatz der Methode der finiten Integration aus Abschnitt 3.3, in dem zunächst nur konservative Randbedingungen im numerischen Modell zugelassen sind, verursacht Wellenleitermoden, die denen im Rechteckhohlleiter ähneln. Darüber hinaus sind

bei *photonischen Kristallfasern* nach Abbildung 5.11d die konservativen Randbedingungen nicht verwendbar, da schon die fundamentalen Wellenleitermoden Leckwellen umfassen, bei denen Leistung in transversale Richtung abgestrahlt wird. Daher wird mit der PML aus Abschnitt 2.2.4.4 eine transparente Randbedingung verwendet – je nach Kristallfaser für propagierende Wellenanteile bei auftretenden Leckwellen oder für evaneszente Wellenanteile bei stark geführten Wellenleitermoden.

Der Abschluss des FIT-Rechengebiets mit der PML entsprechend der Formulierung aus Abschnitt 3.3.1 verursacht wiederum, dass sich die geschichteten unterschiedlichen Materialien der PML wie ein Wellenleiter verhalten, dessen Ausbreitungskonstanten im gleichen Bereich liegen wie die der *photonischen Kristallfaser*. Dies führt zu einem Eigenwertspektrum in dem einerseits die geführten Wellenleitermoden der *photonischen Kristallfaser* und andererseits die ausbreitungsfähigen Wellenleitermoden innerhalb der PML enthalten sind. In den Abbildungen 5.11b und 5.11e sind Beispiele für die in der PML lokalisierten Wellenleitermoden gegeben, während die im Wellenleiterkern geführten Wellenleitermoden nach Abbildungen 5.11c und 5.11f tatsächlich von Interesse sind.

5.3.1. Apriori Wissen und physikalisch motiviertes Selektionskriterium

In [102] wird die Differenzierung der beiden auftretenden Wellenleitermodentypen allein anhand der Wellenzahl vorgenommen, was dazu führt dass dennoch einige unerwünschte Wellenleitermoden berechnet werden. Im Folgenden werden die auftretenden Wellenleitermodentypen anhand des Orts des Leistungsflusses unterschieden, so dass die unerwünschten Wellenleitermoden sicher ausgeschlossen werden können. Bei den geführten Wellenleitermoden nach Abbildungen 5.11c und 5.11f ist der Leistungsfluss im Kern der Faser konzentriert und die Feldstärke klingt über der Querschnittsfläche mit zunehmenden Radius schnell ab, während die durch die PML verursachten Wellenleitermoden am Rand des Rechengebiets geführt werden. Das physikalisch motivierte Selektionskriterium, das für die Auswahl der aktuellen RITZ–Werte in den Algorithmen 4.11–4.13 zum Einsatz kommt, lässt sich gemäß Abschnitt 4.5.1.1 folgendermaßen definieren: Aus jedem RITZ–Vektor werden die normierten POYNTING–Vektoren in der Querschnittsebene berechnet. Dann wird die z -Komponente der POYNTING–Vektoren jeweils über das zuvor festgelegte Kerngebiet integriert, um die Information zu erhalten, welcher Anteil der Leistung der jeweiligen Wellenleitermodenapproximation im festgelegten Kerngebiet geführt wird. Hieraus ergibt sich das physikalisch motivierte Selektionskriterium

$$\frac{\int_{A_1} S_z \, dA}{\int_{A_{\text{gesamt}}} S_z \, dA} = \frac{P_{\text{Kern}}}{P_{\text{gesamt}}} > \iota. \quad (5.1)$$

Sofern der Leistungsfluss im festgelegten Kerngebiet ein vordefiniertes Limit ι erreicht, wird der zugehörige RITZ–Vektor als aktuelle Approximation des Eigenvektors für die weitere Iteration des JACOBI–DAVIDSON–Verfahrens verwendet. Dabei ist zu beachten, dass es sich stets um Approximationen der Eigenvektoren handelt, deren Residuum während der Iteration noch groß ist. So werden drei Dinge während des JACOBI–DAVIDSON–Verfahrens sichergestellt:

1. Der Abstand der berechneten Ausbreitungskonstanten der Wellenleitermoden ist klein zum Zielwert τ .

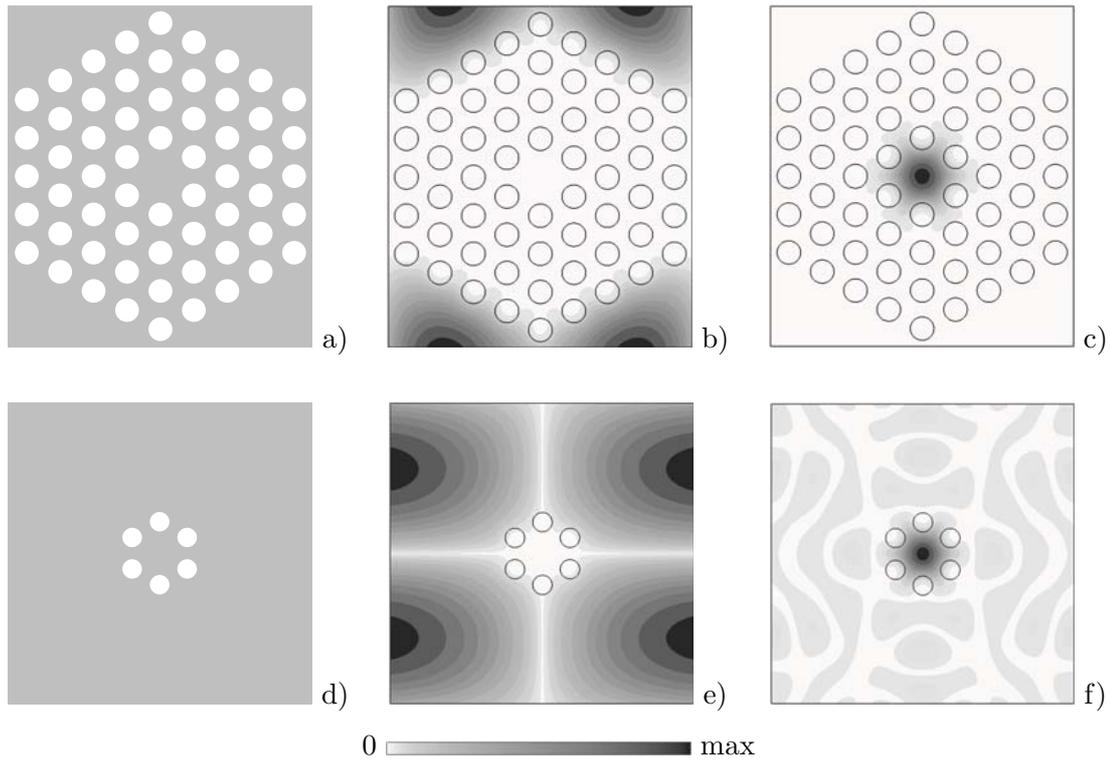


Abbildung 5.11.: *Photonische Kristallfasern* mit a) vier Lochlagen und d) einer Lochlage. Betrag der transversalen elektrischen Feldstärke von jeweils exemplarischen Wellenleitermoden, die in der PML ausbreitungsfähig sind (b und e) sowie von den geführten Wellenleitermoden (c und f), die im Kern geführt werden.

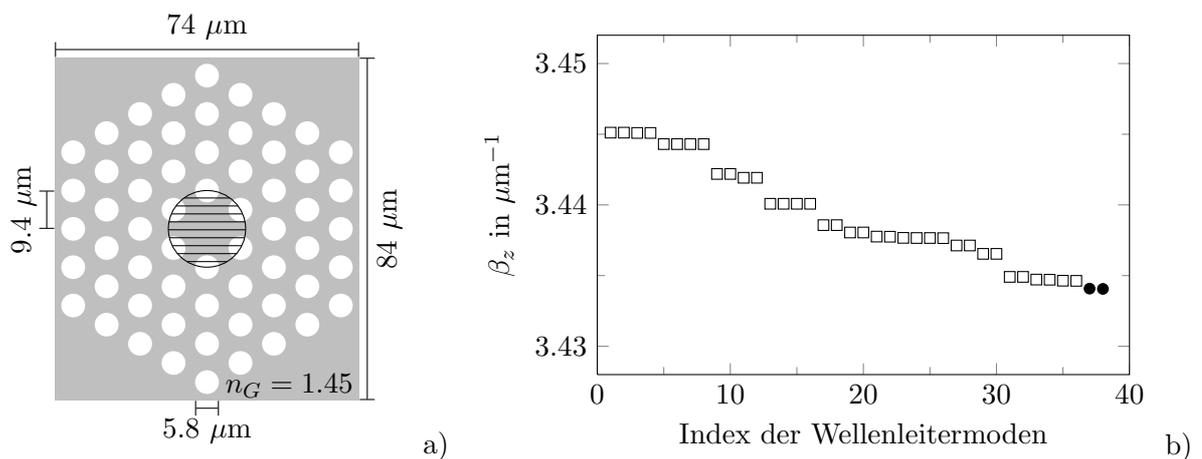


Abbildung 5.12.: a) Abmessungen der *photonischen Kristallfaser* mit vier Lochlagen. Die schraffierte Fläche gibt das Integrationsgebiet A_1 für Gleichung (5.1) an. b) Berechnete Ausbreitungskonstanten für die Struktur aus a). Die Werte \bullet gehören zu den gesuchten Wellenleitermoden mit Feldverteilungen analog zu Abbildung 5.11c, die Werte \square haben Feldverteilungen, die analog zu Abbildung 5.11b in der PML geführt werden.

2. Der Leistungsfluss der berechneten Wellenleitermoden ist im Kern der *photonischen Kristallfaser* konzentriert.
3. Konvergenz ist erreicht, wenn die Norm des Residuums kleiner als die vorgegebene Schranke ϵ ist.

5.3.2. Numerische Konfiguration

Es werden zwei verschiedene *photonische Kristallfasern* betrachtet, die beide bei $2\ \mu\text{m}$ Wellenlänge betrieben werden. Die geometrischen Abmessungen der *photonischen Kristallfaser* mit vier Lochlagen sind in Abbildung 5.12a gegeben, für die *photonische Kristallfaser* mit nur einer Lochlage werden die Außenmaße von $90\ \mu\text{m} \times 90\ \mu\text{m}$ verwendet. Die Ausbreitungskonstanten in den *photonischen Kristallfasern* werden in FIT mit der Eigenwertformulierung für die zweidimensionale Querschnittsebene aus Abschnitt 3.3 berechnet. Beispiele für den Betrag der zu erwartenden transversalen Feldverteilung sind in Abbildung 5.11 enthalten. Das numerische Modell der Querschnittsfläche wird mit CST MICROWAVE STUDIO [123] erzeugt und in MATLAB [126] um eine PML für evaneszente transversale Wellenleitermoden bei der *photonischen Kristallfaser* mit vier Lochlagen erweitert¹⁰. Bei der *photonischen Kristallfaser* mit einer Lochlage ist diese Erweiterung eine PML für transversal propagierende Wellenanteile¹¹. Das numerische Modell der *photonischen Kristallfaser* mit vier Lochlagen besteht aus 64438 Freiheitsgraden, die aufgrund der PML für evaneszente Wellen reell sind, und das Modell der *photonischen Kristallfaser* mit einer Lochlage besteht aus 16012 komplexen Freiheitsgraden. Als Voraussetzung an die Eigenwertlöser wird nun gefordert, jeweils die ersten beiden ausbreitungsfähigen Moden zu berechnen. Abbildung 5.12b zeigt exemplarisch den Teil des Spektrums mit dem kleinsten Realteil¹² bzw. den daraus berechneten Ausbreitungskonstanten β_z des numerischen Modells der *photonischen Kristallfaser* nach Abbildung 5.12a. Die ersten 36 Moden werden innerhalb der PML-Region geführt, und erst die Moden mit der Ordnungsnummer 37 und 38 erfüllen das physikalisch motivierte Selektionskriterium (5.1). Für die *photonische Kristallfaser* nach Abbildung 5.11d haben die ersten beiden geführten Wellenleitermoden die Ordnungsnummern 135 und 136 bezüglich des Endes des Spektrums mit dem kleinsten Realteil.

Von der numerischen Seite betrachtet muss das aus der FIT-Diskretisierung resultierende unsymmetrische Standardeigenwertproblem aus (3.38) gelöst werden, das je nach PML-Typ reell- oder komplexwertig ist. Die Spektren der Standardeigenwertprobleme werden mit Algorithmus 4.1 zunächst verschoben und skaliert, um die Konditionszahl zu verbessern. Da mit dem kleinsten Realteil als Zielwert äußere Eigenwerte berechnet werden, kommt Algorithmus 4.9 mit und ohne dem modifizierten Selektionskriterium aus Algorithmus 4.12 zum Einsatz. Die maximale Anzahl von Iterationen innerhalb des JACOBI-DAVIDSON-Verfahrens wird auf 160 für die *photonische Kristallfaser* mit vier Lochlagen und auf 600 für die Faser mit einer Lochlage gesetzt. Zur Berechnung mit dem modifizierten Selektionsprozess aus Algorithmus 4.12 wird ein kreisrundes Integrationsgebiet A_1 mit Radius $9.4\ \mu\text{m}$ über dem Kern entsprechend Abbildung 5.12a definiert, auf dem das Selektionskriterium (5.1) ausgewertet wird. Dazu werden die Komponenten der elektrischen Gitterspannung $\bar{\mathbf{e}}$ innerhalb des Gebiets A_1 verwendet. Die zur Berechnung des Leistungsflusses notwendigen magnetischen

¹⁰Jeweils vier Rechengitterzellen PML.

¹¹Da aus der Literatur bekannt ist, dass bei diesem Typ der *photonischen Kristallfaser* bereits der fundamentale Wellenleitermode Leckwellen besitzt.

¹²Der Eigenwertlöser erhält als Zielwert $\tau := \min(\text{Re}\{\theta_i\})$.

Gitterspannungen $\hat{\mathbf{h}}$ sind auf dem dualen Rechengitter allokiert, so dass die Fläche des berechneten Leistungsflusses jeweils durch eine primäre und duale Rechengitterkante aufgespannt wird. RITZ–Paare werden als Eigenpaare akzeptiert, wenn die Norm des Residuums unter 10^{-12} fällt. Die Korrekturgleichung (4.7) wird mit einem direkten Verfahren gelöst, was zwar numerisch aufwendig ist, aber zu quadratischer Konvergenz führen kann [46].

Zwei Parameter, die das Konvergenzverhalten des JACOBI–DAVIDSON–Verfahrens beeinflussen, werden im Folgenden untersucht.

Der erste Parameter ist die minimale und maximale Dimension des Unterraums m_{\min} und m_{\max} , die bei l_{\max} gesuchten Eigenlösungen entweder $m_{\min} = l_{\max} + 5$ und $m_{\max} = l_{\max} + 10$ oder $m_{\min} = 1$ und $m_{\max} = 150$ beträgt. Im letztgenannten Fall findet entsprechend selten ein Neustart mit reduziertem Unterraum statt, so dass die Dimension des Unterraums sukzessive ansteigt. Der zweite Parameter, der starken Einfluss auf das Konvergenzverhalten hat, ist die Wahl des Startvektors \mathbf{v}_0 . Sofern keinerlei Information über das erwartete Ergebnis vorliegt, lässt sich ein Zufallsvektor als Startvektor verwenden. Eine Wahl, die sich bereits grob am Ergebnis orientiert, kann erzielt werden, wenn nur Einträge in \mathbf{v}_0 auf Eins gesetzt, die zu Komponenten gehören, die innerhalb des Integrationsgebiets A_1 (schraffiert in Abbildung 5.12a) liegen und somit im physikalisch motivierten Selektionskriterium ausschlaggebend sind. Bei der Berechnung mit dem modifizierten Selektionskriterium aus Algorithmus 4.12 ist darüber hinaus die Wahl der Schranke ι_l von Bedeutung, mit der entschieden wird, ob das Selektionskriterium erfüllt wird oder nicht.

5.3.3. Performanzvergleich

Die vier möglichen Parameterkombinationen sind in Tabelle 5.1 aufgelistet, in der jeder Parametersatz eine Zeile darstellt. Die beiden rechten Spalten stellen die Referenzergebnisse dar. Sie geben an, wie viele Iterationen mit Algorithmus 4.9 notwendig sind, um die beiden ersten kerngeführten ausbreitungsfähigen Wellenleitermoden zu berechnen. Dabei werden im Fall der PCF mit vier Lochlagen 36 unbenötigte Wellenleitermoden (vgl. Abbildung 5.12b) und im Fall der PCF mit einer Lochlage 134 unbenötigte Wellenleitermoden berechnet, bevor die beiden gesuchten Wellenleitermoden auftreten. Dabei ist bemerkenswert, dass der unmodifizierte Algorithmus 4.9 für die Parametersätze zwei (\diamond) und vier (\square) aus Tabelle 5.1 innerhalb von 600 Iterationen nicht gegen die beiden gesuchten Wellenleitermoden konvergiert, wenn die maximale Dimension des Unterraums oft reduziert wird. Dies ist unabhängig vom gewählten Startvektor \mathbf{v}_0 , allerdings werden einige Wellenleitermoden gefunden, die in der PML geführt werden. Mit den Parametersätzen eins (\circ) und drei (\times) aus Tabelle 5.1 wird gefordert, dass die Dimension des Unterraums nur alle 150 Schritte reduziert wird. Es folgt, dass 136 Eigenlösungen (inklusive der beiden kerngeführten Wellenleitermoden) in 301 bzw. 312 Iterationen berechnet werden.

Für die Performanzanalyse mit modifiziertem Selektionskriterium nach Algorithmus 4.12 wird in Abbildungen 5.13 für jede Struktur zusätzlich der Schwellwert ι_l untersucht, der die Schranke zum Erkennen kerngeführter Wellenleitermoden darstellt. Dazu wird für jeden der vier Fälle aus Tabelle 5.1 eine Datenreihe in den Abbildungen 5.13 erzeugt, bei dem ι_l Werte zwischen 0.3 und 0.99 annimmt, wobei die Markierungen der Datenreihen der Zeilenzuordnung entsprechend Spalte 1 in Tabelle 5.1 zu entnehmen sind. Eine Null in der Anzahl der Iterationen in den Abbildungen 5.13 steht für eine Parameterkombination, in welcher der JACOBI–DAVIDSON–Löser die beiden gesuchten kerngeführten Wellenleitermoden nicht in-

nerhalb der Maximalanzahl von Iterationen findet. Es stellt sich heraus, dass die Wahl der Schranke ι_l in Algorithmus 4.12 Einfluss auf das Konvergenzverhalten hat, wenngleich es wenige Werte für ι_l gibt, bei denen Algorithmus 4.12 völlig fehlschlägt. Stärkeren Einfluss auf die Anzahl der Iterationen hat hingegen die Wahl der Parameterkonstellation aus Tabelle 5.1.

Die Ergebnisse für die PCF mit vier Lochlagen sind in Abbildung 5.13a gegeben. Im Falle der Parametersätze drei (\times) und vier (\square) aus Tabelle 5.1, wenn der Startvektor überwiegend Komponenten im Kern besitzt, führt Algorithmus 4.12 für alle Werte von ι_l zum Ziel. Darüber hinaus brauchen die Parametersätze zwei (\diamond) und vier (\square), in denen die maximale Unterraumdimension begrenzt ist, weniger Iterationen als die beiden Parametersätze, in denen auf eine Reduktion des Unterraums verzichtet wird. Im besten Fall von Parametersatz vier (\square) terminiert Algorithmus 4.12 bereits nach fünf Iterationen und die beiden gesuchten kerngeführten Wellenleitermoden liegen in hoher Genauigkeit vor. Dies ist unabhängig von der Wahl von ι_l .

Die Ergebnisse für die PCF mit nur einer Lochlage sind in Abbildung 5.13b gegeben. Der Parametersatz zwei (\diamond) aus Tabelle 5.1 schlägt vollkommen fehl und zwar unabhängig von ι_l . Bessere Ergebnisse resultieren aus den restlichen Parametersätzen, in denen Algorithmus 4.12 zu sicherer Konvergenz führt. Dass keiner der Parametersätze aus Tabelle 5.1 für $\iota_l = 0.99$ konvergiert, liegt an einem zu klein gewählten Integrationsgebiet A_1 entsprechend Abbildung 5.11a, das zum Messen der Lokalisierung des Leistungsflusses in (5.1) zum Einsatz kommt. Bei der PCF nach Abbildung 5.11d gibt es keine kerngeführten Wellenleitermoden, die mindestens 99% der Leistung innerhalb des Integrationsgebiets führen, so dass Algorithmus 4.12 erwartungsgemäss fehlschlagen muss. Vielversprechende Resultate ergibt der Parametersatz vier (\square) aus Tabelle 5.1, in dem eine häufige Reduktion der Unterraumdimension mit einem Startvektor mit Komponenten im Kern gepaart ist: Nur Sechs Iterationen sind notwendig, um die ersten beiden kerngeführten Wellenleitermoden zu finden.

5.3.4. Fazit

In diesem Abschnitt wird gezeigt, dass das physikalisch motivierte Selektionskriterium Algorithmus 4.12 zur Wahl des jeweils aktuellen RITZ-Paars in Algorithmus 4.9 zu drastisch verbessertem Konvergenzverhalten führt, wenn damit die Ausbreitungskonstanten in Wellenleitern berechnet werden, deren Eigenwertspektrum durch den (physikalisch notwendigen Abschluss) mit einer transparenten Randbedingung um eine Vielzahl unerwünschter Wellenleitermoden angereichert ist. Die gezielte Berechnung der gesuchten Wellenleitermoden wird durch die Implementierung der a priori verfügbaren Information, dass die Leistung der gesuchten Wellenleitermoden im Kern geführt wird, erreicht. Die Wahl der im physikalisch motivierten Selektionskriterium vorhandenen Schranke ι_l stellt sich als weitgehend unkritisch heraus, so dass der vorgestellte Algorithmus in der Praxis gut funktioniert. Desweiteren ist die Kontrolle über den Startvektor und die Unterraumdimension im vorgestellten Beispiel ausschlaggebend. Das beste Konvergenzverhalten wird mit einem Startvektor erreicht, der die zu erwartende Feldverteilung bereits grob widerspiegelt. In Tabelle 5.2 wird die Anzahl der benötigten Iterationen von Algorithmus 4.9 mit und ohne physikalisch motiviertem Selektionskriterium Algorithmus 4.12 gegenüber gestellt, so dass das Einsparpotenzial von Rechenzeit- und Speicherressourcen sichtbar wird. Diese Vorteile sind jedoch von der elektromagnetischen Problemstellung abhängig. Die Validierung von Algorithmus 4.12 zeigt sich durch den Vergleich mit dem Algorithmus 4.9. Diese Ergebnisse sind in [139, 143, 138, 144] veröffentlicht.

Reihe	Startvektor- einträge \mathbf{v}_0	Unterraum- dimension m_{\min}, m_{\max}	Anzahl der Iterationen mit JDQR Algorithmus 4.9	
			PCF vier Lochlagen $l_{\max} = 38$	PCF eine Lochlage $l_{\max} = 136$
1 ○	zufällig	1, 150	92	301
2 ◇	zufällig	$l_{\max} + 5, l_{\max} + 10$	81	> 600
3 ×	1 auf A_1 , 0 sonst	1, 150	82	312
4 □	1 auf A_1 , 0 sonst	$l_{\max} + 5, l_{\max} + 10$	69	> 600

Tabelle 5.1.: Parametersätze zu Abbildung 5.13 und Referenzergebnisse von Algorithmus 4.9. Parameter sind die Wahl des Startvektors \mathbf{v}_0 und die Grenzen der Unterraumdimension; l_{\max} ist die Anzahl der zu berechnenden Eigenlösungen.

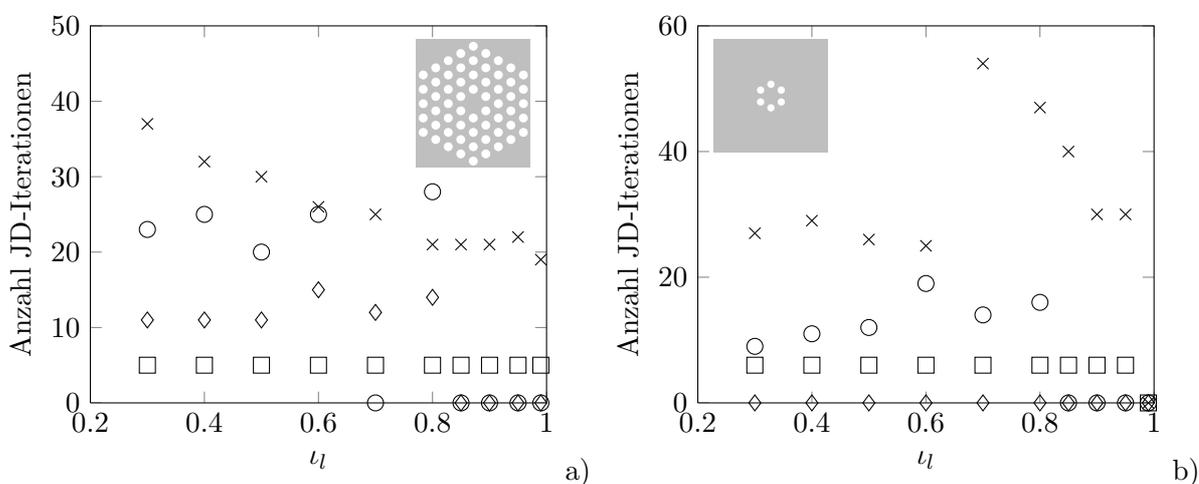


Abbildung 5.13.: Anzahl von Iterationen, die Algorithmus 4.9 mit physikalisch motiviertem Selektionskriterium aus Algorithmus 4.12 in Abhängigkeit von ν_l benötigt, um die ersten beiden kerngeführten Wellenleitermoden von den *photonischen Kristallfasern* a) nach Abbildung 5.11a und b) nach Abbildung 5.11d zu berechnen. Die Parameterzuordnung gilt entsprechend Tabelle 5.1. Eine 0 in der Anzahl der Iterationen gibt an, dass die ersten beiden kerngeführten Wellenleitermoden nicht innerhalb der geforderten Maximalanzahl von Iterationen gefunden werden.

PCF	Minimale Anzahl an Iterationen zur Berechnung zweier kerngeführter Wellenleitermoden	
	Algorithmus 4.9	Algorithmus 4.9 mit Algorithmus 4.12
vier Lochlagen	69	5
eine Lochlage	301	6

Tabelle 5.2.: Minimale Anzahl an Iterationen zur Berechnung zweier kerngeführter Wellenleitermoden mit JDQR-Algorithmus 4.9 sowie JDQR-Algorithmus 4.9 in Kombination mit dem erweiterten Selektionsprozess Algorithmus 4.12.

5.4. Eigenmoden isolierter Nanostrukturen im Freiraum

Numerische Modelle von Strukturen, die sich im Freiraum befinden, können nicht immer durch konservative Randbedingungen abgeschlossen werden, da unter Umständen Feldanteile abgestrahlt werden. Daher kommt in diesen Fällen die PML aus Abschnitt 2.2.4.4 zur Modellierung einer transparenten Randbedingung zum Einsatz, was in der Formulierung aus Abschnitt 3.4 mündet. Die gesuchten Eigenmoden liegen dabei im Innern des Spektrums¹³ und die äußeren Eigenmoden sind häufig so niederfrequent, dass sie im kompletten Rechengebiet schwingen, ohne dabei die geometrische Struktur aufzulösen. A priori verfügbare Informationen können jedoch helfen, die gesuchte Lösung schnell zu berechnen und dabei unerwünschte Lösungen auszublenden. In den folgenden Abschnitten wird einerseits anhand einer nicht-konvergierten Zeitbereichssimulation und einer verfügbaren analytischen Näherungslösung gezeigt, wie diese Informationen zur Generierung von Zielwerten und Startvektoren verwendet werden können. Andererseits lassen sich die a priori verfügbaren Informationen im erweiterten Selektionsprozess des JACOBI–DAVIDSON–Verfahrens aus Abschnitt 4.5 einsetzen, um das Auftreten der gesuchten Lösung während der Iteration zu bevorzugen.

5.4.1. Nanoresonatoren aus photonischen Kristallen

Das räumliche Einfangen¹⁴ von Licht und seine Konzentration ist eine wichtige Voraussetzung zur Anregung von Licht-Materie-Interaktionen in der nichtlinearen Optik [117]. Vom elektrotechnischen Blickwinkel ist dazu ein Resonator notwendig, der bei der geforderten Frequenz des Lichts gerade eine Resonanz besitzt. Mit photonischen Nanoresonatoren nach Abbildung 5.14a lassen sich diese Anforderungen erfüllen. Sie zeigen darüber hinaus eine hohe Güte (2.48) [91]. Nanoresonatoren bestehen aus einem dielektrischen Substrat und einem regulären Strukturgitter aus Löchern mit hexagonaler Elementarzelle. Der eigentliche Resonator entsteht durch das Auslassen einiger Löcher in der Mitte der Struktur, analog zur photonischen Kristallfaser aus Abschnitt 5.3. Experimentell wurden Nanoresonatoren in [91] untersucht. Das umgebende Medium kann meist als Freiraum modelliert werden, so dass abhängig von der Frequenz Strahlungsmoden auftreten können. Der Entwurf und die Optimierung von Nanoresonatoren kann mittels Simulationstechnik durchgeführt werden, so dass für diesen Zweck Zeit- und Frequenzbereichsmethoden vorgestellt werden¹⁵, wobei das Rechengebiet mit der transparenten Randbedingung abgeschlossen wird. Der Schwerpunkt liegt dabei auf der möglichst genauen Bestimmung der Resonanzfrequenzen, Güten und Eigenmoden, während nichtlineares und dispersives Materialverhalten vernachlässigt wird.

5.4.1.1. Zeitbereichsverfahren

Die FIT–Formulierung in Abschnitt 2.2.2 führt auf ein explizites Zeitintegrationsverfahren, bei dem in jedem Schritt Matrix-Vektormultiplikationen durchzuführen sind. Die hohen Güten von Resonatoren verursachen jedoch, dass die ins Rechengebiet eingespeiste Energie über der Zeit nur langsam abklingt, was dazu führt, dass lange Zeitspannen simuliert werden müssen, bevor das Frequenzspektrum mit einer DFT aus dem Zeitsignal bestimmt werden kann. Für

¹³Die Lage im Innern ist unabhängig von der Verschiebung des n_p -fach entarteten Nulleigenwerts.

¹⁴engl. confinement

¹⁵Die Zeitbereichssimulationen sind in Kooperation mit M.Sc. Christoph Claßen erstellt worden.

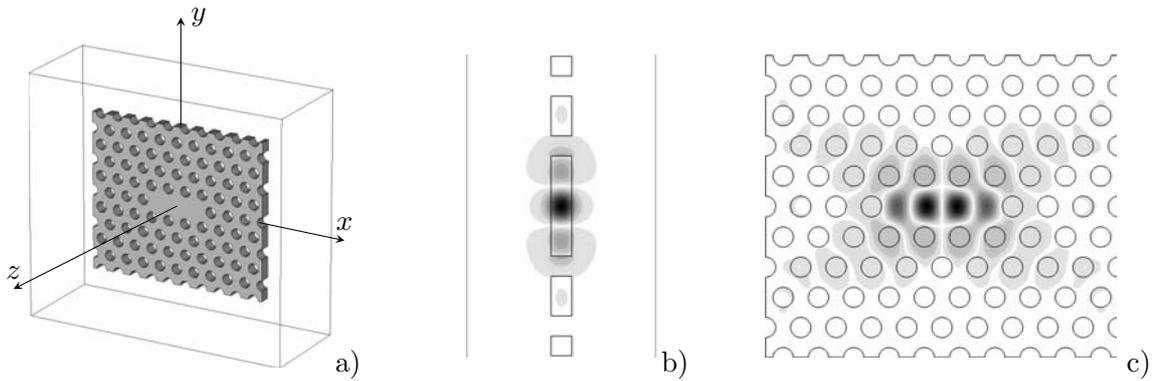


Abbildung 5.14.: a) Nanoresonator als dielektrische Platte mit Löchern mit Lochabstand a_{per} . b) Betrag der E_y -Komponente in der y - z -Ebene und c) Betrag der H_z -Komponente in der x - y -Ebene des $H_{z,4,1}$ -Eigenmodes, mit Maxima in der Nähe von $(a_{\text{per}}/2; 0; 0)$.

die Extraktion von Resonanzfrequenzen und Güten können Methoden der Signalverarbeitung eingesetzt werden, um diese Informationen aus nicht vollständig abgeklungenen Feldern zu berechnen [113]. Hier wird auf die Methode der *harmonischen Inversion* zurückgegriffen [106], die mit dem frei erhältlichen `harminv`-Code vorliegt [125]. Der Ansatz setzt dabei voraus, dass das Zeitsignal $f(t)$ in einem begrenzten Frequenzbereich als eine endliche Summe abklingender Sinus-Schwingungen dargestellt werden kann

$$f(t) = \sum_l A_l \sin(\omega_l t + \varphi_l) e^{-\alpha_l t}.$$

Dazu müssen die Koeffizienten $\{A_l, \omega_l, \varphi_l, \alpha_l\}$ bestimmt werden. Die Resonanzfrequenzen sind unmittelbar die ω_l und die Güten lassen sich aus ω_l und den Dämpfungskonstanten α_l bestimmen. In der Anwendung auf Nanoresonatoren zeigt sich, dass die Resonanzen bereits nach kurzer Simulation des Zeitsignals sicher von `harminv` identifiziert werden können. Die Anregung der Zeitbereich-Simulation muss so ausgelegt sein, dass alle gesuchten Eigenmoden tatsächlich angeregt werden. Es muss daher zumindest a priori verfügbare Informationen über die Feldverteilungen der gesuchten Eigenmoden geben. Neben der Anregung müssen die Orte der Aufzeichnung des Zeitsignals ebenfalls an der Feldverteilung der gesuchten Eigenmoden ausgerichtet sein, so dass gegebenenfalls mehrere Zeitbereichssimulationen notwendig sind.

5.4.1.2. Frequenzbereichsformulierung zur Eigenmoden-Simulation

Der Vorteil der Eigenwert-Formulierung ist, dass keine speziell gewählte Anregung entworfen werden muss, jedoch ist der numerische Aufwand im Allgemeinen groß, da eine Folge von linearen Gleichungssystemen gelöst werden muss. Der Abschluss des Rechengebiets mit einer PML gemäß Abschnitt 2.2.4.4 verursacht, dass das Eigenwertproblem der Resonanzfrequenzen in der FIT-Formulierung (3.41) nichtlinear wird. Da jedoch Eigenwerte innerhalb des Spektrums berechnet werden sollen, ist in jedem Fall eine gut gewählte Schätzfrequenz als Zielwert notwendig, so dass sich (3.41) in (3.42) linearisieren lässt¹⁶. Die resultierende Systemmatrix ist komplex und unsymmetrisch, so dass die JACOBI-DAVIDSON-Verfahren aus Algorithmus 4.9 und 4.10 zum Einsatz kommen. Um rasche Konvergenz zu erreichen, sind Vorkonditionierer und gute Startvektoren notwendig.

¹⁶Der resultierende Fehler wird in Abschnitt 5.4.3 untersucht.

5.4.1.3. Kombinierte Vorgehensweise

In einer kombinierten Vorgehensweise lassen sich Startvektoren für die Berechnung des Eigenwertproblems aus einer Zeitbereichssimulation gewinnen. Der Ablauf ist in Abbildung 5.15 dargestellt, wobei das übergeordnete Ziel ist, möglichst schnell das zugrunde liegende Eigenwertproblem zu lösen. Zunächst werden die Orte der Anregung und der Aufzeichnung für das Zeitbereichsverfahren justiert. Dies geschieht z.B. durch Dipole oder Stromschleifen im Rechengitter, deren Form sich an der Feldverteilung des gesuchten Eigenmodes orientiert. Dem folgt eine kurze Zeitbereichsrechnung von etwa drei Pulsbreiten. Auf das erhaltene Zeitsignal kann die *harmonische Inversion* angewendet werden, woraus sich Approximationen der Resonanzfrequenzen ergeben. Ist die gesuchte Resonanzfrequenz nicht auffindbar, sollte die Feldverteilung untersucht werden, um eine bessere Wahl von Anregungsschleife und Aufzeichnungsschleife treffen zu können. Dieser Vorgang muss in jedem Fall wiederholt werden, bis die gesuchte Resonanzfrequenz auftaucht. Ist sie gefunden, folgt eine zweite Simulation einiger Pulsbreiten im Zeitbereich, bei der die dreidimensionale Feldverteilung im Rechengebiet bei der gesuchten Resonanzfrequenz f_0 mittels DFT aufgezeichnet wird. Da zum Ende der Zeitbereichssimulation die im Rechengebiet verbleibende Energie noch nah am Maximum ist, stellt die aufgezeichnete dreidimensionale Feldverteilung eine grobe Lösung des Eigenmodes dar. Diese wird als Startvektor im JACOBI-DAVIDSON-Verfahren verwendet, wobei die aus der *harmonischen Inversion* ermittelte Resonanzfrequenz als Zielwert verwendet wird.

5.4.1.4. Numerisches Modell

Als Beispiel wird ein Ausschnitt der Struktur aus [91] verwendet¹⁷, die in Abbildung 5.14a gezeigt ist und verlustfrei modelliert wird. Aus der Literatur ist bekannt, dass die E_y -Komponente im Ursprung ein Maximum hat, so dass diese Komponente des Zeitsignals dort aufgezeichnet wird [91] (siehe Abbildung 5.14b. Darüber hinaus kann das Rechengebiet mit drei Symmetriebedingungen auf ein Achtel seiner Ausgangsgröße reduziert werden ($xy, yz: H_{\text{tan}} = 0, xz: E_{\text{tan}} = 0$) und die verbleibenden Ränder werden um transparente Randbedingungen (PML, siehe Abschnitt 2.2.4.4) erweitert.

Für die Zeitbereichssimulation kommt CST STUDIO SUITE [123] zum Einsatz. Der Ort der Anregung ist essentiell für das erfolgreiche Auffinden des gesuchten Eigenmodes und es muss darauf geachtet werden, dass möglichst keine weiteren unerwünschten Eigenmoden angeregt werden. Um die in Abbildung 5.14c gezeigte H_z -Komponente anzuregen, wird eine quadratische Stromschleife mit Kantenlänge $0.1\mu\text{m}$ in der x - y -Ebene um $(a_{\text{per}}/2; 0; 0)$ zentriert. Wie durch Abbildung 5.15 motiviert, ergibt sich diese Position für Anregung und Aufzeichnung aus mehreren Schleifendurchläufen, wobei die Zeitbereichssimulation jeweils für drei Pulsbreiten durchgeführt wird. Die im Rechengebiet vorhandene Restenergie ist dabei erst auf -4.9 dB vom Maximalwert abgeklungen. Aufgezeichnete dreidimensionale Feldverteilungen, die mittels der DFT gewonnen werden, können typischerweise bei Werten von -40 dB bis -60 dB Restenergie im Rechengebiet als konvergiert angesehen werden. Das resultierende Zeitsignal von drei Pulsbreiten Dauer wird mit der *harmonischen Inversion* analysiert, wobei die gesuchte minimale Resonanzfrequenz auf 180 THz und die maximale Resonanzfrequenz auf 200 THz gesetzt wird. Mit Eigenfrequenzen aus diesem Bereich versucht die *harmonische Inversion* das Zeitsignal zu approximieren.

¹⁷Strukturgitterkonstante $a_{\text{per}} = 0.42 \mu\text{m}$ (Abstand Lochmittelpunkte), Lochradius $\varrho_0 = 0.29a_{\text{per}}$, Dielektrikum mit $\varepsilon_r = 11.56$ und einer Dicke von $t = 0.6a_{\text{per}}$.

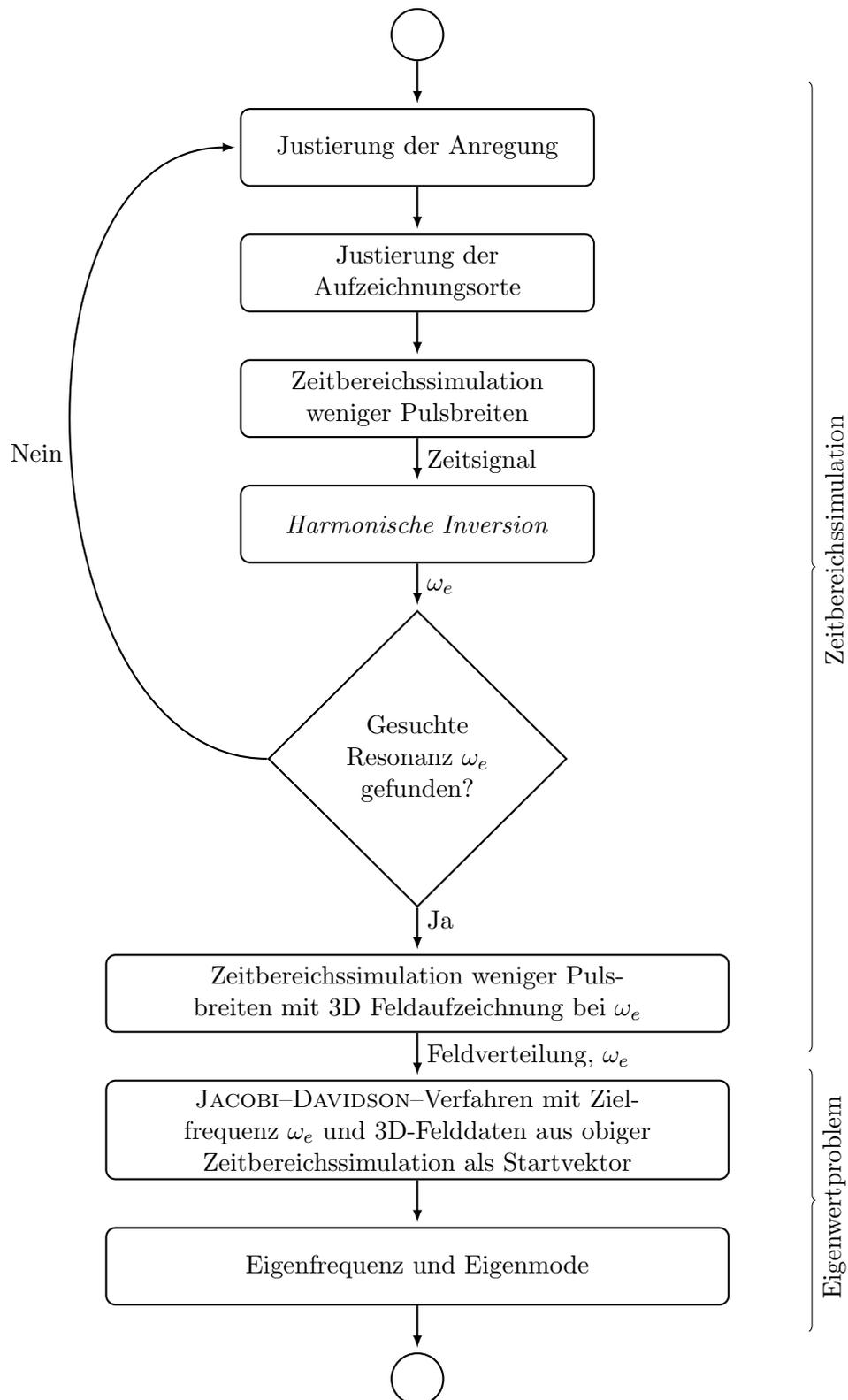


Abbildung 5.15.: Startvektorerzeugung durch Zeitbereichssimulation zur Verwendung in der Berechnung des Eigenwertproblems.

Der Eigenwert-Formulierung aus Abschnitt 3.4 liegt – mit Ausnahme der PML – das gleiche numerische Modell wie der Zeitbereichssimulation zugrunde. Als PML werden in MATLAB [126] sieben absorbierende Lagen hinzugefügt, um die das Rechengebiet in aufsteigender Koordinatenrichtung erweitert wird. Die PML wird bei der mit der *harmonischen Inversion* berechneten Schätzfrequenz ω_e ausgewertet, so dass sich das Standard eigenwertproblem (3.42) ergibt. Als Löser werden die JACOBI–DAVIDSON–Verfahren aus Algorithmus 4.9 und 4.10 verwendet. Die Systemmatrix wird zunächst einer Spektraltransformation nach Algorithmus 4.1 unterzogen.

Zur Vorkonditionierung der Korrekturgleichung kommt eine unvollständige LU–Zerlegung mit verschiedenen Toleranzparametern nach Abschnitt 4.4.3.2 zum Einsatz. Dem geht eine Umsortierung nach dem CUTHILL–MCKEE–Verfahren aus Abschnitt 4.4.4 voraus, um die Anzahl der bei der unvollständigen LU–Zerlegung zusätzlich auftretenden Matrixeinträge zu begrenzen. Im Zusammenspiel mit der unvollständigen LU–Zerlegung liefert das BiCGStab–Verfahren aus Abschnitt 4.4.2 gute Ergebnisse. Ein RITZ–Paar wird als Eigenpaar akzeptiert, wenn die Norm des Residuums unterhalb 10^{-8} liegt. Dies sollte in maximal 100 Iterationen des JACOBI–DAVIDSON–Verfahrens erreicht werden.

5.4.1.5. Ergebnisse

Zunächst wird ein numerisches Modell aus einer Diskretisierung mit zehn Rechengitterlinien pro Wellenlänge verwendet, das auf 190400 Unbekannte führt. In Tabelle 5.3 sind die Ergebnisse für die Resonanzfrequenz f und die Güte Q dargestellt, die aus der *harmonischen Inversion* und der DFT mit verschiedenen langen Zeitsignalen als Eingabedaten berechnet sind. Dabei wird für die *harmonische Inversion* stets die volle Länge des simulierten Zeitsignals verwendet. Das Band, in dem Resonanzfrequenzen extrahiert werden, ist auf 180 bis 200 THz begrenzt. Es zeigt sich, dass die *harmonische Inversion* schon bei Zeitbereichssimulationen von wenigen Pulsbreiten Dauer nah an das Ergebnis der dreidimensionalen Eigenmodenberechnung von 192.10 THz und $Q=3502$ herankommt. Die in Tabelle 5.3 dargestellten Resonanzfrequenzen schwanken dabei mit den unterschiedlichen simulierten Zeitdauern, während sich die Güte ab 20 Pulsbreiten Dauer kaum noch ändert. Bei den Ergebnissen der DFT kann zwar die Resonanzfrequenz bereits geschätzt werden, jedoch weicht die Güte auch für eine größere Anzahl an simulierten Pulsbreiten noch deutlich vom Endergebnis ab. Unterschiede zwischen Eigenwert- und Zeitbereichssimulation sind einerseits auf die unterschiedlichen PML–Implementierungen in CST STUDIO SUITE und MATLAB [126] zurückzuführen und andererseits ist der Einfluss der Dispersion aufgrund der Zeitdiskretisierung¹⁸ in der Eigenwertformulierung nicht vorhanden.

Zur Berechnung des dreidimensionalen Eigenmodes werden die JACOBI–DAVIDSON–Verfahren JDQR und JDQZ aus Abschnitt 4.3 mit unterschiedlichen Vorkonditionierern für die Korrekturgleichung verwendet. In Abbildung 5.16 wird die Anzahl benötigter JD-Iterationen für zufallswertbesetzte Startvektoren sowie für Startvektoren aus unkonvergierten Zeitbereichssimulationen bei verschiedenen ILU-Parametern gegenüber gestellt. Mit der Verwendung von Startvektoren, die aus der Zeitbereichssimulation gewonnen werden, lässt sich die benötigte Anzahl an JD-Iterationen erheblich reduzieren (Datenreihen \square und \circ). Die Anzahl an benötigten JD-Iterationen liegt für JDQR und JDQZ dicht beieinander und der Einfluss der Qualität der ILU erweist sich – zumindest im betrachteten Parameterbereich – als rela-

¹⁸Zur Dispersion durch Zeitdiskretisierung siehe z.B. [32].

Dauer in Pulsbreiten f in THz Q E/E_{\max} in dB	<i>Harmonische Inversion</i>					DFT	
	3	8	20	40	261	40	261
	192.22	192.46	192.22	192.06	192.26	192.17	192.18
	3432	3480	3506	3506	3509	1184	3487
	-4.9	-6.1	-8.7	-13.1	-60	-13.1	-60

Tabelle 5.3.: Die *harmonische Inversion* des im Ursprung aufgezeichneten E_y -Zeitsignals liefert nach Zeitbereichssimulationen von wenigen Pulsbreiten Dauer bereits gute Approximationen der Resonanzfrequenz f und Güte Q . Die dreidimensionale Eigenmodeberechnung liefert $f = 192.10$ THz mit $Q=3502$. Die Güten, die mittels DFT bestimmt werden, weichen für nicht-konvergierte Zeitbereichssimulationen, bei denen die im Rechengebiet vorhandene Energie E/E_{\max} groß ist, stark ab.

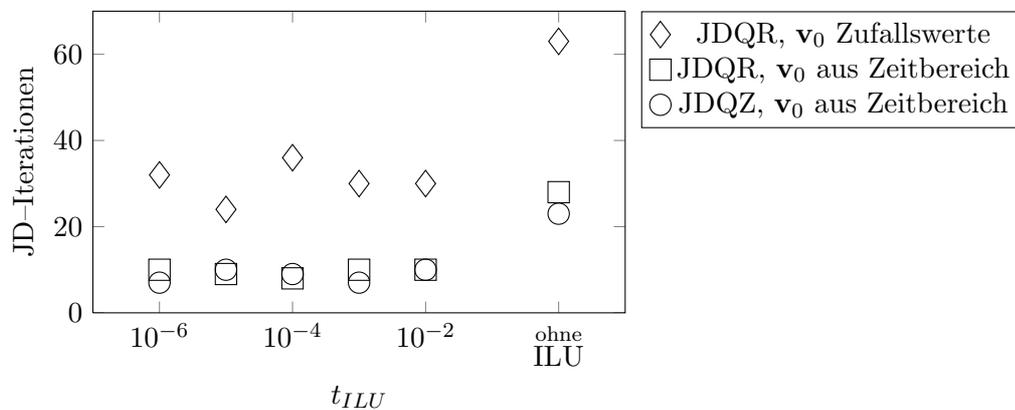


Abbildung 5.16.: Anzahl der benötigten JACOBI-DAVIDSON-Iterationen in Abhängigkeit vom ILU-Schwellwert t_{ILU} . JDQR-Verfahren mit Startvektoren aus Zufallszahlen (\diamond) und Zeitbereichsrechnung (\square). Startvektoren aus Zeitbereichsrechnung und JDQZ (\circ).

Linien pro λ_{el} Unbekannte	10	15	25	35	50
3D Eigenwertproblem f_0 in THz	192.095	191.549	191.397	191.369	191.397
Q	3502	3750	3875	3938	4063
<i>Harmonische Inversion</i> f_0 in THz	192.217	191.445	191.403	191.468	191.502
Q	3432	4962	4879	3668	4335

Tabelle 5.4.: Ergebnisse der *harmonischen Inversion* des Zeitsignals von drei Pulsbreiten Dauer und der dreidimensionalen Eigenmodeberechnung: Resonanzfrequenzen und Güten dargestellt in Abhängigkeit von der Rechengitterauflösung.

tiv gering. Dies bedeutet, dass sich die Korrekturgleichung mit ILU-Vorkonditionierung im betrachteten Parameterbereich nicht wesentlich besser lösen lässt, zumal der BiCGStab als iterativer Löser für das lineare Gleichungssystem stets die maximale Anzahl von 300 Iterationen erreicht. Die benötigte Zeit pro JD-Iteration steigt auf dem verwendeten Rechnersystem¹⁹ von rund 50 s ohne ILU-Vorkonditionierer auf über 280 s bei Verwendung des besten und damit am dichtesten besetzten Vorkonditionierers an. Dabei ist noch nicht berücksichtigt, dass die Berechnungszeit der ILU mit kleiner werdender Toleranzschwelle ansteigt. Für feine Diskretisierungen, die große Systemmatrizen mit sich bringen, muss daher abgewogen werden, inwieweit sich die Verwendung des ILU-Vorkonditionierers lohnt, da mit dem guten Startvektor aus der Zeitbereichssimulation die Anzahl benötigter JD-Iterationen schon mehr als halbiert wird. Abbildungen 5.14b und 5.14c zeigen den Betrag $|E_y|$ und $|H_z|$ des berechneten Eigenmodes in zwei Schnittebenen.

In Tabelle 5.4 ist die Abhängigkeit der Resonanzfrequenzen und Güten von der Rechengitterauflösung dargestellt²⁰. Resonanzfrequenzen und Güten ändern sich dementsprechend mit fortschreitender Gitterverfeinerung immer weniger. Während sich bei der dreidimensionalen Eigenmodeberechnung die Resonanzfrequenzen monoton dem Wert der feinsten Rechengitterauflösung annähern, schwanken die Ergebnisse aus der *harmonischen Inversion* etwas stärker. Jedoch ist die erreichte Genauigkeit für die hier betrachtete Generierung von Eingabedaten für den Eigenwertlöser ausreichend.

5.4.1.6. Fazit

In diesem Abschnitt wird anhand eines Nanoresonators gezeigt, wie sich die Berechnung von Eigenmoden beschleunigen lässt. Dazu werden Resonanzfrequenzen in einem abklingenden Zeitsignal mit Hilfe der Methode der *harmonischen Inversion* berechnet. Bei der erhaltenen Resonanzfrequenz wird die dreidimensionale Feldverteilung während einer Zeitbereichssimulation weniger Pulsbreiten aufgezeichnet. Die dreidimensionale Feldverteilung enthält genügend Anteile vom gesuchten Eigenmode, so dass sie als Startvektor in einer folgenden Lösung des Eigenwertproblems für das gesamte Rechengebiet dienen kann. Es zeigt sich, dass sich der numerische Aufwand, gemessen in der Anzahl der Iterationen des JACOBI-DAVIDSON-Verfahrens, deutlich verringern lässt. Teile der präsentierten Ergebnisse sind in [127] veröffentlicht.

5.4.2. Resonanzen in Microdisks

Als *Microdisks* werden dielektrische Scheiben bezeichnet, deren Durchmesser typischerweise wenige Mikrometer beträgt. Ihre Eigenmoden werden zum räumlichen Einfangen von Licht verwendet, wodurch die Anregung von Licht-Materie-Kopplung ermöglicht wird [117], wozu Eigenmoden mit einer hohen Güte benötigt werden [121]. Sogenannte *whispering gallery modes*²¹ weisen typischerweise eine hohe Güte auf und werden in der Nähe des äußeren Radius geführt. Die Anwendbarkeit von *Microdisks* als Laser oder Einzelphotonenquellen ist dadurch eingeschränkt, dass sie das emittierte Licht nicht gerichtet, sondern gleichförmig über den Um-

¹⁹Vier Intel Xeon 7350 mit insgesamt 16 Kernen.

²⁰Für eine allgemeine Betrachtung der Konvergenzeigenschaften der Methode der finiten Integration siehe z.B. [26, Anhang A.1].

²¹Für die Berechnung von *whispering gallery modes* dielektrischer Scheiben mit einem Durchmesser von etwa einem halben Meter siehe [145].

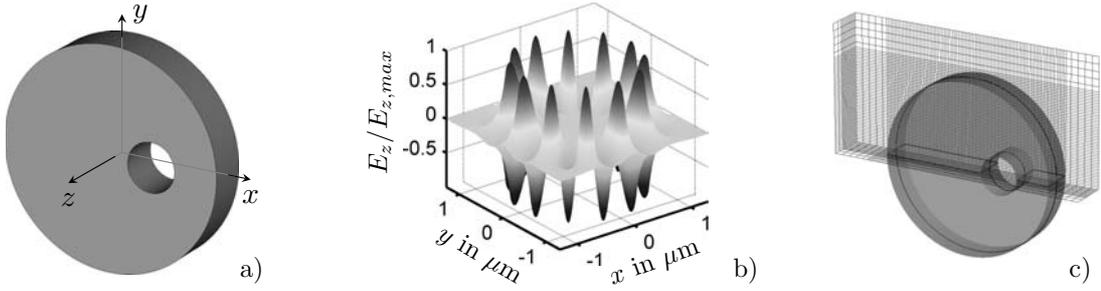


Abbildung 5.17.: a) Struktur der gelochten *Microdisk*. b) Näherungslösung in der $z = 0$ Ebene für den $\text{TM}_{1,12}$ -Eigenmode, berechnet mit der analytischen Approximation aus [116]. c) Rechengitter (ohne PML-Lagen) der Diskretisierung der gelochten *Microdisk*: Zwei Symmetrieebenen reduzieren die Anzahl der Unbekannten auf ein Viertel der ursprünglichen Anzahl.

fang verteilt abstrahlen²². Um eine gerichtete Abstrahlung bestimmter Eigenmoden zu erreichen, werden verschiedene geometrische Modifikationen an *Microdisks* vorgenommen. Beliebige und technologisch möglich ist das Einbringen eines radial versetzten zylindrischen Aussparung nach Abbildung 5.17a [121]. Die Modifikation der ursprünglichen Geometrie verursacht, dass eine Berechnung mit Verfahren für rotationsymmetrische Anordnungen ausscheidet, so dass dreidimensionale Diskretisierungen zum Einsatz kommen. Die am Rand geführten *whispering gallery modes* haben nur eine schwache Wechselwirkung mit dem Loch während Eigenmoden, deren Feldverteilung sich über einen radial größeren Bereich erstreckt, ein Abstrahlverhalten mit stark azimuthaler Abhängigkeit aufweisen können.

5.4.2.1. Analytische Approximation

Zur Berechnung einer Näherungslösung der Resonanzfrequenz von *whispering gallery modes* in homogenen dielektrischen Scheiben kann ein analytischer Ansatz verfolgt werden, der auf der Verwendung eines effektiven Brechungsindex $n_{d,\text{eff}}$ basiert [93, 116]. Die Modellbildung findet in zwei Stufen statt. Eine ausführliche Herleitung befindet sich in [93].

Im ersten Schritt wird die Wellenausbreitung in einer unendlich ausgedehnten dielektrischen Platte mit Brechungsindex $n_d = \sqrt{\epsilon_r}$ betrachtet, siehe Abbildung 5.18a. Das umgebende Medium wird mit Freiraum und $n_0 = 1$ angesetzt. Die geführte Wellenausbreitung innerhalb der dielektrischen Platte der Dicke Δz kann mit Hilfe eines effektiven Brechungsindex $n_{d,\text{eff}}$ beschrieben werden. Für Moden mit niedrigster Ordnung in z -Richtung folgt aus der Erfüllung der Stetigkeitsbedingungen der MAXWELL'schen Gleichungen bei $z = \pm\Delta z/2$ das analytische Eigenwertproblem zu

$$\tan\left(\frac{k\Delta z}{2}\sqrt{n_d^2 - n_{d,\text{eff}}^2}\right) = \nu_1 \frac{n_{d,\text{eff}}^2 - 1}{n_d^2 - n_{d,\text{eff}}^2} \quad \text{mit} \quad \nu_1 = \begin{cases} n_d^2, & \text{für TM Moden,} \\ 1, & \text{für TE Moden.} \end{cases} \quad (5.2)$$

Hierbei ist $k \in \mathbb{R}$ die Wellenzahl. Moden, die eine E_z -Komponente in der Ebene $z = 0$ besitzen, werden als TM-Mode bezeichnet. TE-Moden haben analog dazu H_z -Komponenten in der Ebene $z = 0$.

²²Zu Formulierungen in der ϱ - z Ebene siehe [145]. Zu FIT-Diskretisierungen mit kreiszylindrischen Rechengittern siehe [97, 118, 162].

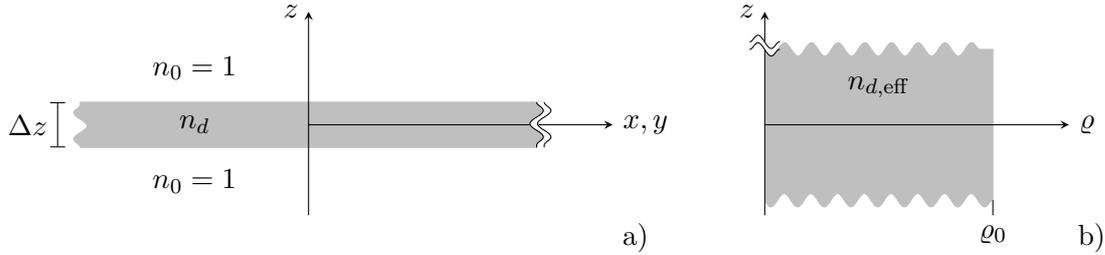


Abbildung 5.18.: a) Unendlich ausgedehnte dielektrische Platte mit Brechungsindex n_d im Querschnitt. b) Axial unendlich ausgedehnter dielektrischer Zylinder mit Brechungsindex $n_{d,eff}$.

Im zweiten Schritt wird ein axial unendlich ausgedehnter dielektrischer Zylinder betrachtet, der aus dem homogenen effektiven Brechungsindex $n_{d,eff}$ besteht (siehe Abbildung 5.18b). Für den Ansatz zylindersymmetrischer Felder folgt aus der Erfüllung der Stetigkeitsbedingungen auf der Zylinderoberfläche bei $\rho = \rho_0$ das analytische Eigenwertproblem

$$\nu_2 \frac{J'_m(n_{d,eff} \underline{k}_{m,l_\rho} \rho_0)}{J_m(n_{d,eff} \underline{k}_{m,l_\rho} \rho_0)} = \frac{H_m^{(2)'}(n_{d,eff} \underline{k}_{m,l_\rho} \rho_0)}{H_m^{(2)}(n_{d,eff} \underline{k}_{m,l_\rho} \rho_0)} \quad \text{mit} \quad \nu_2 = \begin{cases} n_{d,eff}, & \text{für TM Moden,} \\ \frac{1}{n_{d,eff}}, & \text{für TE Moden.} \end{cases} \quad (5.3)$$

Hierbei sind J_m BESSEL- und $H_m^{(2)}$ HANKEL-Funktionen zweiter Art der Ordnung m . $H'_m(x)$ und $J'_m(x)$ sind die jeweils ersten Ableitungen der Funktionen nach ihrem Argument x . Die Lösung ist die komplexe Wellenzahl \underline{k}_{m,l_ρ} mit azimuthaler Ordnung m und radialer Ordnung l_ρ . Eine Approximation der Lösung für die *Microdisk* kann erhalten werden, wenn \underline{k}_{m,l_ρ} Gleichung (5.3) und $\text{Re}\{\underline{k}_{m,l_\rho}\}$ Gleichung (5.2) erfüllt. Aus \underline{k}_{m,l_ρ} lassen sich die komplexe Resonanzfrequenz und die Güte nach (2.49) berechnen. Eine exemplarische Lösung der z -Komponente des $\text{TM}_{1,12}$ -Eigenmode ist in Abbildung 5.17b gegeben. Die erzielbare Genauigkeit in den Resonanzfrequenzen liegt abhängig von Material- und Geometrieparametern sowie der Modenordnung bei einigen Prozent [145]. Eine Approximation der Feldverteilung in der Ebene $z = 0$ lässt sich erhalten, indem unter Verwendung einer Lösung der Gleichungen (5.2) und (5.3) im Innern des Zylinders die BESSEL-Funktion und im Außenraum die HANKEL-Funktion entsprechender azimuthaler Ordnung angesetzt werden. Die komplexen Amplituden ergeben sich nach Abschnitt 2.1.2.2.

5.4.2.2. Numerische Berechnung

Das aus der Diskretisierung mit der Methode der finiten Integration gewonnene numerische Modell für eine dielektrische Scheibe mit transparenten Randbedingungen kann nach Abschnitt 3.4 auf das Standard eigenwertproblem (3.42) linearisiert werden²³. Zur numerischen Berechnung kommen einerseits die JACOBI-DAVIDSON-Verfahren für das Standard eigenwertproblem (Algorithmen 4.9 und 4.10) mit erweitertem Selektionskriterium aus Abschnitt 4.5 zum Einsatz. Desweiteren wird ein Verfahren auf Basis der RAYLEIGH-Quotienten-Iteration (Algorithmus 4.4) verwendet. Das Ziel bleibt dabei stets, einen einzigen Eigenmode zu berechnen, dessen Feldverteilung zumindest teilweise durch die analytische Approximation angenähert werden kann.

²³Der resultierende Fehler wird in Abschnitt 5.4.3 untersucht.

Jacobi–Davidson–Verfahren. Die Auswahl des jeweils aktuellen RITZ–Werts im JACOBI–DAVIDSON–Verfahren erfolgt üblicherweise anhand der Eigenwerte des in den Suchraum projizierten Eigenwertproblems. Die in Abschnitt 4.5 gezeigten Erweiterungen des Selektionsprozesses verwenden darüber hinaus die RITZ–Vektoren, um zu entscheiden, mit welchem RITZ–Paar die Iteration fortgesetzt wird. Bei herkömmlicher Extraktion des RITZ–Paares (JDQR–Algorithmus 4.9), die für die Berechnung äußerer Eigenwerte geeignet ist, kommt der erweiterte Selektionsprozess nach Algorithmus 4.12 zum Einsatz. Der erweiterte Selektionsprozess nach Algorithmus 4.13 wird bei harmonischer Extraktion des RITZ–Paares verwendet und ist für innere Eigenwerte etwas besser²⁴ geeignet (JDQZ–Algorithmus 4.10). Die Ausgestaltung des Kriteriums für den erweiterten Selektionsprozess basiert auf dem Skalarprodukt eines Gewichtungsvektors \mathbf{f} , der die a priori vorhandenen Informationen enthält, mit dem jeweiligen RITZ–Vektor \mathbf{u}_i

$$\mathbf{f}^H \mathbf{u}_i > \iota_i. \quad (5.4)$$

Der Gewichtungsvektor \mathbf{f} kann entweder aus vorangegangenen numerischen Berechnungen oder analytischen Approximationen bestehen und muss nicht vollbesetzt sein. Ein vielversprechender RITZ–Vektor für die Verwendung in den folgenden JACOBI–DAVIDSON–Iterationen zeichnet sich dadurch aus, dass das Skalarprodukt in (5.4) größer als ein zuvor definierter Schwellwert ι_i ist. Der erweiterte Selektionsprozess stellt sich folgendermaßen dar:

1. Sortierung der RITZ–Werte bezüglich ihres Abstandes zum Zielwert τ .
2. Auswertung von (5.4) für alle RITZ–Vektoren.
3. Wenn es RITZ–Paare gibt, die (5.4) erfüllen, werden sie an die Spitze der Sortierung gesetzt.
4. Die JACOBI–DAVIDSON–Iteration wird mit dem ersten RITZ–Paar der Sortierreihenfolge fortgesetzt.

Dementsprechend findet keine Änderung in der Sortierreihenfolge statt, wenn es keine RITZ–Vektoren gibt, die (5.4) erfüllen.

Rayleigh–Quotienten–Iteration. Die RAYLEIGH–Quotienten–Iteration (RQI) nach Algorithmus 4.4 dient der schnellen Berechnung eines einzigen Eigenpaars, allerdings muss das auftretende lineare Gleichungssystem möglichst exakt²⁵ gelöst werden [69]. Das lineare Gleichungssystem wird singular, wenn ein Eigenpaar gefunden ist und die RQI schlägt fehl, falls $\|\mathbf{v}_1\| = 0$ auftritt, so dass nicht normiert werden kann. In Algorithmus 4.5 wird der Startvektor als rechte Seite für die ersten $\iota_{RQI} \geq 1$ Iterationen konstant gehalten, so dass das Konvergenzverhalten in Abhängigkeit von ι_{RQI} untersucht werden kann. Trotz der relativen Einfachheit der RQI besteht eine enge Verwandtschaft zwischen der RQI und dem JACOBI–DAVIDSON–Verfahren, wie aus der Literatur bekannt ist [45]. Die Konvergenzordnung der RQI ist quadratisch und lokal kubisch [69].

²⁴Die Ergebnisse in Tabelle 5.5 zeigen, dass bei dem betrachteten Beispiel und harmonischer Extraktion gegenüber der herkömmlichen Extraktion meist eine Iteration weniger benötigt wird.

²⁵RQI mit inexakten Lösungen des linearen Gleichungssystems werden beispielsweise in [112] gezeigt.

5.4.2.3. Numerische Konfiguration

Für numerische Experimente dient eine gelochte *Microdisk* im Freiraum entsprechend Abbildung 5.17a. Der Brechungsindex beträgt $n_d = 3.3$ und der Radius $\varrho = 1 \mu\text{m}$ bei einer Dicke von $\Delta z = 375 \text{ nm}$. Die kreiszylindrische Aussparung hat einen Radius von 220 nm und ihre Symmetrieachse liegt bei $x = 391 \text{ nm}$. Berechnet werden die Eigenmoden $\text{TM}_{1,12}$ und $\text{TM}_{2,8}$. Die Eigenmoden sind keine Grundmoden und somit sind die entsprechenden Eigenwerte des algebraischen Eigenwertproblems nicht an einem oberen oder unteren Ende des Spektrums lokalisiert. Zur Diskretisierung wird in CST MICROWAVE STUDIO [123] ein Rechengitter mit 14 Rechengitterlinien pro Wellenlänge erzeugt. Zwei Symmetriebedingungen reduzieren die Anzahl der Unbekannten auf ein Viertel der ursprünglichen Problemgröße, wie in Abbildung 5.17c dargestellt ist. Die analytische Approximation liefert stets rotationssymmetrische Lösungen, die für azimutale Ordnung $m > 0$ entartet sind. Um eine eindeutige Lösung zu erhalten, muss die Lage der Nulldurchgänge in azimutaler Richtung ausgerichtet werden. Dies geschieht durch eine Symmetriebedingung $E_z = 0$ in der x - z -Ebene. Dadurch wird auch sichergestellt, dass die numerische Lösung nicht phasenverschoben zur analytischen Approximation berechnet wird. Das Rechengitter wird in Matlab um vier Lagen PML erweitert. Das resultierende algebraische System der Diskretisierung mit 14 Rechengitterlinien pro Wellenlänge hat 270936 komplexe Unbekannte.

Die Systemmatrix wird zunächst umsortiert, wie in Abschnitt 4.4.4 motiviert ist. Anschließend wird das Spektrum mit Algorithmus 4.1 verschoben und normiert. Die Korrekturgleichung im JACOBI-DAVIDSON-Verfahren wird direkt gelöst und die Toleranz für die Norm des Residuums des RITZ-Paares beträgt 10^{-9} . Die maximale Anzahl an JACOBI-DAVIDSON-Iterationen wird auf 50 gesetzt. Die Schätzfrequenz, die aus der analytischen Approximation erhalten wird, beträgt 242.9 THz für den $\text{TM}_{1,12}$ -Eigenmode und 405.1 THz für den $\text{TM}_{2,8}$ -Eigenmode, woraus durch Verschieben und Skalieren jeweils ein Zielwert τ für das JACOBI-DAVIDSON-Verfahren und die RQI gewonnen wird. Die Feldverteilung in der $z = 0$ Ebene der analytischen Approximation aus Abbildung 5.17b wird als Gewichtungsvektor \mathbf{f} für das Selektionskriterium (5.4) verwendet. Darüber hinaus lassen sich aus der Feldverteilung der analytischen Approximation verschiedene Startvektoren \mathbf{v}_0 erzeugen, in dem die Approximation für die $z = 0$ Ebene \mathbf{v}_{ana} auch in Ebenen mit $z > 0$ verwendet wird, zumal sich die Feldverteilung von Ebene zu Ebene nicht abrupt ändern wird.

5.4.2.4. Ergebnisse für den $\text{TM}_{1,12}$ -Eigenmode

Jacobi-Davidson-Verfahren. In Tabelle 5.5 sind die Ergebnisse der JACOBI-DAVIDSON-Löser aus Algorithmus 4.9 und 4.10 mit verschiedenen Startvektoren gezeigt. Der Eintrag in der ersten Spalte steht für die Anzahl der Rechengitterlagen, in denen die analytische Approximation enthalten ist. Beim JDQR-Löser ohne erweitertem Selektionskriterium hängt die Anzahl der benötigten Iterationen davon ab, ob die analytische Approximation im Startvektor enthalten ist oder nicht. Elf bis zwölf zusätzliche Eigenmoden werden berechnet, bevor der $\text{TM}_{1,12}$ -Eigenmode als Lösung auftaucht. Dazu werden 31 bis 40 Iterationen benötigt. Die berechneten Eigenfrequenzen mit den Güten Q sind in Abbildung 5.19a dargestellt: Die gestrichelte Linie markiert die Schätzfrequenz, die aus der analytischen Approximation resultiert. Das Quadrat steht für das Ergebnis des $\text{TM}_{1,12}$ -Eigenmode mit einer Resonanzfrequenz von 253.6 THz und $Q \approx 1.1 \times 10^6$. Zusätzliche Eigenmoden, die vom JDQR ohne erweitertes Selektionskriterium berechnet werden, sind durch Kreise dargestellt. Der Einschub in Abbil-

# \mathbf{v}_{ana} in \mathbf{v}_0	Algorithmus 4.9			Alg. 4.9 und 4.12		Alg. 4.10 und 4.13	
	# JDQR Iterationen	# Zusatz- moden	f / THz	# JDQR Iterationen	f / THz	# JDQZ Iterationen	f / THz
0	40	11	253.6	21	$\neq 253.6$	5	$\neq 253.6$
1	31	11	253.6	6	253.6	4	253.6
3	31	12	253.6	4	253.6	5	253.6
4	31	11	253.6	4	253.6	3	253.6
7	32	11	253.6	5	253.6	4	253.6

Tabelle 5.5.: Ergebnisse für den $\text{TM}_{1,12}$ -Eigenmode erhalten mit dem originalen JDQR-Algorithmus sowie dem JDQR- und dem JDQZ-Algorithmus in Kombination mit dem erweiterten Selektionskriterium (5.4) in Abhängigkeit von verschiedenen Startvektoren \mathbf{v}_0 , die zwischen Null und sieben Lagen der analytischen Approximation beinhalten.

dung 5.19a zeigt den Betrag der E_z Feldverteilung in der $z = 0$ Ebene des $\text{TM}_{1,12}$ -Eigenmodes.

Der JDQR-Löser aus Algorithmus 4.9 mit dem erweiterten Selektionskriterium nach Algorithmus 4.12 (Spalten fünf und sechs in Tabelle 5.5) schlägt erwartungsgemäß fehl, wenn die analytische Approximation nicht im Startvektor \mathbf{v}_0 enthalten ist, da (5.4) nicht erfüllt werden kann und somit keine erweiterte Selektion stattfindet. Bessere Startvektoren enthalten die analytische Approximation in einer oder mehreren z -Ebenen und führen zu Konvergenz innerhalb vier bis sechs Iterationen. Dabei werden keinerlei zusätzliche unerwünschte Eigenmoden berechnet. Als Schwellwert wird $\iota_l = 0.2$ gewählt.

Das annähernd gleiche Konvergenzverhalten lässt sich beim JDQZ-Löser nach Algorithmus 4.10 mit erweitertem Selektionsprozess nach Algorithmus 4.13 beobachten (Spalten sieben und acht in Tabelle 5.5), wobei wieder als Schwellwert $\iota_l = 0.2$ gewählt wird. In den meisten Fällen konvergiert das JDQZ-Verfahren etwas schneller als das JDQR-Verfahren, wie es für innere Eigenwerte erwartet wird [56].

Die Untersuchung des Einflusses des Schwellwerts ι_l in (5.4) wird für den Fall durchgeführt, dass der Startvektor \mathbf{v}_0 eine einzige Lage der analytischen Approximation enthält. Es zeigt sich, dass der JDQR mit erweitertem Selektionsprozess nach sechs Iterationen konvergiert, wenn $\iota_l \in [0.2, 0.95]$ gilt. Das Kriterium (5.4) kann für $\iota_l > 0.97$ nicht mehr erfüllt werden, so dass unerwünschte Eigenpaare berechnet werden.

Eine Besonderheit im JACOBI-DAVIDSON-Verfahren ist, dass die Korrekturgleichung nicht exakt gelöst werden muss, sondern dass näherungsweise Lösungen möglich sind, z.B. mit einem iterativen Löser. In diesem Fall können RITZ-Vektoren auftreten, die (5.4) erfüllen, aber deren RITZ-Werte weit vom Zielwert τ entfernt sind. Um zu vermeiden, dass dadurch zu große Sprünge durch das Spektrum entstehen, ist es ratsam, nicht zu kleine Werte für ι_l zu wählen. Die Kombination von $\iota_l = 0.8$ mit dem BiCGStab-Löser zur Lösung der JD-Korrekturgleichung führt zu Konvergenz am gesuchten Eigenmode nach 80 Iterationen.

Rayleigh-Quotienten-Iteration. Wenn die Kombination aus Anzahl der Unbekannten und der verfügbaren Arbeitsspeicher- und Rechenressourcen die direkte Lösung der auftretenden linearen Gleichungssysteme erlaubt, ist die RQI eine Alternative zum JD – zumindest

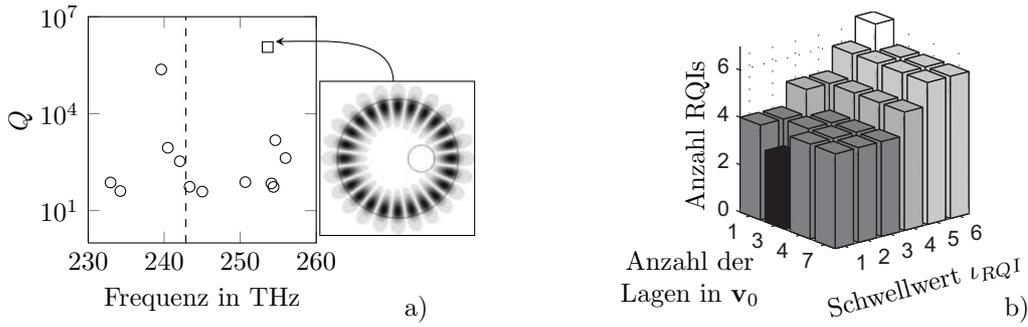


Abbildung 5.19.: a) Eigenfrequenzen und Güte Q : Reelle Schätzfrequenz aus der analytischen Approximation (gestrichelte Linie), gesuchter $TM_{1,12}$ -Eigenmode (\square) und zusätzliche Eigenmoden des diskreten Modells (\circ). b) Anzahl benötigter Iterationen der RQI für verschiedene Startvektoren \mathbf{v}_0 und Schwellwerte ι_{RQI} .

solange ein einziges Eigenpaar gesucht ist. Die Ergebnisse der RQI sind im Diagramm in Abbildung 5.19b dargestellt. Dabei wird die Anzahl der Lagen der analytischen Approximation, die im Startvektor \mathbf{v}_0 enthalten sind, untersucht. Darüber hinaus gibt der Schwellwert ι_{RQI} an, für wie viele Iterationen der Startvektor \mathbf{v}_0 als rechte Seite in der RQI verwendet wird, wobei zumindest die erste Lösung mit \mathbf{v}_0 als rechter Seite erfolgen muss. Die geringste Anzahl an benötigten Iterationen ergibt sich für $\iota_{RQI} = 1, 2, 3$. Allerdings fällt die Norm des Residuums ebenso für die durchgängige Verwendung von \mathbf{v}_0 als rechte Seite ($\iota_{RQI} = 6$) schließlich wie gefordert unter 10^{-9} .

Die Robustheit der RQI kann untersucht werden, indem verschieden gute Startvektoren generiert werden. Dazu werden die Startvektoren \mathbf{v}_0 aus der konvergierten Lösung \mathbf{v}_1 gewonnen, auf die gleichförmig verteilte Zufallswerte addiert werden. Als Maß dient das Skalarprodukt von \mathbf{v}_1 und \mathbf{v}_0

$$\cos(\phi) := \frac{\mathbf{v}_1 \cdot \mathbf{v}_0}{\|\mathbf{v}_1\| \|\mathbf{v}_0\|}.$$

Es zeigt sich, dass für Werte bei $\cos(\phi) = 0.19 + 0.001j$ noch der gesuchte Eigenmode gefunden wird. Dies ist unabhängig von den Parametern ι_{RQI} und der Anzahl der in \mathbf{v}_0 enthaltenen Lagen der analytischen Approximation, wobei die Anzahl benötigter Iterationen zwischen drei und sieben schwankt. Bei Werten unterhalb von $\cos(\phi) = 0.147 + 0.001j$ sind zu wenige Komponenten enthalten, als dass noch Konvergenz zum gesuchten Eigenmode erreicht werden könnte.

5.4.2.5. Ergebnisse für den $TM_{2,8}$ -Eigenmode

Eigenmoden der *Microdisk* mit radial höherer Ordnung haben im Innenbereich insgesamt größere Feldanteile, die nicht zu vernachlässigen sind. Dadurch hat das in der *Microdisk* enthaltene Loch bei $x > 0$ einen Einfluss auf die möglichen Eigenmoden (vgl. Abbildung 5.17a). Eine analytische Approximation der Feldverteilung ist in der Nähe des Lochs nicht möglich, wohl aber bei $x < 0$, dem Halbraum ohne Loch. Exemplarisch soll der $TM_{2,8}$ -Eigenmode der gelochten *Microdisk* aus Abbildung 5.17a berechnet werden. Die Implementierung der analytischen Approximation liefert für diesen Eigenmode eine Schätzfrequenz von 405.1 THz und

# \mathbf{v}_{ana} in \mathbf{v}_0	Algorithmen 4.9 und 4.12		Algorithmen 4.10 und 4.13	
	# JDQR Iterationen	f / THz	# JDQZ Iterationen	f / THz
1	31	$\neq 240.5$	6	$\neq 240.5$
3	7	240.5	49	$\neq 240.5$
4	10	240.5	7	240.5
5	7	240.5	11	240.5
7	10	240.5	11	$\neq 240.5$

Tabelle 5.6.: Ergebnisse für die Berechnung des $\text{TM}_{2,8}$ -Eigenmodes. JDQR und JDQZ mit erweitertem Selektionskriterium (5.4) in Abhängigkeit von verschiedenen Startvektoren \mathbf{v}_0 , die zwischen eins und sieben Lagen der analytischen Approximation nach Abbildung 5.20a enthalten.

die zugehörige Feldverteilung – beides bezogen auf eine homogene *Microdisk* ohne Loch. Aus dieser Feldverteilung lässt sich wie oben beschrieben ein Startvektor für die Eigenwertlöser generieren.

Die RQI schlägt für diese Aufgabe in dem Sinn fehl, dass für alle sinnvoll erscheinenden Parameterkombinationen stets ein anderer als der gesuchte $\text{TM}_{2,8}$ -Eigenmode berechnet wird. Die JACOBI-DAVIDSON-Algorithmen ohne erweiterten Selektionsprozess berechnen ebenfalls unerwünschte Eigenmoden, da die Schätzfrequenz von 405.1 THz stark von der gesuchten Resonanzfrequenz abweicht.

Für die JACOBI-DAVIDSON-Verfahren mit erweitertem Selektionskriterium wird $\nu_l = 0.75$ gesetzt. Als Gewichtungsvektor \mathbf{f} im Selektionskriterium (5.4) wird die analytische Approximation der Halbebene $z = 0, x \leq 0$ verwendet, die in Abbildung 5.20a dargestellt ist. Diese Wahl wird getroffen, da sich das Loch in der *Microdisk* bei $x > 0$ befindet und keine Komponenten der analytischen Approximation aus diesem Bereich für das Selektionskriterium verwendet werden können. Die Ergebnisse, die mit einer Schätzfrequenz von 405.1 THz erhalten werden, sind in Tabelle 5.6 dargestellt. Der JDQR mit erweitertem Selektionskriterium konvergiert gegen die gesuchte Lösung, wenn mehr als eine Ebene der analytischen Lösung im Startvektor enthalten ist. Das verhältnismäßig schlechte Abschneiden des JDQZ-Verfahrens gegenüber dem JDQR-Verfahren lässt sich damit erklären, dass der Zielwert in diesem Verfahren einen größeren Einfluss hat und dieser insbesondere am Anfang der Iterationen relativ weit von der gesuchten Lösung entfernt liegt. Abbildung 5.20b zeigt schließlich den gesuchten $\text{TM}_{2,8}$ -Eigenmode, der aufgrund des Loches in der *Microdisk* eine schwache Abstrahlung in $+x$ -Richtung bei einer Güte von $Q = 871$ aufweist.

5.4.2.6. Vollständiges Eigenwertspektrum bei grober Diskretisierung

Die Berechnung des vollständigen Eigenwertspektrums wird schnell sehr speicherintensiv, weshalb eine grobe Diskretisierung verwendet wird. Abbildung 5.21 zeigt das vollständige Spektrum eines numerischen Modells der gelochten *Microdisk* aus Abbildung 5.17a mit lediglich sieben Linien pro Wellenlänge. Inklusive PML entsteht ein Eigenwertproblem mit 48888 komplexen Freiheitsgraden. Abbildung 5.21a zeigt die Eigenwerte, die aus der QR-Zerlegung resultieren. Das Eigenwertspektrum ist mit Algorithmus 4.1 transformiert worden, um die Kondition zu verbessern, so dass für den betragsgrößten Eigenwert etwa $|\lambda_{\max}| \cong 1$ gilt. In Abbildung 5.21b sind die Eigenwerte als Güten über der Frequenz dargestellt. Der etwa

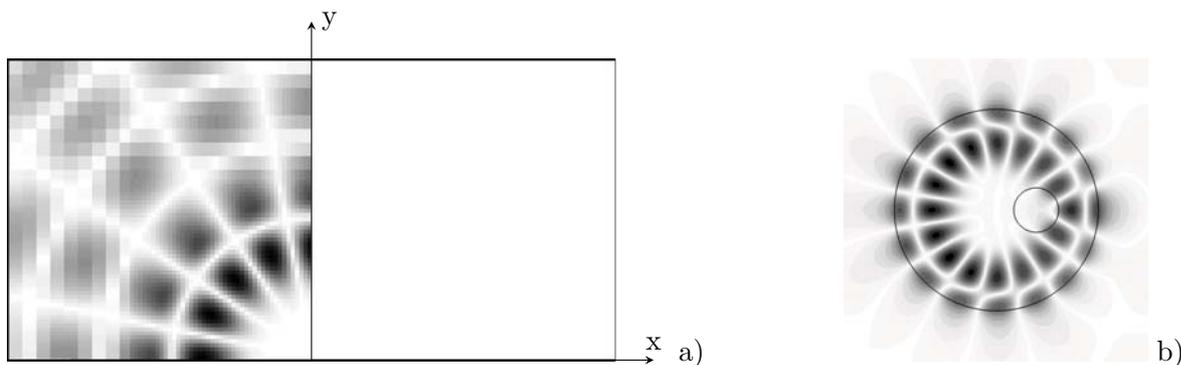


Abbildung 5.20.: a) Analytische Approximation von E_z des $TM_{2,8}$ -Eigenmodes in der $z = 0$ Ebene für $x < 0$. b) $|E_z|$ des numerisch berechneten $TM_{2,8}$ -Eigenmodes in der $z = 0$ Ebene.

n_p -fach²⁶ entartete statische Eigenwert bei Null ist in dieser Darstellung gut zu erkennen [28], und die Abweichungen von Null sind numerisch bedingt. Die Markierung des $TM_{1,12}$ -Eigenmode zeigt, dass dieser Eigenmode tatsächlich zu den Eigenmoden mit den höchsten Güten im unteren Frequenzbereich gehört. Der $TM_{1,12}$ -Eigenmode befindet sich bei betragmäßiger Sortierung an der 175. Stelle neben den Nulleigenwerten²⁷ und gehört somit zu den inneren Eigenwerten.

5.4.2.7. Fazit

In diesem Abschnitt wird gezeigt, wie a priori verfügbare Informationen über gesuchte Eigenvektoren im erweiterten Selektionsprozess innerhalb des JACOBI-DAVIDSON-Verfahrens mit Standardextraktion oder harmonischer Extraktion verwendet werden können. Der erweiterte Selektionsprozess basiert auf einer Umsortierung der RITZ-Paare, die entsprechend der a priori bekannten Eigenschaften gewichtet werden. Die Wahl des Schwellwerts im Selektionskriterium stellt sich als unkritisch für den Erfolg heraus. Dies gilt sowohl für die direkte Lösung der Korrekturgleichung im JD-Verfahren als auch für die näherungsweise Lösung, bei der im Allgemeinen eine größere Anzahl Iterationen notwendig ist, bevor das Konvergenzkriterium erreicht ist. Sofern die direkte Lösung der auftretenden linearen Gleichungssysteme möglich ist, bietet sich die RQI als Alternative zum JD an – zumindest wenn ein einzelnes Eigenpaar gesucht ist. Die Anzahl an benötigten Iterationen ist vergleichbar gering. In beiden Verfahren wird die meiste Berechnungszeit für die Lösung der linearen Gleichungssysteme benötigt. Andere Gewichtungskriterien sind denkbar, die auf sekundären Größen wie z.B. der Energiedichte beruhen. Diese Ergebnisse sind teilweise in [141] veröffentlicht.

5.4.3. Fehlereinfluss durch linearisierte PML

Die in Abschnitt 2.2.4.4 eingeführte transparente Randbedingung (PML) verursacht aufgrund der eingebrachten frequenzabhängigen Parameter, dass das algebraische Eigenwertproblem in der Methode der finiten Integration nach Gleichung (3.39) nichtlinear ist. Als Ausweg bietet sich eine Linearisierung durch die Auswertung der PML-Parameter bei einer Schätzfrequenz

²⁶Hier: 15680 von 48888.

²⁷Die Nulleigenwerte können ggf. mit dem grad-div-Term spektral verschoben werden [28].

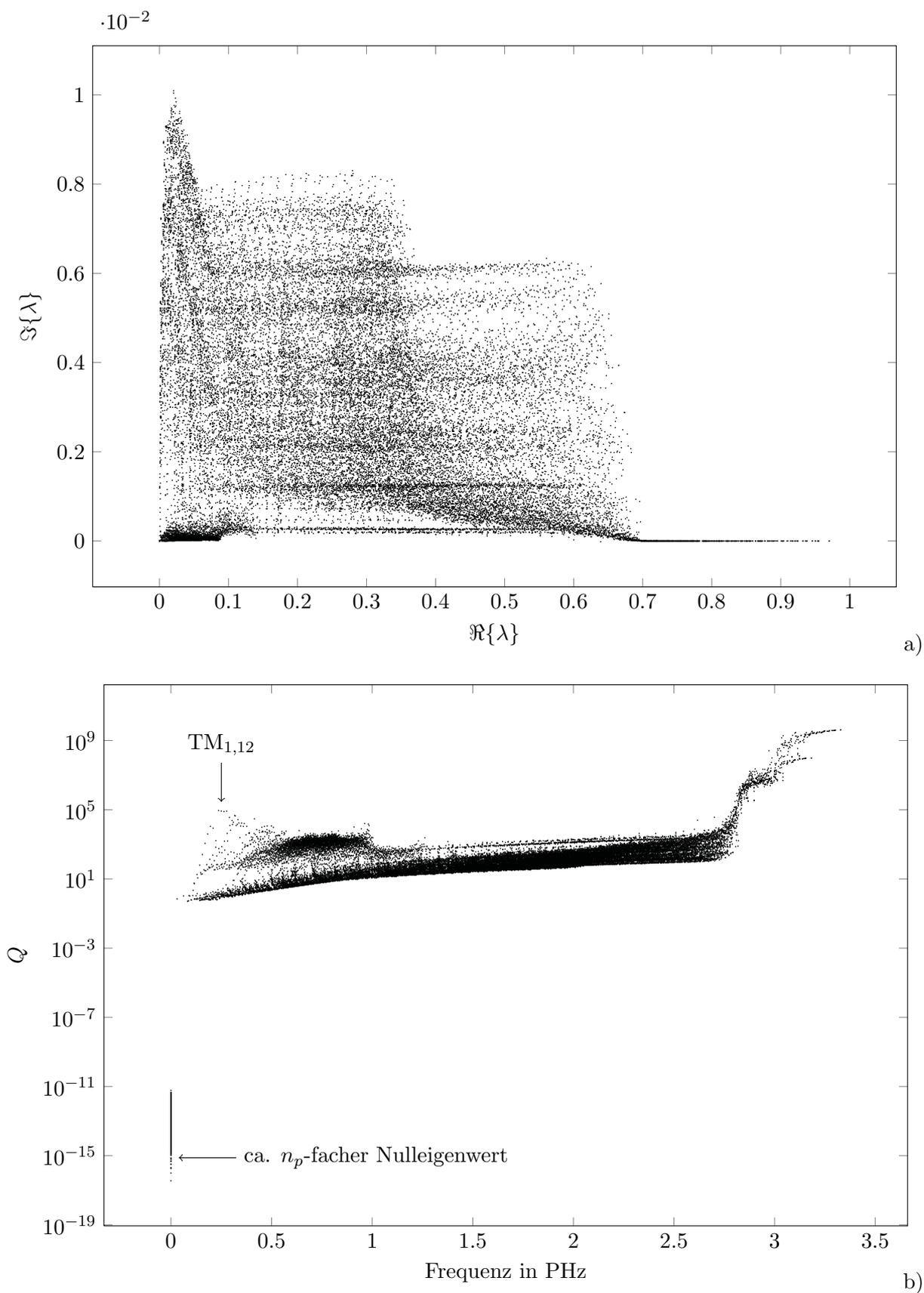


Abbildung 5.21.: Vollständiges Eigenwertspektrum eines numerischen Modells der gelochten *Microdisk* mit 48888 Freiheitsgraden a) und daraus berechnete Güten Q über der Frequenz b).

ω_e an, die in der Nähe der gesuchten Resonanzfrequenz ω liegen muss (vgl. (3.42)). Die Größenordnung des dadurch eingeführten Fehlers soll im Folgenden untersucht werden.

5.4.3.1. Beispiel

Als Beispiel dient die numerische Berechnung des $\text{TM}_{2,8}$ -Eigenmodes der gelochten *Microdisk* aus Abschnitt 5.4.2.5. Der $\text{TM}_{2,8}$ -Eigenmode weist mit $Q = 871$ eine geringere Güte als der betrachtete $\text{TM}_{1,12}$ -Eigenmode mit $Q \approx 1.1 \times 10^6$ auf, so dass mehr Feldanteile abgestrahlt werden und sich der $\text{TM}_{2,8}$ -Eigenmode für die Untersuchung der PML-Randbedingung eignet.

5.4.3.2. Referenzlösung

Zur Erzeugung der Referenzlösung ω_{ref} wird (3.42) mit einer Fixpunkt-Iteration bestehend aus dem JDQR (Algorithmus 4.9) mit erweitertem Selektionskriterium nach Algorithmus 4.12 mehrere Male gelöst, wobei die im i -ten Schritt berechnete Eigenfrequenz ω_i als Auswertefrequenz ω_e im $i + 1$ -ten Schritt verwendet wird (vgl. Algorithmus 4.16). Die erste Lösung erfolgt analog zu Abschnitt 5.4.2.5. Die Fixpunkt-Iteration wird abgebrochen, wenn die relative Abweichung $\frac{|\omega_e - \omega_i|}{|\omega_e|}$ kleiner als 10^{-12} ist. Die Ergebnisse sind in Tabelle 5.7 angegeben: Nach vier Iterationen wird die Toleranzschwelle unterschritten.

Iteration	Auswertefrequenz $\frac{\omega_e}{2\pi}$ / THz	Eigenfrequenz $\frac{\omega_i}{2\pi}$ / THz	$\frac{ \omega_e - \omega_i }{ \omega_e }$
1	405.060 + j0.000	240.484 + j0.138	4.06×10^{-1}
2	240.484 + j0.138	240.483 + j0.136	7.67×10^{-6}
3	240.483 + j0.136	240.483 + j0.136	2.54×10^{-10}
4	240.483 + j0.136	240.483 + j0.136	2.69×10^{-13}

Tabelle 5.7.: Ergebnisse der Fixpunkt-Iterationen zur Berechnung einer Referenz-Lösung des linearisierten Eigenwertproblems (3.42) am Beispiel des $\text{TM}_{2,8}$ -Eigenmodes aus Abschnitt 5.4.2.5.

5.4.3.3. Abweichung zur Referenzlösung

Für die Untersuchung der Abweichung zur Referenzlösung wird von einer reellen Schätzfrequenz ω_e ausgegangen bei der die PML-Parameter jeweils ausgewertet werden. In Abbildung 5.22 ist die resultierende relative Abweichung der berechneten Resonanzfrequenz ω_i zur Referenzlösung ω_{ref} in Abhängigkeit der relativen Abweichung der Schätzfrequenz ω_e zu $\text{Re}\{\omega_{\text{ref}}\}$ dargestellt. Es zeigt sich, dass die Abweichung der berechneten Eigenfrequenz ω_i zur Referenzlösung ω_{ref} für einen großen Schwankungsbereich der reellen Schätzfrequenz ω_e unterhalb von 10^{-5} liegt. Bei einer Abweichung der Schätzfrequenz von $\text{Re}\{\omega_e - \omega_{\text{ref}}\} / \text{Re}\{\omega_{\text{ref}}\} < 0.03$ liegt die Abweichung der berechneten Eigenfrequenz ω_i zur Referenzlösung ω_{ref} unter 10^{-6} . Für grobe Diskretisierungen bei guten Schätzfrequenzen (wie z.B. in Abschnitt 5.4.1) kann der Fehler durch die Linearisierung von (3.42) vernachlässigt werden.

Für die Güte gilt mit Gleichung (2.49) $Q = \text{Re}\{\omega\} / 2 \text{Im}\{\omega\}$. Die Berechnung von sehr hohen Güten ist äquivalent zur Berechnung von Resonanzfrequenzen, deren Imaginärteil klein gegenüber dem Realteil ist. Ist das Ziel, eine möglichst exakte Bestimmung der Güten, werden

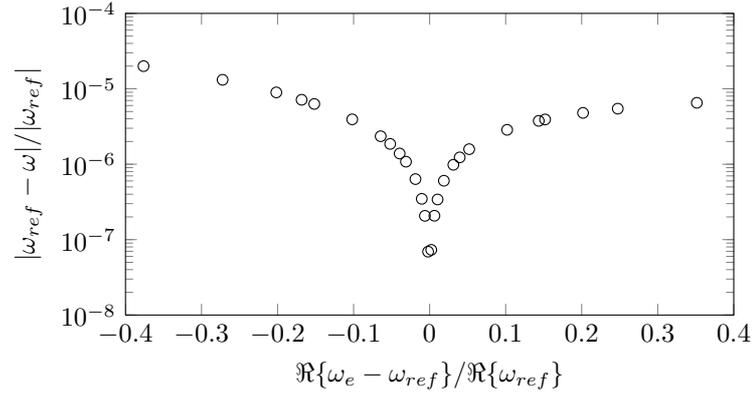


Abbildung 5.22.: Relative Abweichung der berechneten Resonanzfrequenz ω_i zur Referenzlösung ω_{ref} in Abhängigkeit der relativen Abweichung der reellen Schätzfrequenz ω_e zu $\text{Re}\{\omega_{\text{ref}}\}$.

hinreichend viele Dezimalstellen der Eigenwerte benötigt. Dazu muss die Norm des Residuums des Eigenwertproblems hinreichend klein werden. Darüber hinaus wirkt sich bei dieser Problemstellung die Linearisierung der PML stärker aus, so dass in diesem Fall eine Fixpunkt-Iteration ratsam ist.

5.5. Eigenwertberechnung zweidimensionaler Nanoresonatoren mit FLAME

Neben der Methode der finiten Integration wird im Folgenden als zweites Simulationsverfahren die Methode der flexiblen lokalen Approximation aus Abschnitt 2.3 zur Berechnung von Eigenfrequenzen verwendet. Originär ist FLAME für die Berechnung von Randwertproblemen entworfen worden [37], jedoch lassen sich Eigenwertformulierungen ableiten wie in Abschnitt 3.5 gezeigt ist. Die resultierenden FLAME-Eigenwertprobleme (3.47) sind nichtlinear, so dass zu ihrer Lösung Verfahren aus Abschnitt 4.6 zum Einsatz kommen.

5.5.1. Nanoresonator in FLAME–Diskretisierung

Als geometrische Struktur wird eine Schnittebene des Nanoresonators aus [91] verwendet. Der Resonator entsteht, indem im Strukturgitter basierend auf der hexagonalen Elementarzelle nach Abbildung 5.23a drei Elementarzellen vollständig ausgefüllt werden. Abbildung 5.23b zeigt das zweidimensionale Modell des Nanoresonators mit den gleichen Parametern²⁸ wie in Abschnitt 5.4.1. Für ein numerisches Modell in FLAME des Nanoresonators ist die Verwendung von Zylinderfunktionen der Ordnung m aus (3.45) als Basisfunktionen naheliegend, so dass für dielektrische Zylinder mit Radius ϱ_0

$$\begin{aligned} \psi_m^{\text{cyl,int}} &= c_n J_m(k_{\text{cyl,int}} \varrho) j^{-m} e^{jm\phi}, & \varrho &\leq \varrho_0 \\ \psi_m^{\text{cyl,ext}} &= (J_m(k_{\text{cyl,ext}} \varrho) + d_n H_m^{(2)}(k_{\text{cyl,ext}} \varrho)) j^{-m} e^{jm\phi}, & \varrho_0 < \varrho \leq 1.5\varrho_0 \end{aligned} \quad (5.5)$$

²⁸Strukturgitterkonstante $a_{\text{per}} = 0.42 \mu\text{m}$ (Abstand der Lochmittelpunkte), Lochradius $\varrho_0 = 0.29a_{\text{per}}$ und Dielektrikum mit $\varepsilon_r = 11.56$.

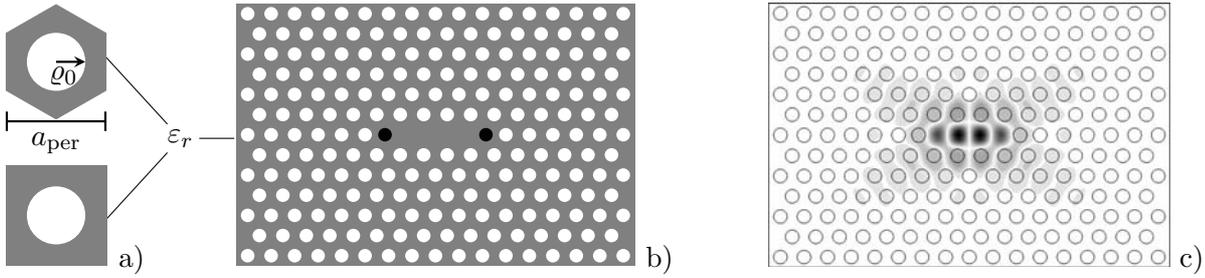


Abbildung 5.23.: a) Hexagonale Elementarzelle (oben) und kartesische Elementarzelle mit gleichem Lochmittelpunktsabstand (unten). b) Nanoresonator aus Dielektrikum mit luftgefüllten Löchern. Die schwarz markierten Löcher sind um $0.15a$ nach außen verschoben. c) Betrag der H_z -Komponente des $H_{z,4,1}$ -Eigenmodes.

gewählt wird. Die Koeffizienten c_n und d_n werden aus den Stetigkeitsbedingungen der MAXWELL'schen Gleichungen (2.11) und (2.12) berechnet. In Gebieten mit $1.5\rho_0 < \rho$ werden als Basisfunktionen ebene Wellen $\psi_m^{pw}(\omega, x, y) = e^{j\frac{\omega\sqrt{\epsilon_r}}{c_0}(\cos(m\frac{\pi}{4})x + \sin(m\frac{\pi}{4})y)}$ mit $m \in \{0, \dots, 7\}$ aus Gleichung (3.44) angesetzt. Je nach Radius ρ_0 , Materialeigenschaft ϵ_r , Wellenzahl und Polarisation lassen sich unterschiedliche Konvergenzordnungen mit FLAME erzielen.

5.5.2. Konvergenzordnung für die Streuung am Zylinder

Die Untersuchung der Konvergenzordnung erfordert eine Referenzlösung. Für die Streuung ebener elektromagnetischer Wellen am einzelnen Zylinder folgt mittels analytischer Multipolentwicklung die Lösung u_{ex} , wie in Abschnitt 2.1.2.2 gezeigt ist. Als Testproblem dient die Lösung des Randwertproblems

$$\mathbf{A}_{\text{FLAME}}(h)\mathbf{u}_h = \mathbf{r}_{ex} \quad (5.6)$$

für die kartesische Elementarzelle aus Abbildung 5.23a. Die Systemmatrix $\mathbf{A}_{\text{FLAME}}$ und die Lösung \mathbf{u}_h hängen von der Rechengitterschrittweite h ab. \mathbf{r}_{ex} enthält die exakte Lösung u_{ex} an den Randwerten. Betrachtet wird der äquidistante 3×3 -Differenzenstern auf einem Teilgebiet $\Omega^{(i)}$. An neun Knoten werden acht Basisfunktionen ausgewertet, so dass eine 9×8 -Knotenmatrix (2.119) entsteht. Die Wahl der Basisfunktionen innerhalb der Elementarzelle ist in Abbildung 5.24a exemplarisch für elf Rechengitterpunkte pro Richtung veranschaulicht:

- Innerhalb des dielektrischen Zylinders gilt $\psi_m^{cyl,int}$ aus (5.5) mit $m \in \{-3, \dots, 4\}$.
- Im Nahbereich des Zylinders gilt $\rho \leq 1.5\rho_0$ wird $\psi_m^{cyl,ext}$ aus (5.5) mit $m \in \{-3, \dots, 4\}$.
- In einiger Entfernung zum Zylinder mit $\rho > 1.5\rho_0$ werden ebene Wellen ψ_m^{pw} mit $m \in \{0, \dots, 7\}$ aus (3.44) als Basisfunktion verwendet.
- Die Randwerte der Elementarzelle sind durch die analytische Lösung u_{ex} gegeben und bilden die rechte Seite in (5.6).

Gleichung (5.6) wird für kleiner werdende h bei fester Frequenz f bzw. Wellenzahl k gelöst. Der globale Approximationsfehler²⁹ $\|\mathbf{u}_h - \mathbf{u}_{ex}\|_2 / \|\mathbf{u}_{ex}\|_2$ ist in Abbildung 5.24b für einen

²⁹Vgl. Abschnitt 2.3.3.1

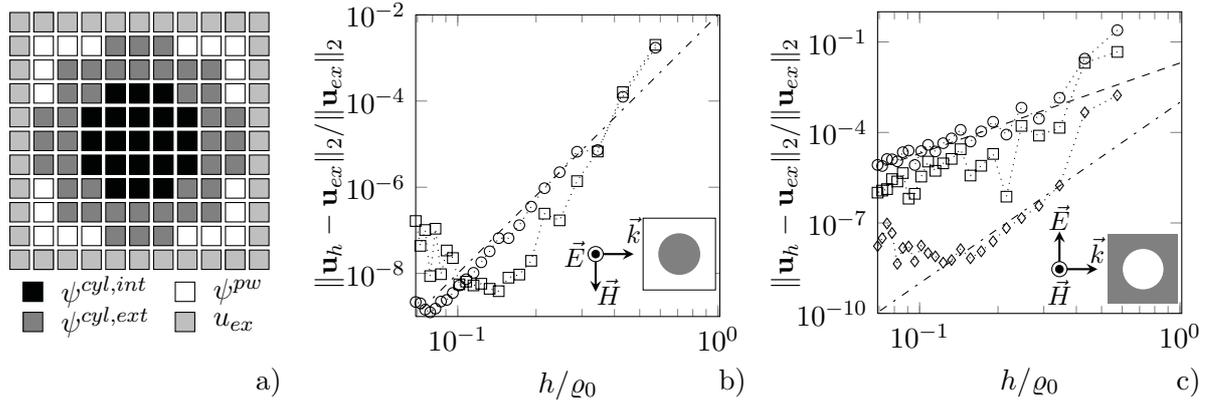


Abbildung 5.24.: a) Zuordnung der Basisfunktionen und der Randwerte im Rechengitter für eine Diskretisierung mit elf Punkten pro Richtung. Approximationsfehler in Abhängigkeit der Rechengitterschrittweite: b) Dielektrischer Zylinder ($\varepsilon_r = 11.56$) im Freiraum bei E_z -Polarisation; c) Loch im homogenen Dielektrikum ($\varepsilon_r = 11.56$) bei H_z -Polarisation, Legende: $k_0\rho_0 = \{0.13(\diamond), 0.40(\square), 1.02(\circ)\}$, h^3 (- -), h^6 (· ·).

dielektrischen Zylinder im Freiraum bei E_z -Polarisation dargestellt. Die aus der Literatur bekannte Konvergenzordnung von sechs [34] kann dabei für $k_0\rho_0 = 1.02(\circ)$ nachvollzogen werden. Das Konvergenzverhalten für $k_0\rho_0 = 0.40(\square)$ ist uneinheitlicher und wird schon bei größeren Werten von h/ρ_0 numerisch instabil.

Deutlichere Unterschiede im Konvergenzverhalten sind für die Elementarzelle zu sehen, bei der die dielektrischen Eigenschaften von Zylinder und Umgebung vertauscht sind, so dass wie in Abbildung 5.23a der Zylinder die Dielektrizitätszahl ε_0 und die Umgebung $\varepsilon_0\varepsilon_r$ besitzt. Abbildung 5.24c zeigt diesen Fall bei H_z -Polarisation. Für $k_0\rho_0 = 0.13(\diamond)$ wird Konvergenz sechster Ordnung erreicht. Für $k_0\rho_0 = 0.40(\square)$ nähert sich die Kurve (bis auf einige Ausnahmen) der Konvergenz dritter Ordnung an. Im Fall von $k_0\rho_0 = 1.02(\circ)$ schließlich resultiert bereits nach den ersten Schritten Konvergenz von lediglich dritter Ordnung.

Für die größten Werte von h/ρ_0 wird der Fehlereinfluss durch die mangelnde lokale Approximation der gewählten Basisfunktionen und die relativ große Fehlanpassung zwischen Gebieten unterschiedlicher Basisfunktionen verursacht. Bei einer Gitterauflösung von $h/\rho_0 \approx 0.34$ ist die Zuordnung der Basisfunktionen durch Abbildung 5.24a gegeben, woraus die anfangs noch grobe Diskretisierung ersichtlich wird. Im Fall von $k_0\rho_0 = 1.02(\circ)$ begrenzen die in den Basisfunktionen enthaltenen Ordnungen der Zylinderfunktionen die Konvergenzordnung.

5.5.3. Lösung des FLAME–Eigenwertproblems

Betrachtet wird die planare Struktur aus Abschnitt 5.4.1 in zwei Dimensionen, die aus den hexagonalen Elementarzellen in Abbildung 5.23a zusammengesetzt ist und in Abbildung 5.23b dargestellt ist. Der $H_{z,4,1}$ -Eigenmode entsprechend Abbildung 5.23c wird bei unterschiedlichen Gitterauflösungen berechnet. Die Ergebnisse für die Konvergenzordnung aus dem vorherigen Abschnitt für die H_z -Polarisation in Abbildung 5.24c gelten somit näherungsweise, da diese sich auf die kartesische Elementarzelle aus Abbildung 5.23a beziehen. Die Zuordnung der Basisfunktionen geschieht analog zu Abbildung 5.24a.

Für die Eigenwertformulierung aus Abschnitt 3.5 geht das Randwertproblem (5.6) in (3.47) aus Abschnitt 3.5 über. Da keine erweiterte Randbehandlung implementiert ist, werden die Randwerte zu Null gesetzt. Ein Eigenpaar (ω_e, \mathbf{x}_e) von $\mathbf{A}_{\text{FLAME}}(\omega_e)\mathbf{x}_e = \mathbf{0}$ (3.47) ist gefunden, wenn $\mathbf{x}_e \in \ker \mathbf{A}_{\text{FLAME}}(\omega_e)$ erfüllt ist. Die Lösung erfolgt mit dem Sekantenverfahren aus Algorithmus 4.17. Die darin auftretenden Standardeigenwertprobleme werden mit dem JDQR aus Algorithmus 4.9 gelöst. Beim Auftreten mehrerer Eigenwerte nahe Null von $\mathbf{A}_{\text{FLAME}}(\omega_e)$ empfiehlt es sich, entweder aus mehreren Eigenvektoren den Gesuchten auszuwählen oder den JDQR mit erweitertem Selektionskriterium nach Algorithmus 4.12 zu verwenden.

5.5.3.1. Ergebnisse

Algorithmus 4.17 erfordert Startwerte, die hinreichend dicht am gesuchten Eigenwert liegen. Aufgrund der geringen Dimension, der aus der FLAME-Formulierung hervorgehenden Systemmatrizen von $\dim \mathbf{A}_{\text{FLAME}} \propto 10^4$, ist der Bereich des gesuchten Eigenwerts recht schnell durch mehrere Läufe bei unterschiedlichen Startwerten auffindbar. Lösungen, die eine ähnliche Feldverteilung wie der gesuchte Eigenmode aufweisen, treten auch auf, wenn der Zielwert τ kein Eigenwert λ des nichtlinearen Eigenwertproblems ist. Um die Konvergenz zu unerwünschten Eigenwerten zu unterbinden, kann innerhalb von Algorithmus 4.17 ab dem zweiten Aufruf des Löser für das Standardeigenwertproblem (hier: JDQR-Löser) die zuletzt berechnete Lösung als Startwert verwendet werden, was das Konvergenzverhalten deutlich verbessert.

Die berechnete Eigenfrequenz der Struktur aus Abbildung 5.23b im Bereich von 155.8 THz entspricht gerade dem Fall $k_0 \varrho_0 = 0.40(\square)$ in Abbildung 5.24c. Dort wird für die kartesische Elementarzelle Konvergenz dritter Ordnung erzielt.

Abbildung 5.25a zeigt die Ergebnisse der mit FIT (\times) und FLAME (\circ) berechneten Resonanzfrequenzen des $H_{z,4,1}$ -Eigenmodes des Nanoresonators aus Abbildung 5.23b. Schon für relativ grobe Diskretisierungen von $h/\varrho_0 > 0.5$ strebt die mit FLAME berechnete Resonanzfrequenz schnell gegen den endgültigen Wert. Die relative Abweichung der letzten beiden berechneten FLAME-Lösungen liegt bei 5.2×10^{-5} . Das Sekantenverfahren in Algorithmus 4.17 benötigt für die Ergebnisse in Abbildung 5.25a fünf bis acht Funktionsauswertungen, was in diesem Fall jeweils der Lösung von fünf bis acht Standardeigenwertproblemen entspricht.

Beim Vergleich der Anzahl der Unbekannten in Abbildung 5.25b ist darauf zu achten, dass die FLAME-Formulierung ausschließlich für die H_z -Polarisation in der Ebene gültig ist, während mit der verwendeten FIT-Formulierung prinzipiell die Berechnung dreikomponentiger Felder möglich ist. Das äquidistante FIT-Rechengitter hat zwei Lagen Rechengitterpunkte in z -Richtung und seine Randwerte erfüllen $H_{\text{tan}} = 0$.

Während die vollständige Struktur aus Abbildung 5.23b in FLAME simuliert wird, erfolgt in FIT eine Reduktion der Anzahl der Unbekannten durch Symmetriebedingungen auf etwa ein Viertel. Abbildung 5.25b zeigt die Anzahl der Unbekannten für das FIT-Rechengitter ohne Symmetrie. Ein wichtiger Unterschied im Speicherbedarf wie in der Berechnung der Eigenwertprobleme ist dadurch gegeben, dass die Systemmatrix des FIT-Eigenwertproblems mit konservativen Randbedingungen aus (2.93) reell ist, während die Systemmatrix des FLAME-Eigenwertproblems komplex ist.

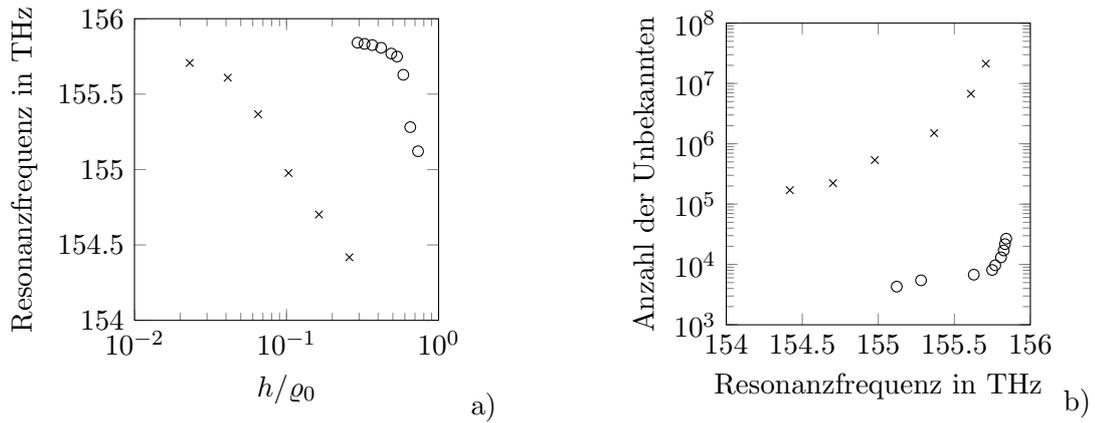


Abbildung 5.25.: a) Mit FIT (\times) und FLAME (\circ) berechnete Resonanzfrequenz des $H_{z4,1}$ -Eigenmodes in Abhängigkeit der Rechengitterauflösung. b) Anzahl der Unbekannten ohne Berücksichtigung von Symmetrien in Abhängigkeit der berechneten Resonanzfrequenz.

5.5.4. Fazit

Dieser Abschnitt beschreibt, wie nichtlineare Eigenwertprobleme, die aus einer FLAME-Diskretisierung resultieren, gelöst werden können. Das Eigenwertproblem wird als Nullstellenproblem formuliert. Das Nullstellenproblem kann mit einem Algorithmus auf Basis des Sekantenverfahrens gelöst werden, wozu Startwerte nah am gesuchten Eigenwert notwendig sind. Die Stärke der FLAME-Formulierung liegt in der höheren Anfangsgenauigkeit für grobe Diskretisierungen. Bei sehr kleinen Rechengitterschrittweiten wird die FLAME-Formulierung numerisch instabil, was nur durch Auswertung der verwendeten Basisfunktionen an mehr als 16 Dezimalstellen vermieden werden kann, wie aus der Literatur bekannt ist [95].

6. Zusammenfassung und Ausblick

Diese Arbeit beschäftigt sich mit der Erweiterung von Berechnungsmethoden für Eigenwertprobleme, die in der numerischen Simulation elektromagnetischer Felder auftreten. Der Schwerpunkt liegt auf Algorithmen für periodische Strukturen und für isolierte Strukturen, die als separate Objekte im freien Raum modelliert werden. Ausgehend von einer Analyse mit bekannten Methoden für Strukturen mit Abmessungen im Zentimeterbereich werden Erweiterungen aufgezeigt, die für die Berechnung von Resonanzeffekten in Nanostrukturen ausgelegt sind. Zur numerischen Modellbildung kommt die Methode der finiten Integration zum Einsatz.

Dispersionseigenschaften von Elementarzellen im Zentimeterbereich. Die Zielsetzung bei der elektromagnetischen Analyse von *Metamaterial*-Elementarzellen mit Abmessungen im Zentimeterbereich ist die Untersuchung der auftretenden BLOCH-Moden im räumlich unendlich ausgedehnten Strukturgitter und die Konsequenzen, die sich aus dem modalen Spektrum für die Beschreibung mit effektiven Materialparametern ergeben. Dazu wird das elektromagnetische Verhalten eines exemplarischen periodischen *Metamaterial*-Strukturgitters bestehend aus verlustfreien Elementarzellen mit zwei verschiedenen Methoden analysiert.

In der ersten Methode wird für das numerische Modell der Elementarzelle das Eigenwertproblem für die unbekanntenen Resonanzfrequenzen aufgestellt, wobei zwei gegenüberliegende Randflächen durch eine feste Phasenverschiebung in Relation gesetzt werden. Das Eigenwertproblem wird für verschiedene Phasenverschiebungen gelöst, woraus sich die BLOCH-Moden als dreidimensionale Feldverteilung in der Elementarzelle sowie die Dispersionskurven ergeben. Die BLOCH-Moden stellen die Referenzlösung dar, da diese Formulierung an den periodischen Randflächen eine beliebige Superposition von Wellenleitermoden auf der Randfläche ermöglicht. Zur Quantifizierung des Einflusses von Wellenleitermoden höherer Ordnung wird die Feldverteilung eines BLOCH-Modens auf der Randfläche des Rechengebiets in Wellenleitermoden entwickelt. Es stellt sich heraus, dass an den periodischen Randflächen kleiner Elementarzellen eine große Anzahl an Wellenleitermoden höherer Ordnung existiert, während die BLOCH-Moden größerer Elementarzellen weniger Anteile von Wellenleitermoden höherer Ordnung besitzen. Die Ursache liegt in den Feldsingularitäten durch metallische Ecken und Kanten begründet, die im Fall der größeren Elementarzelle weiter vom periodischen Rand entfernt sind, so dass sich die Feldverteilung am Rand besser durch eine ebene Welle approximieren lässt als im Fall der kleineren Elementarzelle.

In der zweiten Methode werden die Dispersionseigenschaften einer Elementarzelle aus der Streumatrix berechnet, die durch Frequenzbereichsimulation des numerischen Modells mit der Methode der finiten Integration erhalten werden kann. Die Dispersionskurve lässt sich aus den Eigenpaaren der Transfer-Streumatrix berechnen. Für eine steigende Anzahl an berücksichtigten Wellenleitermoden höherer Ordnung geht die Dispersionskurve aus dem Streumatrix-Ansatz in die Dispersionskurve aus dem periodisch berandeten Eigenwertproblem über, so

dass beide Verfahren im Rahmen der numerischen Genauigkeit die gleichen Ergebnisse liefern.

Die Konsequenz ist, dass die Charakterisierung von Strukturgittern aus kleinen Elementarzellen anhand des Dispersionsdiagramms immer möglich ist. Das Konzept der effektiven Materialparameter basiert typischerweise auf einer Betrachtung, der ebene Wellen zugrunde liegen und kann daher den Freiheitsgrad einer möglicherweise auftretenden Modenkonversion nicht berücksichtigen. Daher schlägt das Konzept der effektiven Materialparameter fehl, wenn eine starke Modenkonversion zu Wellenleitermoden höherer Ordnung innerhalb der Elementarzelle stattfindet.

Dispersionseigenschaften von Elementarzellen im Nanometerbereich. In einem weiteren Schritt wird die Analyse für verlustfreie Elementarzellen im Zentimeterbereich auf verlustbehaftete Strukturen im Nanometerbereich erweitert, da es bei Raumtemperatur keine Metalle gibt, die bei optischen Frequenzen in hinreichender Genauigkeit als ideale Leiter modelliert werden können. Anstelle idealer Leitfähigkeiten sind für Metalle bei optischen Frequenzen Modelle für dispersive Dielektrika zur Materialbeschreibung notwendig. Für den Streumatrix-Ansatz ergibt sich als einziger Unterschied, dass der Löser, mit dem die Streumatrix berechnet wird, passende Dispersionsmodelle beherrschen muss. Bei verlustbehafteten Elementarzellen kann gut eine Zeitbereichssimulation zur Erzeugung der Streumatrix verwendet werden, da die ins Rechengebiet eingespeiste Energie durch die vorhandenen Materialverluste und die teilweise Materialtransparenz schneller wieder abnimmt als im Fall idealer Leiter. Jedoch kann nicht a priori bestimmt werden, wie viele Wellenleitermoden höherer Ordnung in der Streumatrix berücksichtigt werden müssen, um eine hinreichend genaue Dispersionskurve zu erhalten.

Die Formulierung für das periodisch berandete Eigenwertproblem ist für verlustbehaftete Elementarzellen aufwendiger, da ein polynomiales Eigenwertproblem für eine Reihe verschiedener Frequenzen zu lösen ist. Zur Lösung wird ein Eigenwertlöser auf Basis des JACOBI-DAVIDSON-Verfahrens vorgestellt, der für rasche Konvergenz eine gezielte Schätzung des Zielwerts benötigt. In der Eigenwertformulierung sind alle Wellenleitermoden höherer Ordnung inbegriffen. Der Streumatrix-Ansatz mit der Berücksichtigung von hinreichend vielen Wellenleitermoden höherer Ordnung und die Lösung des polynomialen Eigenwertproblems, resultierend aus der periodisch berandeten Elementarzelle, liefern für das Beispiel einer *Fisshnet*-Elementarzelle die Dispersions- und Dämpfungskurven in sehr guter Übereinstimmung. Somit sind zwei Verfahren zur Berechnung der Dispersions- und Dämpfungskurven verlustbehafteter Elementarzellen eingeführt worden.

Berechnung von Wellenleitermoden in photonischen Kristallfasern. Bei der Berechnung von Wellenleitermoden in photonischen Kristallfasern treten unabhängig von den gewählten transversalen Randbedingungen Wellenleitermoden auf, die nicht im Kern der photonischen Kristallfaser geführt werden. Die Zielsetzung ist, ein Verfahren anzugeben, mit dem ausschließlich die im Kern geführten Wellenleitermoden berechnet werden. Dazu wird ein erweiterter Selektionsprozess innerhalb des JACOBI-DAVIDSON-Verfahrens verwendet, in dem a priori bekannte Eigenschaften der gesuchten Eigenpaare als Kriterium dienen. Anstatt die im Verfahren auftretenden Eigenwert-Approximationen ausschließlich bezüglich eines Maximal- oder Zielwerts zu sortieren, werden zusätzlich die Eigenschaften der Eigenvektor-Approximationen analysiert. Am Beispiel der photonischen Kristallfaser zeigt

sich, dass der eingeführte Löser mit erweitertem Selektionsprozess zu drastisch verbessertem Konvergenzverhalten führt, wenn Ausbreitungskonstanten in Wellenleitern berechnet werden, deren Eigenwertspektrum durch den modellierungsbedingten Abschluss mit einer transparenten Randbedingung um eine Vielzahl unerwünschter Wellenleitermoden angereichert ist. Die gezielte Berechnung der gesuchten Wellenleitermoden wird durch die Auswertung der a priori verfügbaren Information erreicht, so dass sichergestellt wird, dass die Leistung der berechneten Wellenleitermoden im Kern geführt wird. Die Wahl des im erweiterten Selektionskriterium vorhandenen Schwellwerts stellt sich als weitgehend unkritisch heraus. Großen Einfluss hingegen hat die Wahl des Startvektors. Die besten Konvergenzeigenschaften werden mit einem Startvektor erreicht, der die zu erwartende Feldverteilung bereits grob widerspiegelt. Da keine unerwünschten Eigenpaare berechnet werden, besitzt der Löser ein problemabhängiges Einsparpotenzial für Rechenzeit- und Speicherressourcen. Während ohne erweitertes Selektionskriterium im konkreten Beispiel eine Vielzahl unerwünschter Wellenleitermoden auftreten, ergeben sich mit dem erweiterten Selektionskriterium keine unerwünschten Wellenleitermoden.

Berechnung von Eigenmoden isolierter Nanostrukturen im Freiraum. Die Problemstellung der Berechnung von Eigenmoden isolierter Nanostrukturen im Freiraum umfasst einerseits das Auffinden von Eigenmoden höherer Ordnung und andererseits die Unterdrückung von Eigenmoden, deren Feldverteilung ausschließlich innerhalb der transparenten Randbedingung lokalisiert ist. Daraus folgt die Zielsetzung, Verfahren zu entwickeln, die mittels a priori bekannter Eigenschaften den gesuchten Eigenmode berechnen können.

Für die Berechnung von Eigenmoden höherer Ordnung in dreidimensionalen Strukturen, sind Zielwerte notwendig, die bereits nah am erwarteten Eigenwert liegen und Startvektoren, die bereits grob die erwartete Feldverteilung widerspiegeln. Am Beispiel eines dielektrischen Nanoresonators mit *photonischer Kristallstruktur* im Freiraum wird gezeigt, wie sich die Berechnung von dreidimensionalen Eigenmoden beschleunigen lässt. Dazu wird eine Zeitbereichssimulation von wenigen Pulsbreiten Dauer durchgeführt. Die Resonanzfrequenzen im abklingenden Zeitsignal werden mit Hilfe der Methode der *harmonischen Inversion* approximiert. Bei der so erhaltenen Resonanzfrequenz wird die dreidimensionale Feldverteilung während einer weiteren Zeitbereichssimulation von wenigen Pulsbreiten Dauer aufgezeichnet. Die dreidimensionale Feldverteilung enthält bereits genügend Anteile vom gesuchten Eigenmode, so dass sie als Startvektor in einer folgenden Lösung des Eigenwertproblems für das gesamte Rechengebiet dienen kann. Es zeigt sich, dass sich der numerische Aufwand, gemessen in der Anzahl der Iterationen des JACOBI-DAVIDSON-Verfahrens, deutlich verringern lässt, wenn die aus der Zeitbereichssimulation gewonnenen Startvektoren und Zielwerte verwendet werden. Am konkreten Beispiel des dielektrischen Nanoresonators wird die Anzahl der benötigten JACOBI-DAVIDSON-Iterationen durch die Wahl des Startvektors mehr als halbiert.

Wenn die Feldverteilung von gesuchten Eigenmoden in dreidimensionalen Strukturen a priori durch analytische Betrachtungen approximiert werden kann, lassen sich diese Informationen im oben eingeführten erweiterten Selektionsprozess innerhalb des JACOBI-DAVIDSON-Verfahrens mit Standardextraktion oder harmonischer Extraktion der RITZ-Werte verwenden. Der erweiterte Selektionsprozess basiert auf einer Umsortierung der RITZ-Paare, die entsprechend der a priori bekannten Eigenschaften gewichtet werden. Als Beispiel dient eine aus einer homogenen dielektrischen Scheibe bestehende *Microdisk* im Freiraum. Für die Approximation von Eigenmoden dieser Anordnung existiert ein Ansatz auf Basis zweier gekoppelter analytischer Eigenwertprobleme. Die erhaltene Lösung ist eine Approximation des Feldes in

der Querschnittsebene senkrecht zur Zylinderachse, da das Feldverhalten an den Kanten der dielektrischen Scheibe nicht exakt beschrieben wird. Das Ergebnis der analytischen Approximation lässt sich jedoch im eingeführten erweiterten Selektionskriterium verwenden, um die Sortierung der RITZ–Paare zu steuern. Das JACOBI–DAVIDSON–Verfahren mit erweitertem Selektionsprozess bevorzugt Eigenmoden, die das Selektionskriterium erfüllen. Hierdurch wird die benötigte Anzahl an Iterationen bis zur Konvergenz deutlich reduziert. Dies gilt sowohl für die direkte Lösung der Korrekturgleichung im JACOBI–DAVIDSON–Verfahren als auch für die näherungsweise Lösung mittels iterativem Gleichungssystemlöser, bei der meist eine größere Anzahl an Iterationen notwendig ist, bevor Konvergenz erreicht ist. Wenn die direkte Lösung des auftretenden linearen Gleichungssystems der Korrekturgleichung möglich ist, bietet sich die RAYLEIGH–Quotienten–Iteration als Alternative zum JACOBI–DAVIDSON–Verfahren an, sofern ein einzelnes Eigenpaar gesucht ist. In allen betrachteten Verfahren wird die meiste Rechenzeit für die Lösung der linearen Gleichungssysteme benötigt. Der Gewichtungsvektor innerhalb des erweiterten Selektionskriteriums für das JACOBI–DAVIDSON–Verfahren muss nicht zwangsläufig eine komplette Schnittebene des numerischen Modells umfassen. Für die gelochte *Microdisk* wird gezeigt, dass sich die Feldverteilung eines Eigenmodes mit radial höherer Ordnung in der lochabgewandten Halbebene mit der analytischen Approximation für die homogene *Microdisk* in erster Näherung beschreiben lässt. Der Zielwert von 405 THz aus der analytischen Approximation für den $\text{TM}_{2,8}$ –Eigenmode liegt weit entfernt vom tatsächlichen Ergebnis von rund 240 THz. Dennoch sind die JACOBI–DAVIDSON–Verfahren mit erweitertem Selektionsprozess in der Lage den $\text{TM}_{2,8}$ –Eigenmode in 7 bis 10 Iterationen zu berechnen.

Eigenwertberechnung zweidimensionaler Nanoresonatoren mit FLAME. Als Alternative zur Methode der finiten Integration wird anhand eines zweidimensionalen Nanoresonators mit *photonischer Kristallstruktur* die Verwendbarkeit der Methode der flexiblen lokalen Approximation (FLAME) untersucht werden. Da für die Beschreibung der Wellenausbreitung in FLAME Exponential- und BESSEL–Funktionen angesetzt werden, kann gezeigt werden, dass die Einträge der FLAME–Systemmatrix bei symbolischer Auswertung des Kerns der Knotenmatrix aus Termen zusammengesetzt sind, die diese Funktionen enthalten. Das Argument der Basisfunktionen beinhaltet die Wellenzahl oder Frequenz, so dass das Eigenwertproblem der Methode der flexiblen lokalen Approximation nichtlinear wird. Die hier betrachtete Lösungsmöglichkeit ist die Formulierung als Nullstellenproblem. Das Nullstellenproblem wird mit einem Algorithmus auf Basis des Sekantenverfahrens gelöst, wozu Startwerte nah am gesuchten Eigenwert notwendig sind. Im Vergleich mit der Methode der finiten Integration liegt der Vorteil der Methode der flexiblen lokalen Approximation in der höheren Anfangsgenauigkeit für grobe Diskretisierungen. Bei sehr kleinen Rechengitterschrittweiten wird diese Formulierung jedoch numerisch instabil. Die Auswertung der verwendeten Basisfunktionen mit mehr als 16 Dezimalstellen kann Instabilitäten vermeiden, jedoch bringt sie erheblichen Aufwand mit sich.

Ausblick

Bei der Berechnung der Dispersionskurven aus der Eigenwertformulierung einer dreidimensionalen Diskretisierung wird das resultierende Eigenwertproblem für verschiedene Parameterwerte der Frequenz bzw. des Phasenvorschubs aufgestellt. Die Einträge in den Systemma-

trizen ändern sich nur für die Randflächen mit Periodizitätsbedingung und die resultierenden Eigenwerte und BLOCH-Moden ändern sich bei hinreichend kleiner Abtastschrittweite nur geringfügig. Für die Folge der zu berechnenden Eigenwertprobleme sollten sich Methoden finden lassen, die den Gesamtrechenaufwand reduzieren – möglicherweise mittels Methoden zur Reduktion der Modellordnung [100].

Ein weiterer interessanter Typ von periodischen Strukturgitter entsteht durch die Betrachtung von Elementarzellen, die modellierungsbedingt an einer oder mehreren Randflächen mit einer transparenten Randbedingung abgeschlossen werden müssen. Dazu gehören beispielsweise Antennen-Arrays, die in eine Richtung als unendlich ausgedehnt betrachtet werden [120]. Die numerische Eigenmode-Berechnung in solchen Strukturen ist anspruchsvoll, da die transparente Randbedingung und die Periodizität verursacht, dass das resultierende Eigenwertproblem komplex und nichtlinear wird.

Für den erweiterten Selektionsprozess in den JACOBI-DAVIDSON-Verfahren lässt sich die Auswertung des Selektionskriteriums möglicherweise direkt im niederdimensionalen auf den Unterraum projizierten Eigenwertproblem durchführen, anstatt alle RITZ-Vektoren in voller Dimension zu berechnen. Andere als die hier verwendeten Gewichtungskriterien sind denkbar, die auf anderen sekundären Größen wie z.B. der Energiedichte beruhen.

In der Methode der flexiblen lokalen Approximation wird das lokal bekannte analytische Verhalten von elektromagnetischen Feldern ausgenutzt, um mit grober Diskretisierung relativ genaue numerische Modelle zu erhalten. Eine Erweiterung auf dreidimensionale Rechengitter und drei Feldkomponenten wird in [110] diskutiert. Dabei zeigt sich, dass die auftretenden Knotenmatrizen auf überbestimmte Gleichungssysteme führen, die durch Minimierung der Summe der Fehlerquadrate gelöst werden. In [146] wird gezeigt, wie analytische Kantenmodelle als lokale Modifikation der Materialmatrizen in der Methode der finiten Integration eingearbeitet werden können. Für weitere einfache Geometrien wie z.B. Kugeln oder Zylinder sollte ebenfalls eine lokale Modifikation der Materialmatrizen möglich sein, um den Einfluss des Treppeneffekts in der Diskretisierung zu minimieren.

A. Anhang

A.1. Berechnung der Dispersionskurve aus der Streumatrix

Mit der Vernachlässigung der Ausschreibung der Frequenzabhängigkeit, der Voraussetzung der Reziprozität aufgrund Verlustlosigkeit der betrachteten Struktur und $x := e^{-j\phi}$ folgt für das direkte Auflösen von Gleichung (3.3) nach [27]

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} a_2 \frac{1}{x} \\ a_1 x \end{pmatrix} \quad (\text{A.1})$$

$$S_{11} = \frac{(\frac{1}{x} - S_{12})(x - S_{21})}{S_{22}} \quad (\text{A.2})$$

$$S_{11}S_{22} = 1 - S_{21}\frac{1}{x} - S_{12}x + S_{12}S_{21} \quad (\text{A.3})$$

$$x^2 + \frac{\det S - 1}{2S_{12}}x + \underbrace{\frac{S_{21}}{S_{12}}}_{=1} = 0 \quad (\text{A.4})$$

mit der Lösung

$$x_{1,2} = -\frac{\det S - 1}{2S_{12}} \pm \sqrt{\left(\frac{\det S - 1}{2S_{12}}\right)^2 - 1}. \quad (\text{A.5})$$

Zu zeigen: *Die Eigenwerte der Transfer-Streumatrix sind das gesuchte Ergebnis.*

Mit der Transfer-Streumatrix nach (2.66) unter der Annahme eines einzigen Modes und nicht ausgeschriebener Frequenzabhängigkeit folgt

$$\begin{pmatrix} b_1 \\ a_1 \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = \begin{pmatrix} \frac{-\det S}{S_{12}} & \frac{S_{22}}{S_{12}} \\ -\frac{S_{11}}{S_{12}} & \frac{1}{S_{12}} \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}. \quad (\text{A.6})$$

Die Eigenwerte $x_{1,2}$ der Transfer-Streumatrix ergeben sich zu

$$\det(T - xI) = 0 \quad (\text{A.7})$$

$$\left(\frac{-\det S}{S_{12}} - x\right) \left(\frac{1}{S_{12}} - x\right) + \frac{S_{11}S_{22}}{S_{12}^2} = 0 \quad (\text{A.8})$$

$$x^2 + \frac{\det S - 1}{2S_{12}}x + \underbrace{\frac{S_{21}}{S_{12}}}_{=1} = 0 \quad (\text{A.9})$$

mit der Lösung wie (A.5)

$$x_{1,2} = -\frac{\det S - 1}{2S_{12}} \pm \sqrt{\left(\frac{\det S - 1}{2S_{12}}\right)^2 - 1}. \quad (\text{A.10})$$

A.2. Matrizendefinitionen

Definition A.2.1 (Givens–Rotationsmatrix): Eine $n \times n$ GIVENS–Rotationsmatrix hat die Form

$$\mathbf{G}_{il}(\phi) = \begin{bmatrix} \mathbf{I}_{i-1} & & & & & \\ & \cos(\phi) & & \sin(\phi) & & \\ & & \mathbf{I}_{l-i-1} & & & \\ & -\sin(\phi) & & \cos(\phi) & & \\ & & & & & \mathbf{I}_{n-l} \end{bmatrix},$$

wobei für die Indices $i < l$ gilt und der Winkel $\phi \in [\pi/2, \pi/2)$ ist.

Der Winkel kann immer so gewählt werden, dass die l -te Komponente von $\mathbf{G}_{il}(\phi)\mathbf{x}$ für einen beliebigen Vektor $\mathbf{x} \in \mathbb{R}^n$ zu Null gesetzt wird. GIVENS–Rotationsmatrizen sind orthogonal und wirken nur auf zwei Zeilen (Spalten) einer Matrix \mathbf{A} , wenn \mathbf{A} von links (rechts) mit $\mathbf{G}(\phi)$ multipliziert wird.

Definition A.2.2 (Hessenberg–Matrix): Eine Matrix $\mathbf{H} \in \mathbb{C}^{n \times n}$ wird dann obere HESSENBERG–Matrix genannt, wenn alle Einträge unterhalb der ersten Nebendiagonalen gleich Null sind

$$h_{il} = 0 \quad \text{für } i > l + 1.$$

A.3. Methoden zur Eigenpaarberechnung von Davidson und Jacobi

Der Darstellung in [86] folgend werden die fundamentalen Ideen der Methoden von DAVIDSON und JACOBI eingeführt.

Methode von Davidson

Die Methode von DAVIDSON [50] ist eine iterative Unterraummethode zur Berechnung von Eigenlösungen und eignet sich insbesondere für diagonaldominante Matrizen \mathbf{A} . Der Suchraum \mathbf{V} soll entsprechend Definition 4.3.1 eine orthonormale Basis des $\mathbb{C}^{n \times k}$ sein und aus den Vektoren $\mathbf{v}_1 \dots \mathbf{v}_k$ bestehen. Die Idee DAVIDSONs zur sukzessiven Erweiterung des Suchraums ist die Folgende: Wenn der gesuchte Eigenwert λ von vornherein bekannt wäre, dann würde für die optimale Erweiterung des Suchraums \mathbf{z}

$$\mathbf{A}(\mathbf{u}_j + \mathbf{z}) = \lambda(\mathbf{u}_j + \mathbf{z})$$

gelten und ließe sich daher aus dem linearen Gleichungssystem

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{z} = \lambda\mathbf{u}_j - \mathbf{A}\mathbf{u}_j$$

berechnen. Das Residuum verschwände im nächsten Iterationsschritt und es resultierte der exakte Eigenwert. Da der exakte Eigenwert λ jedoch nicht a priori bekannt ist, wird als j -te Approximation der RITZ–Wert θ_j verwendet. Um darüber hinaus den Aufwand für die

Lösung des linearen Gleichungssystems gering zu halten, wird \mathbf{A} durch seine Hauptdiagonale \mathbf{D}_A approximiert. Dies führt dann zu einer Erweiterung des Suchraums entsprechend

$$(\mathbf{D}_A - \theta_j \mathbf{I}) \tilde{\mathbf{z}} = -\mathbf{r}_j,$$

mit $\mathbf{r}_j := \mathbf{A}\mathbf{u}_j - \theta_j \mathbf{u}_j$. Der Vektor \mathbf{z} wird gegen $\mathbf{v}_1 \dots \mathbf{v}_k$ orthogonalisiert und schließlich als neue Suchrichtung \mathbf{v}_{k+1} dem Suchraum zugefügt. Dieser Vorgang wird wiederholt bis die Norm des Residuums eine vorgegebene Schranke unterschreitet ($\|\mathbf{z}\| < \epsilon$). Dann wird das Approximationspaar (θ_j, \mathbf{u}_j) als Eigenpaar (λ, \mathbf{x}) akzeptiert. Eine Stagnation der Methode kann auftreten, wenn \mathbf{A} beispielsweise diagonal ist, da sich dann keine neuen Suchrichtungen mehr berechnen lassen. Darüber hinaus ist die Methode nur für hinreichend diagonaldominante Matrizen tauglich.

Die grundlegenden Ideen der Methode von DAVIDSON lassen sich anhand von Algorithmus A.2 nachvollziehen. Zur Orthogonalisierung wird der modifizierte GRAM-SCHMIDT-Algorithmus entsprechend Algorithmus A.1 verwendet.

Algorithmus A.1 Modifizierte Gram-Schmidt-Orthogonalisierung mit Verfeinerung [46].

Eingabe: Basis $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_{m-1}]$, zu orthogonalisierender Vektor \mathbf{t}

Ausgabe: orthogonalisierter Vektor \mathbf{t}

- 1: Wähle $\alpha < 1$, z.B. $\alpha = 0.25$
 - 2: $\tau_{in} = \|\mathbf{t}\|$
 - 3: **for** $i = 1 \dots m - 1$ **do**
 - 4: $\mathbf{t} = \mathbf{t} - (\mathbf{v}_i^H \mathbf{t}) \mathbf{v}_i$
 - 5: **end for**
 - 6: **if** $\|\mathbf{t}\| / \tau_{in} \leq \alpha$ **then**
 - 7: **for** $i = 1 \dots m - 1$ **do**
 - 8: $\mathbf{t} = \mathbf{t} - (\mathbf{v}_i^H \mathbf{t}) \mathbf{v}_i$
 - 9: **end for**
 - 10: **end if**
-

Jacobi's Orthogonalkomplementkorrektur

Bereits 1846 beschreibt JACOBI ein Verfahren [66], das basierend auf Rotationen geeignet ist, die Diagonaldominanz und somit die Lösbarkeit von linearen Gleichungssystemen bestehend aus symmetrischen Matrizen zu verbessern. Hierbei ist \mathbf{A} eine diagonaldominante Matrix und $a_{1,1} = \alpha$ ist das größte Diagonalelement. Somit ist α ebenso Approximation für den betragsgrößten Eigenwert λ und $\mathbf{e}_1 = (1, 0, 0, \dots, 0)^T$ ist eine Approximation für den zugehörigen Eigenvektor \mathbf{x} . Dann kann das Verfahren von JACOBI in Matrixnotation als

$$\mathbf{A} \begin{pmatrix} 1 \\ \mathbf{z} \end{pmatrix} = \begin{pmatrix} \alpha & \mathbf{c}^T \\ \mathbf{b} & \mathbf{M} \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{z} \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ \mathbf{z} \end{pmatrix} \quad (\text{A.11})$$

dargestellt werden. Hierbei sind die Vektoren \mathbf{b} und \mathbf{c}^T entsprechende Ausschnitte der Zeilen und Spalten von \mathbf{A} . Gesucht ist der Eigenwert λ , der sich in der Nähe von α befindet und der zugehörige Eigenvektor $\mathbf{x} = (1, \mathbf{z}^T)^T$ mit der Komponente $(0, \mathbf{z}^T)$ orthogonal zu \mathbf{e}_1 . Das Problem (A.11) ist äquivalent zu

$$\lambda = \alpha + \mathbf{c}^T \mathbf{z} \quad (\text{A.12})$$

Algorithmus A.2 Methode von Davidson ohne Restarts und mit der Möglichkeit, ein Eigenpaar zu berechnen. In diesem Beispiel wird der betragsgrößte Eigenwert berechnet.

Eingabe: Matrix \mathbf{A} , Startvektor \mathbf{v}_0 , Toleranz ϵ

Ausgabe: Eigenpaar (λ, \mathbf{x})

```

1:  $\mathbf{D}_A = \text{diag}(\text{diag}(\mathbf{A}))$ 
2:  $\mathbf{v}_0 = \mathbf{v}_0 / \|\mathbf{v}_0\|$ ,  $\mathbf{v}_A = \mathbf{A}\mathbf{v}_0$ 
3:  $\mathbf{V} = \mathbf{v}_0$ ,  $j = 1$                                      % Initialisiere Suchraum
4:  $\mathbf{M} = \mathbf{v}_0^H \mathbf{v}_A$                                % Initialisiere projizierte Matrix  $\mathbf{M}$ 
5: while 1 do
6:    $[\mathbf{W}, \mathbf{D}_\theta] = \text{eig}(\mathbf{M})$ 
7:    $[\theta, i_\theta] = \text{max}(\text{diag}(\mathbf{D}_\theta))$              % Bestimme größten RITZ-Wert  $\theta$ 
8:    $\mathbf{u} = \mathbf{V}\mathbf{W}(:, i_\theta)$                          % Bestimme RITZ-Vektor
9:    $\mathbf{r} = \mathbf{V}_A \mathbf{W}(:, i_\theta) - \theta \mathbf{u}$        % Bestimme Residuum
10:  if ( $\|\mathbf{r}\| < \epsilon$ ) then                       % Stoppkriterium
11:     $\lambda = \theta$ ,  $\mathbf{x} = \mathbf{u}$ , Stoppe
12:  end if
13:   $\mathbf{z} = -(\mathbf{D}_A - \theta \mathbf{I})^{-1} \mathbf{r}$                  % Berechne Korrektur  $\mathbf{z}$ 
14:   $\mathbf{z} = \text{mgs}(\mathbf{V}, \mathbf{z})$ ,  $\mathbf{z} = \mathbf{z} / \|\mathbf{z}\|$      % Orthonormalisiere  $\mathbf{z}$  gegen  $\mathbf{V}$ 
15:   $\mathbf{V} = [\mathbf{V}, \mathbf{z}]$                                % Füge  $\mathbf{z}$  zum Suchraum hinzu
16:   $\mathbf{v}_A = \mathbf{A}\mathbf{z}$ 
17:   $\mathbf{V}_A = [\mathbf{V}_A, \mathbf{v}_A]$ 
18:   $\mathbf{M} = [\mathbf{M}, \mathbf{V}^H \mathbf{v}_A; \mathbf{v}_A^H \mathbf{V}, \mathbf{v}_A^H \mathbf{z}]$  % Aktualisiere  $\mathbf{M}$ 
19:   $j = j + 1$ 
20: end while

```

$$(\mathbf{F} - \lambda \mathbf{I})\mathbf{z} = -\mathbf{b} \quad (\text{A.13})$$

JACOBI schlägt vor, (A.13) iterativ zu lösen, beispielsweise mit dem GAUSS-JACOBI-Verfahren zur Lösung linearer Gleichungssysteme. λ wird dabei in jedem Iterationsschritt entsprechend (A.12) aktualisiert mit $\mathbf{z}_0 = 0$ als Startvektor. Eine Abwandlung von JACOBI's Orthogonal-komplementkorrektur ist in Algorithmus A.3 angegeben. In der Originalarbeit wird λ nur in jedem zweiten Schritt aktualisiert und anstatt das lineare Gleichungssystem für \mathbf{z} zu lösen, wird lediglich das Inkrement $\Delta \mathbf{z}$ bestimmt [86].

Algorithmus A.3 Orthogonalkomplementkorrektur nach Jacobi

Eingabe: Diagonaldominante Matrix \mathbf{A}

Ausgabe: Eigenpaar (λ, \mathbf{x})

```

1:  $n = \text{dim}(\mathbf{A})$ 
2:  $\mathbf{M} = \mathbf{A}(2:n, 2:n)$ ,  $\mathbf{D} = \text{diag}(\text{diag}(\mathbf{M}))$ 
3:  $\mathbf{b} = \mathbf{A}(2:n, 1)$ 
4:  $\mathbf{c}^T = \mathbf{A}(1, 2:n)$ 
5:  $\alpha = \mathbf{A}(1, 1)$ 
6:  $\mathbf{z} = \text{zeros}(n - 1, 1)$ 
7: while Konvergenz nicht erreicht do
8:    $\tilde{\lambda}_k = \alpha + \mathbf{c}^T \mathbf{z}_k$ 
9:    $(\mathbf{D} - \tilde{\lambda}_k \mathbf{I})\mathbf{z}_{k+1} = (\mathbf{D} - \mathbf{M})\mathbf{z}_k - \mathbf{b}$  % Löse nach  $\mathbf{z}_{k+1}$ 
10: end while
11:  $\mathbf{x} = [1; \mathbf{z}]$ ,  $\lambda = \tilde{\lambda}$ 

```

Gegenüberstellung der beiden Methoden

Suchraum

In der Methode von DAVIDSON wird das RITZ-Paar (λ, \mathbf{u}) als jeweils aktuelle Approximation des Eigenpaares geführt. Dabei liegt \mathbf{u} in dem Unterraum, der durch alle bisher berechneten Erweiterungen des Unterraums aufgespannt wird. In der Methode von JACOBI wird dagegen der aktuelle Eigenvektor durch eine Linearkombination des Startvektors \mathbf{u}_1 , der letzten Approximation \mathbf{u}_k und der Korrektur $[0; \mathbf{z}]$ approximiert.

Orthogonalisierung

In der Methode von JACOBI ist die Orthogonalisierung implizit vorhanden, da alle Korrekturen orthogonal zu $\mathbf{u}_1 = \mathbf{e}_1$ sind. In der Methode von DAVIDSON werden die Korrekturen zur Suchraumerweiterung stets explizit zum kompletten bisherigen Suchraum orthogonalisiert.

Eigenwertapproximation

In der Methode von DAVIDSON approximiert ein RITZ-Wert jeweils den Eigenwert, während in der Methode von JACOBI die Eigenwertapproximation aus (A.12) berechnet wird.

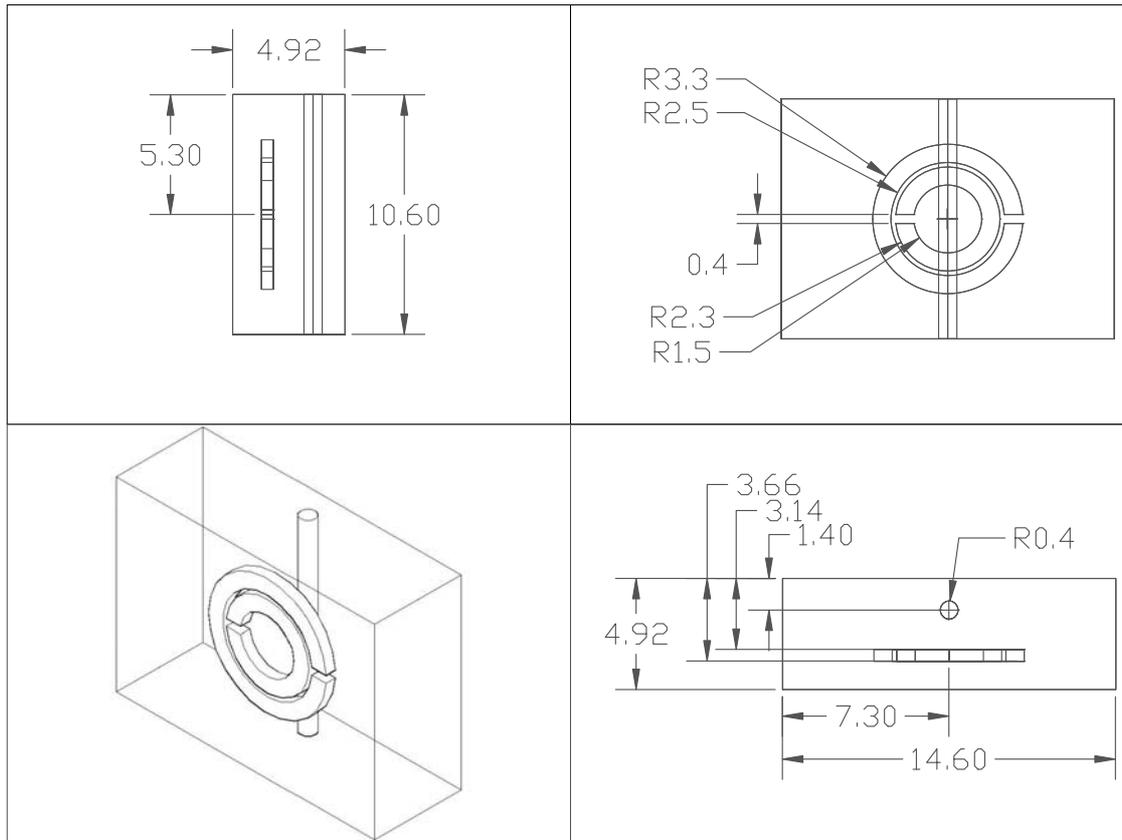


Abbildung A.1.: Elementarzelle aus *Split-Ring-Resonator* [109] und Dipol mit 14.6 mm Ausdehnung in Ausbreitungsrichtung [119].

A.4. Geometrie der Elementarzellen im Zentimeterbereich

Die Geometrie der in Abschnitt 5.1 betrachteten Elementarzellen stammt aus [119]. Die hier untersuchten Abmessungen sind in den Abbildungen A.1 und A.2 angegeben. Die metallischen Bestandteile der Elementarzelle sind in CST MICROWAVE STUDIO als verlustlos simuliert worden.

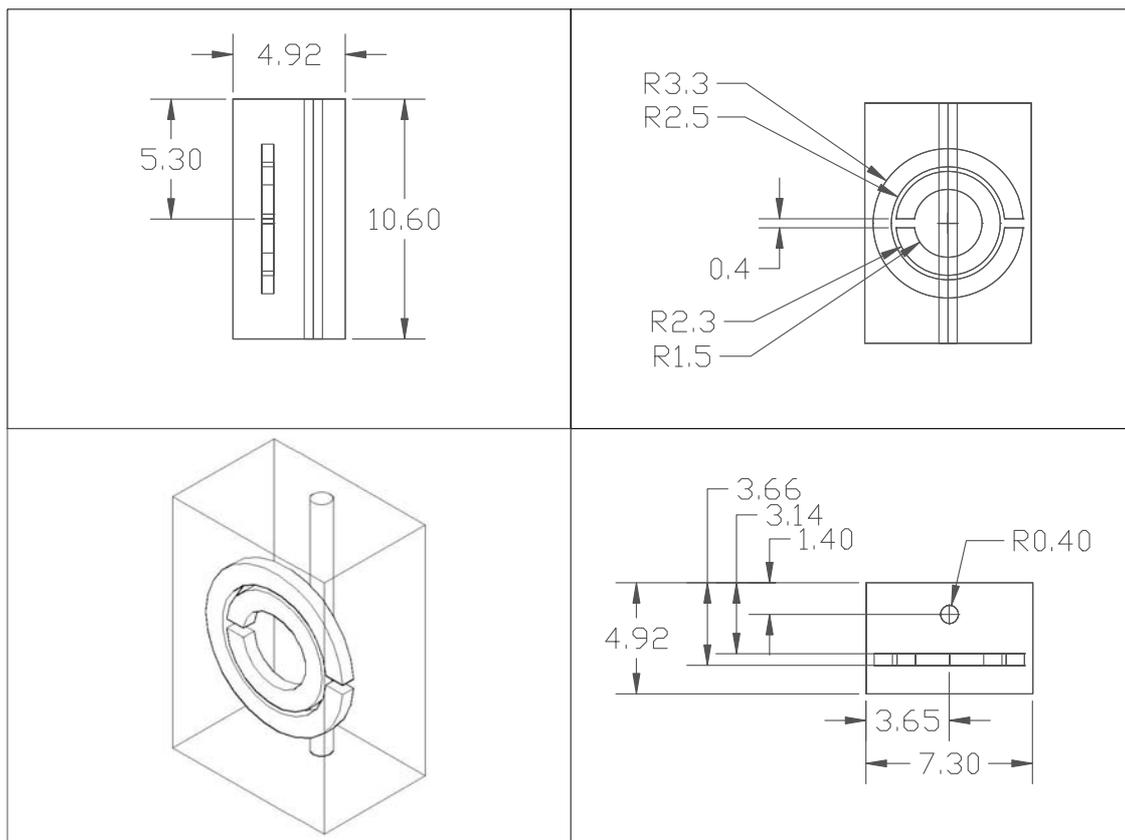


Abbildung A.2.: Elementarzelle aus *Split-Ring-Resonator* [109] und Dipol mit reduzierter Ausdehnung von 7.3 mm in Ausbreitungsrichtung.

Abkürzungs- und Symbolverzeichnis

Abkürzungen

DFT	Diskrete FOURIER-Transformation.
FIT	Methode der finiten Integration [38] (engl. finite integration technique).
FLAME	Methode der flexiblen lokalen Approximation (engl. flexible local approximation methods).
FLOPS	Gleitkommaoperationen pro Sekunde (engl. floating point operations per second).
JD	JACOBI-DAVIDSON.
JDPEP	JACOBI-DAVIDSON-Verfahren für polynomiale Eigenwertprobleme nach Algorithmus 4.14.
JDQR	JACOBI-DAVIDSON-Verfahren mit Standard-Extraktion der RITZ-Werte nach Algorithmus 4.9.
JDQZ	JACOBI-DAVIDSON-Verfahren mit harmonischer Extraktion der RITZ-Werte nach Algorithmus 4.10.
LHM	Linkshändiges <i>Metamaterial</i> .
PBE	Periodisch berandetes Eigenwertproblem.
PCF	Photonische Kristallfaser (engl. photonic crystal fibre).
PEC	Perfekt elektrisch leitend (engl. perfect electric conducting).
PEP	Polynomiales Eigenwertproblem.
PMC	Perfekt magnetisch leitend (engl. perfect magnetic conducting).
PML	Perfekt (impedanz)angepasste Lagen (engl. perfectly matched layers).
RQI	RAYLEIGH-Quotienten-Iteration.
SMA	Streumatrix-Ansatz.
TEM	Transversal elektromagnetisch.

Allgemeine Konventionen

- Vektoren im \mathbb{R}^3 werden durch einen Pfeil gekennzeichnet, z.B. \vec{E} (2.1).
- Vektoren im \mathbb{C}^n sind fett-gesetzte Kleinbuchstaben.
- Matrizen im $\mathbb{C}^{n \times n}$ sind fett-gesetzte Grossbuchstaben.
- Die Zeitabhängigkeit komplexer Phasoren ist $e^{j\omega t}$ entsprechend (2.15).
- Komplexe Größen sind nicht besonders gekennzeichnet, sofern keine Verwechslung mit ähnlichen reellen Größen befürchtet werden muss.
- In Algorithmen ist die Notation an MATLAB [126] orientiert — insbesondere die Adressierung von Matrixelementen.

Allgemeine Symbole

$ \cdot $	Betrag
$\ \cdot\ $	Norm
$\ \cdot\ _2$	Spektralnorm
δ_{il}	KRONECKER Delta
e	EULER'sche Zahl, $2.718\dots \lim_{n \rightarrow \infty} (1 + \frac{1}{n})^n$, $n \in \mathbb{N}$
\mathbb{C}	Menge der komplexen Zahlen
j	Imaginäre Einheit $j = \sqrt{-1}$
$\mathcal{O}()$	LANDAU ¹ -Symbol
\mathbb{N}	Menge der natürlichen Zahlen
π	Kreiszahl, $3.1415\dots$
\mathbb{R}	Menge der reellen Zahlen
$\bar{\mathbf{x}}$	Tensorielle Größe
$z = x + jy$	Komplexe Zahl
$\text{Re}\{z\}, \text{Im}\{z\}$	Real- und Imaginärteil der komplexen Zahl z
\mathbb{Z}	Menge der ganzen Zahlen

Elektromagnetische Wellenausbreitung

a_n	Einfallende Wellenamplitude am Tor n	Seite 18
\mathbf{a}_n	Vektor der einfallenden Wellenamplituden an Tor n	Seite 20
a_{per}	Strukturgitterkonstante	Seite 12
A, \vec{A}	Flächeninhalt, Flächenelement	Seite 7
b_n	Auslaufende Wellenamplitude am Tor n	Seite 19
\mathbf{b}_n	Vektor der auslaufenden Wellenamplituden an Tor n	Seite 20
\vec{B}	Magnetische Flussdichte	Seite 7
c_0	Lichtgeschwindigkeit im Vakuum	Seite 9
c_n	Entwicklungskoeffizient	Seite 10
d_n	Entwicklungskoeffizient	Seite 10
\vec{D}	Verschiebungsstromdichte	Seite 7
\vec{E}	Elektrische Feldstärke	Seite 7
$\vec{\mathcal{E}}_n(x,y)$	Transversale Wellenleitermodenfunktion für \vec{E}	Seite 18
f	Frequenz	Seite 10
$H_n^{(1,2)}$	HANKEL-Funktion erster oder zweiter Art und Ordnung n	Seite 10
\vec{H}	Magnetische Feldstärke	Seite 7
$\vec{\mathcal{H}}_n(x,y)$	Transversale Wellenleitermodenfunktion für \vec{H}	Seite 18
\mathcal{I}	Komplexe Amplitude des modalen Stromes	Seite 19
J_n	BESSEL-Funktion der Ordnung n	Seite 10
\vec{J}	Stromdichte	Seite 7
\vec{J}_F	Flächenstromdichte	Seite 8
\vec{K}	Translationsvektor im reziproken Gitter	Seite 13
\vec{K}_B	BLOCH-Wellenvektor	Seite 12
k	Wellenzahl	Seite 9
m_e	Masse des Elektrons	Seite 15
\vec{M}	Magnetisierung	Seite 8

¹Edmund Georg Hermann Landau (1877-1938), deutscher Mathematiker.

N	Dichte pro Volumeneinheit	Seite 14
\vec{p}	Dipolmoment	Seite 14
\vec{P}	Polarisation	Seite 8
\bar{p}_v	Zeitlich gemittelte elektrische Verlustleistungsdichte	Seite 16
q	Ladung	Seite 14
q_e	Ladung des Elektrons	Seite 15
Q	Gütefaktor	Seite 17
\vec{r}	Ortsvektor	Seite 7
\vec{R}	Translationsvektor im Strukturgitter	Seite 12
s_{nm}	Streuparameter	Seite 19
\mathcal{S}	Streumatrix multimodal	Seite 20
\vec{S}	POYNTING–Vektor	Seite 16
\vec{s}	Wegelement	Seite 7
t	Zeitpunkt	Seite 7
\mathcal{T}	Transfer–Streumatrix multimodal	Seite 20
\mathcal{U}	Komplexe Amplitude der modalen Spannung	Seite 19
\vec{v}	Geschwindigkeitsvektor	Seite 8
v_g	Gruppengeschwindigkeit	Seite 11
v_p	Phasengeschwindigkeit	Seite 11
V	Volumen	Seite 7
w_{eff}	Effektive Energiedichte	Seite 17
\bar{w}_e	Zeitlich gemittelte elektrische Energiedichte	Seite 16
\bar{w}_m	Zeitlich gemittelte magnetische Energiedichte	Seite 16
Z_L	Leitungsimpedanz	Seite 19
α_t	Dämpfungsfaktor oder Abklingkonstante (zeitlich)	Seite 15
α_z	Dämpfungsfaktor oder Abklingkonstante (räumlich)	Seite 18
α_{PML}	PML–Parameter	Seite 27
β_z	Ausbreitungskonstante eines Wellenleitermodes	Seite 18
Δ	LAPLACE–Operator	Seite 9
ε_0	Permittivität des Vakuums	Seite 8
$\bar{\varepsilon}$	Tensorielle Permittivität	Seite 8
ε_c	Komplexe Permittivität	Seite 14
ε_∞	Permittivität im Grenzfall hoher Frequenzen	Seite 14
ε'	Realteil der komplexen Permittivität	Seite 14
ε''	Imaginärteil der komplexen Permittivität	Seite 14
ε_r	Relative Permittivität	Seite 10
λ_{el}	Wellenlänge	Seite 10
μ_0	Permeabilität des Vakuums	Seite 8
$\bar{\mu}$	Tensorielle Permittivität	Seite 8
∇	Nabla–Operator	Seite 9
ω	Kreisfrequenz $\omega = 2\pi f$	Seite 9
ω_P	Plasmafrequenz	Seite 15
φ	Phasenverschiebung	Seite 9
ϕ	Azimutale Koordinate im zylindrischen Koordinatensystem	Seite 10
ρ_L	Raumladungsdichte	Seite 7
ϱ	Radiale Koordinate im zylindrischen Koordinatensystem	Seite 10
σ_F	Flächenladungsdichte	Seite 8

κ	Elektrische Konduktivität	Seite 16
$\bar{\kappa}$	Tensorielle Konduktivität	Seite 8

Methode der finiten Integration

A, \tilde{A}	Primäre, duale Gitterfläche	Seite 21
$\widehat{\mathbf{b}}, \widehat{\mathbf{b}}$	Element, Vektor der magnetischen Gitterflusses	Seite 22
$\mathbf{C}, \tilde{\mathbf{C}}$	Diskreter Rotationsoperator auf dem primären, dualen Gitter	Seite 22
$\widehat{\mathbf{d}}, \widehat{\mathbf{d}}$	Element, Vektor des elektrischen Gitterflusses	Seite 22
$\widehat{\mathbf{e}}, \widehat{\mathbf{e}}$	Element, Vektor der elektrischen Gitterspannung	Seite 22
G, \tilde{G}	Primäres, duales Gitter	Seite 21
$\bar{\mathbf{G}}$	EUKLID'scher Metriktensor	Seite 28
$\widehat{\mathbf{h}}, \widehat{\mathbf{h}}$	Element, Vektor der magnetischen Gitterspannung	Seite 22
I	Indexmenge	Seite 20
$\widehat{\mathbf{j}}, \widehat{\mathbf{j}}$	Element, Vektor des elektrischen Gitterstromes	Seite 22
L, \tilde{L}	Primäre, duale Gitterkante	Seite 21
\mathbf{L}_φ	Diskreter Operator für periodische Randbedingung	Seite 25
\mathbf{M}_ε	Materialmatrix der elektrischen Permittivität	Seite 23
\mathbf{M}_κ	Materialmatrix der elektrischen Konduktivität	Seite 23
\mathbf{M}_μ	Materialmatrix der magnetischen Permeabilität	Seite 23
n_p	Anzahl der Gitterpunkte des Rechengitters	Seite 21
$\mathbf{P}_x, \mathbf{P}_y, \mathbf{P}_z$	Diskrete partielle Differenzierungsoperatoren	Seite 23
q, \mathbf{q}	Element, Vektor der elektrischen Gitterladung	Seite 22
$\mathbf{S}, \tilde{\mathbf{S}}$	Diskreter Divergenzoperator auf dem primären, dualen Gitter	Seite 22
s_ζ	Streckungsfaktor, PML	Seite 27
$\bar{\mathbf{S}}$	PML-Metriktensor	Seite 28
V, \tilde{V}	Primäres, duales Gitterzellvolumen	Seite 22
$\bar{\Gamma}$	Abbildung des Ortsvektors von reell nach komplex, PML	Seite 28
$\bar{\varepsilon}$	Gemittelte elektrische Permittivität	Seite 23
$\bar{\mu}$	Gemittelte magnetische Permeabilität	Seite 23
$\bar{\kappa}$	Gemittelte elektrische Konduktivität	Seite 23
Ω	Gebiet, Rechengebiet	Seite 20
σ_ζ	Dämpfungsprofil, PML	Seite 27

Methode der flexiblen lokalen Approximation

$\mathbf{A}_{\text{FLAME}}$	FLAME-Systemmatrix	Seite 33
$\mathbf{c}^{(i)}$	Koeffizientenvektor auf Teilgebiet (i)	Seite 31
h	Äquidistante Rechengitterschrittweite	Seite 30
$\mathbf{N}^{(i)}$	Knotenmatrix auf Teilgebiet (i)	Seite 31
$\mathbf{s}^{(i)}$	Kern der Knotenmatrix auf Teilgebiet (i) nach (2.122)	Seite 31
u	Lösung des kontinuierlichen Randwertproblems	Seite 30
$\mathbf{u}^{(i)}$	Lokale Lösung auf Teilgebiet (i)	Seite 31
$\epsilon_a^{(i)}$	Approximationsfehler auf Teilgebiet (i)	Seite 33
$\epsilon_k^{(i)}$	Konsistenzfehler auf Teilgebiet (i)	Seite 33
$\Omega^{(i)}$	Teilgebiet des Rechengebiets	Seite 30
$\{\psi_l^{(i)}\}$	Basisfunktionen auf Teilgebiet (i)	Seite 30

Formulierungen

$\mathbf{L}_\varphi^{(g)H}$	Periodischer Randoperator	Seite 42
$P_n(\mathbf{v})$	Leistungsfluss am Tor n	Seite 39
\mathbf{v}	Eigenvektor des Eigenzustands der Transfer–Streumatrix	Seite 39
α_0, α_1	Parameter im dispersiven Materialmodell (3.10)	Seite 40
β_0, β_1	Parameter im dispersiven Materialmodell (3.10)	Seite 40
γ	Eigenwert des Eigenzustands der Transfer–Streumatrix	Seite 39
Δz	Strecke in z -Richtung	Seite 37
ψ_m^{cyl}	Zylindrische Basisfunktionen	Seite 46
ψ_m^{pw}	Ebene Wellen als Basisfunktionen	Seite 46
$\Psi(\omega)$	Matrixpolynom in ω	Seite 41
σ_S	Spektrale Verschiebung	Seite 40

Numerische Methoden

$\bar{\mathbf{C}}$	entsprechend Definition 4.1.1	Seite 51
ℓ	Polynomgrad	Seite 52
$\mathbf{0}$	Nullvektor	Seite 51
\mathbf{f}	Gewichtungsvektor	Seite 80
\mathbf{I}	Identitäts- oder Einheitsmatrix	Seite 52
$L(\lambda)$	Linearisiertes Matrixpolynom	Seite 53
l_{\max}	Anzahl gesuchter Eigenpaare	Seite 71
\mathbf{M}	In Unterraum projizierte Systemmatrix	Seite 68
m_{\max}	Minimale Unterraumdimension	Seite 70
m_{\min}	Minimale Unterraumdimension	Seite 69
n	Matrixdimension	Seite 51
$p_\Psi(\lambda)$	Charakteristisches Polynom in λ zum Matrixpolynom Ψ	Seite 54
\mathbf{p}	Ausgewertete Ableitung des Matrixpolynoms $\Psi(\lambda)$	Seite 75
\mathbf{Q}	Orthogonale Projektionsmatrix	Seite 63
\mathbf{r}	Residuum	Seite 56
\mathbf{s}	Rechter Eigenvektor eines niederdimensionalen Eigenwertproblems	Seite 66
\mathbf{t}	Korrekturvektor	Seite 67
\mathbf{R}, \mathbf{T}	Obere Dreiecksmatrix	Seite 63
\mathbf{u}	RITZ–Vektor, Approximation von \mathbf{x}	Seite 66
\mathbf{v}_0	Startvektor	Seite 60
\mathbf{V}_m	Unterraum der $\in \mathbb{C}^{n \times m}$	Seite 66
\mathbf{W}_m	Unterraum der $\in \mathbb{C}^{n \times m}$	Seite 72
\mathbf{x}	Rechter Eigenvektor	Seite 52
\mathbf{y}	Linker Eigenvektor	Seite 52
α_S, α_{opt}	Faktor zur spektralen Skalierung von PEP–Linearisierungen	Seite 58
$\Delta\Psi$	Störung von Ψ	Seite 57
η	Fehlergröße	Seite 57
ϵ_O	Skalierungsfaktor	Seite 75

ι	Schwellwert für das erweiterte Selektionskriterium im JACO-BI-DAVIDSON-Verfahren	Seite 80
\varkappa	Konditionszahl	Seite 57
λ	Eigenwert	Seite 24
$\Psi(\lambda)$	Matrixpolynom in λ	Seite 52
$\rho_{\mathbf{A}}(\mathbf{v})$	RAYLEIGH-Quotient	Seite 55
$\rho(\mathbf{A})$	Spektralradius von \mathbf{A}	Seite 55
$\sigma(\mathbf{A})$	Spektrum von \mathbf{A}	Seite 54
σ_S	Spektrale Verschiebung	Seite 61
τ	Zielwert	Seite 68
θ	RITZ-Wert	Seite 66
ν	Transformierter Eigenwert	Seite 54
$\Xi(\lambda)$	Nichtlineare Matrixfunktion in λ	Seite 51

Ergebnisse und Anwendungen

c_i	Entwicklungskoeffizient	Seite 97
E	Energie	Seite 119
n_0	Brechungsindex des Freiraums	Seite 122
n_d	Brechungsindex eines Dielektrikums	Seite 122
$n_{d,\text{eff}}$	effektiver Brechungsindex	Seite 122
\mathbf{v}_{ana}	Analytische Approximation aus Abschnitt 5.4.2.1, die in einer oder mehreren z -Ebenen verwendet wird.	Seite 124
δ_e	Eindringtiefe	Seite 101
ν_1, ν_2	Koeffizient	Seite 121
$\Psi_{\mathbf{A}}$	Matrixpolynom mit Koeffizientenmatrizen hoher Dimension	Seite 103
$\Psi_{\mathbf{M}}$	Matrixpolynom mit Koeffizientenmatrizen kleiner Dimension	Seite 103
ϱ_0	Radius	Seite 122

Literaturverzeichnis

Analytische Modellierung elektromagnetischer Wellenausbreitung

- [1] COLLIN, R. E. ; DORATO, P. (Hrsg.): *Field Theory of Guided Waves*. Wiley-IEEE Press, 1990
- [2] HARRINGTON, R. F.: *Time-Harmonic Electromagnetic Fields*. New York : McGraw-Hill Book Company, 1961
- [3] JACKSON, J. D. ; WITTE, C. (Hrsg.): *Klassische Elektrodynamik*. 4th. Walter de Gruyter, 2006
- [4] JOANNOPOULOS, J. : *Photonic crystals: molding the flow of light*. Princeton University Press, 2008. – ISBN 9780691124568
- [5] LAMB, H. : On group-velocity. In: *Proceedings London Mathematical Society* s2-1 (1904), Nr. 1, S. 473–479. <http://dx.doi.org/10.1112/plms/s2-1.1.473>. – DOI 10.1112/plms/s2-1.1.473
- [6] LANDAU, L. D. ; LIFSHITZ, E. ; PITAEVSKII, L. P.: *Course of Theoretical Physics*. Bd. 8: *Electrodynamics of Continuous Media*. 2nd. Butterworth-Heinemann, 1984
- [7] MAXWELL, J. C.: A Dynamical Theory of the Electromagnetic Field. In: *Philosophical Transactions of the Royal Society of London* 155 (1865), 459-512. <http://dx.doi.org/10.1098/rstl.1865.0008>. – DOI 10.1098/rstl.1865.0008
- [8] MAXWELL, J. C.: *A Treatise on Electricity and Magnetism*. Clarendon Press, 1873
- [9] POCKLINGTON, H. C.: Growth of a wave group when the group velocity is negative. In: *Nature* 71 (1905), S. 607–608
- [10] SIHVOLA, A. : *Electromagnetic mixing formulas and applications*. IEE Publishing, 1999. – Errata <http://users.tkk.fi/asihvola/iee-painovirheet.html>
- [11] SLATER, J. C.: Microwave Electronics. In: *Rev. Mod. Phys.* 18 (1946), Oct, Nr. 4, S. 441–512. <http://dx.doi.org/10.1103/RevModPhys.18.441>. – DOI 10.1103/RevModPhys.18.441
- [12] SOLYMAR, L. (Hrsg.) ; SHAMONINA, E. (Hrsg.): *Waves in Metamaterials*. Oxford University Press, 2009
- [13] THIEDE, A. : *Vorlesungsskript Werkstoffe der Elektrotechnik*. Warburger Str. 100 33098 Paderborn, 2010

Diskrete Modellierung elektromagnetischer Wellenausbreitung

- [14] BÉRENGER, J.-P. : A perfectly matched layer for the absorption of electromagnetic waves. In: *Journal of Computational Physics* 114 (1994), Nr. 2, 185-200. <http://dx.doi.org/10.1006/jcph.1994.1159>. – DOI 10.1006/jcph.1994.1159. – ISSN 0021–9991
- [15] ČAJKO, F. : *Nano-Focusing of Light: Electromagnetic Analysis and Simulation*, The University of Akron, Diss., 2009
- [16] CLEMENS, M. : *Zur Berechnung zeitlich langsam veränderlicher elektromagnetischer Felder mit der Finite-Integrations-Methode*, TU Darmstadt, Diss., 1998
- [17] DAI, J. : *Simulation of Multiobject Nanoscale Systems*, The University of Akron, Diss., 2009
- [18] DOHLUS, M. : *Ein Beitrag zur numerischen Berechnung elektromagnetischer Felder im Zeitbereich*, TH Darmstadt, Diss., 1992
- [19] HALVERSON, D. ; FRIEDMAN, G. ; TSUKERMAN, I. : Local approximation matching for open boundary problems. In: *Magnetics, IEEE Transactions on* 40 (2004), July, Nr. 4, S. 2152–2154. <http://dx.doi.org/10.1109/TMAG.2004.832160>. – DOI 10.1109/TMAG.2004.832160. – ISSN 0018–9464
- [20] HAMMES, P. : *Numerische Simulation von Streumatrizen im Hochfrequenzbereich*, TU Darmstadt, Diss., 1999
- [21] KOCH, W. : *Raum- und frequenzadaptive numerische Simulation zeitharmonischer elektromagnetischer Felder in translatorisch periodischen Strukturen*, TU Darmstadt, Diss., 2008
- [22] LUBKOWSKI, G. : *Simulation of Electromagnetic Fields in Double Negative Metamaterials*, TU Darmstadt, Diss., 2009
- [23] MUR, G. : Absorbing Boundary Conditions for the Finite-Difference Approximation of the Time-Domain Electromagnetic-Field Equations. In: *Electromagnetic Compatibility, IEEE Transactions on EMC-23* (1981), nov., Nr. 4, S. 377 –382. <http://dx.doi.org/10.1109/TEMC.1981.303970>. – DOI 10.1109/TEMC.1981.303970. – ISSN 0018–9375
- [24] SCHMITT, D. : *Zur numerischen Berechnung von Resonatoren und Wellenleitern*, TH Darmstadt, Diss., 1994
- [25] SCHUHMANN, R. ; WEILAND, T. : Conservation of Discrete Energy and Related Laws in the Finite Integration Technique. In: *Progress In Electromagnetics Research* 32 (2001), S. 301–316
- [26] SCHUHMANN, R. : *Die nichtorthogonale Finite-Integrations-Methode zur Simulation elektromagnetischer Felder*, TU Darmstadt, Diss., 1999
- [27] SCHUHMANN, R. : *Effective material parameters in metamaterials from transfer matrices*. 20.05.2002
- [28] SCHUHMANN, R. ; WEILAND, T. : *Vorlesungsskript Elektromagnetische Feldsimulation*. Warburger Str. 100 33098 Paderborn, 2009
- [29] SMITH, D. R. ; VIER, D. C. ; KROLL, N. ; SCHULTZ, S. : Direct calculation of permeability and permittivity for a left-handed metamaterial. In: *Applied Physics Let-*

- ters 77 (2000), Nr. 14, 2246-2248. <http://dx.doi.org/10.1063/1.1314884>. – DOI 10.1063/1.1314884
- [30] TEIXEIRA, F. ; CHEW, W. : General closed-form PML constitutive tensors to match arbitrary bianisotropic and dispersive linear media. In: *Microwave and Guided Wave Letters, IEEE* 8 (1998), Jun, Nr. 6, S. 223–225. <http://dx.doi.org/10.1109/75.678571>. – DOI 10.1109/75.678571. – ISSN 1051–8207
- [31] TEIXEIRA, F. ; CHEW, W. : Complex Space Approach to Perfectly Matched Layers: A Review and Some New Developments. In: *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields* 13 (2000), Nr. 5, S. 441–455
- [32] THOMA, P. : *Zur numerischen Lösung der Maxwellschen Gleichungen im Zeitbereich*, TU Darmstadt, Diss., 1997
- [33] TREFFTZ, E. : Ein Gegenstück zum Ritzschen Verfahren. In: MEISSNER, E. (Hrsg.): *Verhandlungen des 2. Internationalen Kongresses für Technische Mechanik*. Zürich : Orell Füssli Verlag, 1926, S. 131–137
- [34] TSUKERMAN, I. : Electromagnetic applications of a new finite-difference calculus. In: *Magnetics, IEEE Transactions on* 41 (2005), July, Nr. 7, S. 2206–2225. <http://dx.doi.org/10.1109/TMAG.2005.847637>. – DOI 10.1109/TMAG.2005.847637. – ISSN 0018–9464
- [35] TSUKERMAN, I. : A new computational method for plasmon resonances of nanoparticles and for wave propagation. In: *Wireless Communications and Applied Computational Electromagnetics, 2005. IEEE/ACES International Conference on*, 2005, S. 909–912
- [36] TSUKERMAN, I. : A class of difference schemes with flexible local approximation. In: *J. Comput. Phys.* 211 (2006), Nr. 2, 659-699. <http://dx.doi.org/10.1016/J.JCP.2005.06.011>. – DOI 10.1016/J.JCP.2005.06.011. – ISSN 0021–9991
- [37] TSUKERMAN, I. ; LOCKWOOD, D. J. (Hrsg.): *Computational Methods for Nanoscale Applications: Particles, Plasmons and Waves*. Springer, 2007 (Nanostructure Science and Technology)
- [38] WEILAND, T. : Eine Methode zur Lösung der Maxwellschen Gleichungen für sechskomponentige Felder auf diskreter Basis. In: *Archiv für Elektronik und Übertragungstechnik* 31 (1977), S. 116–120
- [39] WEILAND, T. : Time Domain Electromagnetic Field Computation with Finite Difference Methods. In: *International Journal of Numerical Modelling* 9 (1996), Nr. 4, S. 295–319. [http://dx.doi.org/10.1002/\(SICI\)1099-1204\(199607\)9:4<295::AID-JNM240>3.0.CO;2-8](http://dx.doi.org/10.1002/(SICI)1099-1204(199607)9:4<295::AID-JNM240>3.0.CO;2-8). – DOI 10.1002/(SICI)1099–1204(199607)9:4<295::AID-JNM240>3.0.CO;2–8
- [40] WEILAND, T. : *Numerische Lösung des Eigenwellenproblems längshomogener Wellenleiter*, TH Darmstadt, Diss., 1977
- [41] WEILAND, T. : Modes in infinitely Repeating Structures of Cylindrical Symmetry. In: *Proceedings of the International Linear Accelerator Conference (LINAC 1986)*, pp. 292-295
- [42] WEILAND, T. : High Precision Eigenmode Computation. In: *Particle Accelerators (PAC)* 56 (1996), 61-82. <http://cdsweb.cern.ch/record/1120276/files/p61.pdf>

- [43] YEE, K. S.: Numerical Solution of Initial Boundary Value Problems Involving Maxwell's Equations in Isotropic Media. In: *IEEE Transactions on Antennas and Propagation* 14 (1966), S. 302
- [44] ZHAO, L. ; CANGELLARIS, A. : GT-PML: Generalized Theory of Perfectly Matched Layers and its Application to the Reflectionless Truncation of Finite-Difference Time-Domain Grids. In: *IEEE Transactions on Microwave Theory and Techniques* 44 (1996), Nr. 12, S. 2555–2563. <http://dx.doi.org/10.1109/22.554601>. – DOI 10.1109/22.554601. – ISSN 0018–9480

Numerische Methoden zur Eigenwertberechnung

- [45] ARBENZ, P. ; HOCHSTENBACH, M. E.: A Jacobi–Davidson Method for Solving Complex Symmetric Eigenvalue Problems. In: *SIAM Journal on Scientific Computing* 25 (2004), Nr. 5, 1655–1673. <http://dx.doi.org/10.1137/S1064827502410992>. – DOI 10.1137/S1064827502410992
- [46] BAI, Z. (Hrsg.) ; DEMMEL, J. (Hrsg.) ; DONGARRA, J. (Hrsg.) ; RUHE, A. (Hrsg.) ; VAN DER VORST, H. (Hrsg.): *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*. Philadelphia : SIAM, 2000
- [47] BARRETT, R. ; BERRY, M. ; CHAN, T. F. ; DEMMEL, J. ; DONATO, J. ; DONGARRA, J. ; EIJKHOUT, V. ; POZO, R. ; ROMINE, C. ; VORST, H. V.: *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, 2nd Edition*. Philadelphia, PA : SIAM, 1994
- [48] BETCKE, T. : Optimal Scaling of Generalized and Polynomial Eigenvalue Problems. In: *SIAM J. Matrix Anal. Appl.* 30 (2008), Nr. 4, S. 1320–1338. <http://dx.doi.org/10.1137/070704769>. – DOI 10.1137/070704769. – ISSN 0895–4798
- [49] CUTHILL, E. ; MCKEE, J. : Reducing the bandwidth of sparse symmetric matrices. In: *Proceedings of the 1969 24th national conference*. New York, NY, USA : ACM, 1969, S. 157–172
- [50] DAVIDSON, E. : The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices. In: *J. Comput. Phys.* 17 (1975), S. 87–94
- [51] DAVIS, T. A.: Algorithm 832: UMFPACK V4.3—an unsymmetric-pattern multifrontal method. In: *ACM Trans. Math. Softw.* 30 (2004), Nr. 2, S. 196–199. – ISSN 0098–3500
- [52] DAVIS, T. A. ; GILBERT, J. R. ; LARIMORE, S. I. ; NG, E. G.: Algorithm 836: COLAMD, a column approximate minimum degree ordering algorithm. In: *ACM Trans. Math. Softw.* 30 (2004), Nr. 3, S. 377–380. <http://dx.doi.org/http://doi.acm.org/10.1145/1024074.1024080>. – DOI <http://doi.acm.org/10.1145/1024074.1024080>. – ISSN 0098–3500
- [53] DEUFLHARD, P. ; HOHMANN, A. : *Numerische Mathematik 1. Eine algorithmisch orientierte Einführung*. Berlin, New York : de Gruyter, 2008
- [54] FEIGH, S. ; CLEMENS, M. ; SCHUHMAN, R. ; WEILAND, T. : Preconditioned Jacobi–Davidson Algorithm for Nonsymmetric, Complex Eigenvalue Problems Arising with

- Gyrotropic Materials in Resonator Cavities. In: *Proceedings of the 8th European Particle Accelerator Conference (EPAC 2002)* (2002), S. 1625–1627
- [55] FOKKEMA, D. R. ; SLEIJPEN, G. L. G. ; VAN DER VORST, H. A.: Accelerated Inexact Newton Schemes for Large Systems of Nonlinear Equations. In: *SIAM J. Sci. Comput.* 19 (1998), Nr. 2, S. 657–674. <http://dx.doi.org/10.1137/S1064827595296148>. – DOI 10.1137/S1064827595296148. – ISSN 1064–8275
- [56] FOKKEMA, D. R. ; SLEIJPEN, G. L. G. ; VAN DER VORST, H. A.: Jacobi-Davidson Style QR and QZ Algorithms for the Reduction of Matrix Pencils. In: *SIAM Journal on Scientific Computing* 20 (1998), Nr. 1, S. 94–125. <http://dx.doi.org/10.1137/S1064827596300073>. – DOI 10.1137/S1064827596300073
- [57] GENSEBERGER, M. ; SLEIJPEN, G. L. G.: Alternative correction equations in the Jacobi-Davidson method. In: *Numer. Linear Algebra Appl.* 6 (1999), Nr. 3, S. 235–253. – ISSN 1099–1506
- [58] GOHBERG, I. ; LANCASTER, P. ; RODMAN, L. : *Matrix polynomials*. Academic Press, 1982 (Computer science and applied mathematics). – ISBN 9780122871603
- [59] GOLUB, G. H. ; VAN DER VORST, H. A.: Eigenvalue Computation in the 20th Century. In: *Journal of Computational and Applied Mathematics* 123 (2000), S. 35–65
- [60] GOLUB, G. H. ; VAN LOAN, C. F.: *Matrix computations (3rd ed.)*. Johns Hopkins University Press <http://portal.acm.org/citation.cfm?id=248979>. – ISBN 0801854148
- [61] GOURLAY, A. R. ; WATSON, G. A.: *Computational Methods for Methods for Matrix Eigenproblems*. New York : John Wiley and Sons, 1973
- [62] HANKE-BOURGEOIS, M. : *Grundlagen der Numerischen Mathematik und des Wissenschaftlichen Rechnens*. Vieweg+Teubner, 2009 (Mathematics and Statistics). – ISBN 978-3-8348-9309-3
- [63] HOCHSTENBACH, M. : *Jacobi-Davidson Gateway*. Website. <http://www.win.tue.nl/casa/research/topics/jd/>
- [64] HOCHSTENBACH, M. E. ; SLEIJPEN, G. L. G.: Harmonic and refined Rayleigh-Ritz for the polynomial eigenvalue problem. In: *Numerical Linear Algebra with Applications* 15 (2008), Nr. 1, 35–54. <http://dx.doi.org/10.1002/nla.562>. – DOI 10.1002/nla.562
- [65] HWANG, T.-M. ; LIN, W.-W. ; LIU, J.-L. ; WANG, W. : Jacobi-Davidson methods for cubic eigenvalue problems. In: *Numerical Linear Algebra with Applications* 12 (2005), Nr. 7, 605–624. <http://dx.doi.org/10.1002/nla.423>. – DOI 10.1002/nla.423
- [66] JACOBI, C. : Über ein leichtes Verfahren, die in der Theorie der Säcularstörungen vorkommenden Gleichungen numerisch aufzulösen. In: *J. Reine und Angew. Math.* (1846), S. 51–94
- [67] KRESSNER, D. : *Lecture Notes in Computational Science and Engineering*. Bd. 46: *Numerical Methods for General and Structured Eigenvalue Problems*. Heidelberg : Springer, 2005
- [68] LÖCHEL, D. : *Numerical methods for eigenvalue problems in the description of instabilities in the plasma edge*, Universität Düsseldorf, Diss., 2009

- [69] OSTROWSKI, A. M.: On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors. V. In: *Archive for Rational Mechanics and Analysis* 3 (1959), 472-481. <http://dx.doi.org/10.1007/BF00284194>. – ISSN 0003–9527. – ibid Bände I–VI
- [70] PAIGE, C. C. ; PARLETT, B. N. ; VORST, H. A. d.: Approximate solutions and eigenvalue bounds from Krylov subspaces. In: *Numer. Linear Algebra Appl.* 2 (1995), Nr. 2, 115–133. <http://dx.doi.org/10.1002/nla.1680020205>. – ISSN 1099–1506
- [71] RIENEN, U. van: *Numerical Methods in Computational Electrodynamics - Linear Systems in Practical Applications*. Springer, 2001 (Springer Lecture Notes in Computational Science and Engineering)
- [72] SAAD, Y. : *Numerical Methods for Large Eigenvalue Problems: Theory and Algorithms*. New York : John Wiley and Sons, 1992
- [73] SAAD, Y. : *Iterative Methods for Sparse Linear Systems, 2nd edition*. Philadelphia, PA : SIAM, 2003
- [74] SCHENK, O. ; GÄRTNER, K. : Solving unsymmetric sparse systems of linear equations with PARDISO. In: *Future Gener. Comput. Syst.* 20 (2004), Nr. 3, S. 475–487. – ISSN 0167–739X
- [75] SCHREIBER, K. : *Nonlinear Eigenvalue Problems: Newton-type Methods and Nonlinear Rayleigh Functionals*, TU Berlin, Diss., 2008. <http://opus.kobv.de/tuberlin/volltexte/2008/1875/>
- [76] SLEIJPEN, G. L. G. ; FOKKEMA, D. R.: Bi-CGSTAB(l) for linear equations involving unsymmetric matrices with complex spectrum. In: *Elec. Trans. Numer. Anal.* 1 (1993), S. 11–32
- [77] SLEIJPEN, G. L. G. ; VAN DER VORST, H. A.: The Jacobi-Davidson method for eigenvalue problems and its relation with accelerated inexact Newton schemes. In: MARGENOV, S. D. (Hrsg.) ; VASSILEVSKI, P. S. (Hrsg.): *Iterative Methods in Linear Algebra II*, 1996 (IMACS Ser. Comput. Appl. Math. 3), S. 377–389
- [78] SLEIJPEN, G. L. G. ; BOOTEN, A. G. L. ; FOKKEMA, D. R. ; VAN DER VORST, H. A.: Jacobi-davidson type methods for generalized eigenproblems and polynomial eigenproblems. In: *BIT Numerical Mathematics* 36 (1996), 595-633. <http://dx.doi.org/10.1007/BF01731936>. – ISSN 0006–3835. – 10.1007/BF01731936
- [79] SLEIJPEN, G. L. G. ; VAN DER VORST, H. A.: A Jacobi–Davidson Iteration Method for Linear Eigenvalue Problems. In: *SIAM Review* 42 (2000), Nr. 2, 267-293. <http://dx.doi.org/10.1137/S0036144599363084>. – DOI 10.1137/S0036144599363084
- [80] SLEIJPEN, G. L. G. ; VORST, H. A. d.: A Jacobi-Davidson iteration method for linear eigenvalue problems. In: *SIAM J. Matrix Anal. Appl.* 17 (1996), Nr. 2, S. 401–425. – ISSN 0895–4798
- [81] SPELLUCCI, P. : *Vorlesungsskript Numerik des Matrizen-Eigenwertproblems*. Schloßgartenstr. 7 64289 Darmstadt, 2007
- [82] STRUTT, J. W.: Some General Theorems relating to Vibrations. In: *Proc. London Math. Soc.* s1-4 (1871), Nr. 1, 357-368. <http://dx.doi.org/10.1112/plms/s1-4.1.357>. – DOI 10.1112/plms/s1-4.1.357

- [83] TISSEUR, F. ; HIGHAM, N. J.: Structured Pseudospectra for Polynomial Eigenvalue Problems, with Applications. In: *SIAM Journal on Matrix Analysis and Applications* 23 (2001), Nr. 1, 187-208. <http://dx.doi.org/10.1137/S0895479800371451>. – DOI 10.1137/S0895479800371451
- [84] TISSEUR, F. : Backward error and condition of polynomial eigenvalue problems. In: *Linear Algebra and its Applications* 309 (2000), Nr. 1-3, 339 - 361. [http://dx.doi.org/10.1016/S0024-3795\(99\)00063-4](http://dx.doi.org/10.1016/S0024-3795(99)00063-4). – DOI 10.1016/S0024-3795(99)00063-4. – ISSN 0024-3795
- [85] ÜBERHUBER, C. : *Computer-Numerik 2*. Springer, 1995
- [86] VAN DER VORST, H. A.: Computational methods for large eigenvalue problems. Version: 2002. [http://dx.doi.org/10.1016/S1570-8659\(02\)08003-1](http://dx.doi.org/10.1016/S1570-8659(02)08003-1). In: *Solution of Equations in \mathbb{R}^n (Part 4), Techniques of Scientific Computing (Part 4), Numerical Methods for Fluids (Part 2)* Bd. 8. Elsevier, 2002. – DOI 10.1016/S1570-8659(02)08003-1. – ISSN 1570-8659, S. 3 – 179
- [87] VAN DER VORST, H. A. ; GOLUB, G. H.: *150 Years Old and Still Alive: Eigenproblems*. <http://www.math.uu.nl/publications/preprints/981.ps.gz>. Version: 1997
- [88] WATKINS, D. S.: A case where balancing is harmful. In: *Electron. Trans. Numer. Anal.* 23 (2006), 1-4. <http://etna.mcs.kent.edu/>
- [89] WIKIPEDIA: *Eigenvalue, eigenvector and eigenspace*. Webseite. http://en.wikipedia.org/wiki/Eigenvalue,_eigenvector_and_eigenspace#History. – Zuletzt abgerufen am 12. Dezember 2010.
- [90] WILKINSON, J. H.: *The Algebraic Eigenvalue Problem*. Oxford, England : Clarendon Press, 1965

Ergebnisse und Anwendungen

- [91] AKAHANE, Y. ; ASANO, T. ; SONG, B.-S. ; NODA, S. : High-Q photonic nanocavity in a two-dimensional photonic crystal. In: *Nature* 425 (2003), Okt., Nr. 6961, 944-947. <http://dx.doi.org/10.1038/nature02063>. – ISSN 0028-0836
- [92] BIRNER, A. ; WEHRSPHON, R. B. ; GÖSELE, U. ; BUSCH, K. : Silicon-based photonic crystals. In: *Advanced Materials* 13 (2001), Nr. 6, S. 377-388
- [93] BITTNER, S. ; DIETZ, B. ; MISKI-UGLU, M. ; IRIARTE, P. O. ; RICHTER, A. ; SCHÄFER, F. : Experimental test of a two-dimensional approximation for dielectric microcavities. In: *Phys. Rev. A* 80 (2009), Aug, Nr. 2, S. 023825. <http://dx.doi.org/10.1103/PhysRevA.80.023825>. – DOI 10.1103/PhysRevA.80.023825
- [94] CAI, W. ; SHALAEV, V. : *Optical Metamaterials — Fundamentals and Applications*. New York : Springer, 2010
- [95] ČAJKO, F. ; TSUKERMAN, I. : Flexible Approximation Schemes for Wave Refraction in Negative Index Materials. In: *Magnetics, IEEE Transactions on* 44 (2008), Nr. 6, S. 1378-1381. <http://dx.doi.org/10.1109/TMAG.2007.916167>. – DOI 10.1109/TMAG.2007.916167. – ISSN 0018-9464

- [96] CHEN, X. ; GRZEGORCZYK, T. M. ; BAE-IAN, W. ; PACHECO, J. ; KONG, J. A.: Robust method to retrieve the constitutive effective parameters of metamaterials. In: *Physical Review E* 70 (2004), S. 016608
- [97] DEHLER, M. : *Numerische Lösung der Maxwell'schen Gleichungen auf kreiszylindrischen Gittern*, TH Darmstadt, Diss., 1993
- [98] DOLLING, G. ; ENKRICH, C. ; WEGENER, M. ; SOUKOULIS, C. M. ; LINDEN, S. : Low-loss negative-index metamaterial at telecommunication wavelengths. In: *Opt. Lett.* 31 (2006), Nr. 12, 1800–1802. <http://ol.osa.org/abstract.cfm?URI=ol-31-12-1800>
- [99] ELEFThERIADES, G. V. (Hrsg.) ; BALMAIN, K. G. (Hrsg.): *Negative-Refraktion Metamaterials – Fundamental Principles and Applications*. IEEE PRESS, 2005
- [100] ELGER, M. : *Adaptive mehrdimensionale rationale Interpolation zur schnellen Bestimmung komplexer Größen bei der Charakterisierung von HF-Komponenten*, Technische Universität Darmstadt, Diss., 2009
- [101] ENGHETA, N. (Hrsg.) ; ZIOLKOWSKI, R. (Hrsg.): *Metamaterials (Physics and Engineering Explorations)*. Wiley & Sons, 2006
- [102] HEBERMEHL, G. ; HÜBNER, F.-K. ; SCHLUNDT, R. ; TISCHLER, T. ; ZSCHEILE, H. ; HEINRICH, W. : Eigen mode computation of microwave and laser structures including PML. In: SCHILDERS, W. H. A. (Hrsg.) ; MATEN, E. J. W. (Hrsg.) ; HOUBEN, S. H. M. J. (Hrsg.): *Scientific Computing in Electrical Engineering* Bd. 4, Springer Verlag (Mathematics in Industry), 196–205
- [103] JOHN, S. : Strong localization of photons in certain disordered dielectric superlattices. In: *Phys. Rev. Lett.* 58 (1987), Jun, Nr. 23, S. 2486–2489. <http://dx.doi.org/10.1103/PhysRevLett.58.2486>. – DOI 10.1103/PhysRevLett.58.2486
- [104] JOHNSON, P. B. ; CHRISTY, R. W.: Optical Constants of the Noble Metals. In: *Phys. Rev. B* 6 (1972), Dec, Nr. 12, S. 4370–4379. <http://dx.doi.org/10.1103/PhysRevB.6.4370>. – DOI 10.1103/PhysRevB.6.4370
- [105] KLEIN, M. W. ; ENKRICH, C. ; WEGENER, M. ; SOUKOULIS, C. M. ; LINDEN, S. : Single-slit split-ring resonators at optical frequencies: limits of size scaling. In: *Opt. Lett.* 31 (2006), Nr. 9, 1259–1261. <http://ol.osa.org/abstract.cfm?URI=ol-31-9-1259>
- [106] MANDELSHTAM, V. A. ; TAYLOR, H. S.: Harmonic inversion of time signals. In: *J. Chem. Phys.* 107 (1997), S. 6756–6769. – Erratum, *ibid*, Bd. 109, Nr. 10, S. 4128 (1998)
- [107] MARKOŠ, P. ; SOUKOULIS, C. M.: Transmission properties effective electromagnetic parameters of double negative metamaterials. In: *Optics Express* 11 (2003), Nr. 7, S. 649
- [108] NICOLSON, A. M. ; ROSS, G. F.: Measurement of the Intrinsic Properties of Materials by Time-Domain Techniques. In: *IEEE Transactions on Instrumentation and Measurement* IM-19 (1970), Nr. 4, S. 377
- [109] PENDRY, J. B. ; HOLDEN, A. J. ; ROBBINS, D. J. ; STEWART, W. J.: Magnetism from conductors and enhanced nonlinear phenomena. In: *IEEE T-MTT* 47 (1999), Nr. 11, S. 2075
- [110] PINHEIRO, H. ; WEBB, J. : A FLAME Molecule for 3-D Electromagnetic Scattering. In: *Magnetics, IEEE Transactions on* 45 (2009), Nr. 3, S. 1120–1123. <http://dx.doi.org/10.1109/TMAG.2009.4800000>

- doi.org/10.1109/TMAG.2009.2012648. – DOI 10.1109/TMAG.2009.2012648. – ISSN 0018–9464
- [111] RUSSELL, P. S.: Photonic-Crystal Fibers. In: *J. Lightwave Technol.* 24 (2006), Nr. 12, 4729–4749. <http://jlt.osa.org/abstract.cfm?URI=JLT-24-12-4729>
- [112] SCHMITT, D. ; STEFFEN, B. ; WEILAND, T. : 2D and 3D computations of lossy eigenvalue problems. In: *Magnetics, IEEE Transactions on* 30 (1994), Nr. 5, S. 3578–3581
- [113] SCHUHMAN, R. ; WITTIG, T. ; MUNTEANU, I. ; TRAPP, B. ; WEILAND, T. : Time and Frequency Domain Simulations of Highly Resonant RF-Devices. In: *International Conference on Electromagnetics in Advanced Applications (ICEAA)*, 2001, S. 125–128
- [114] SHELBY, R. A. ; SMITH, D. R. ; SCHULTZ, S. : Experimental Verification of a Negative Index of Refraction. In: *Science* 292 (2001), Nr. 5514, 77-79. <http://dx.doi.org/10.1126/science.1058847>. – DOI 10.1126/science.1058847
- [115] SMITH, D. R. ; VIER, D. C. ; KROLL, N. ; SCHULTZ, S. : Direct calculation of permeability and permittivity for a left-handed metamaterial. In: *Applied Physics Letters* 77 (2000), Nr. 14, S. 2246
- [116] SMOTROVA, E. ; NOSICH, A. ; BENSON, T. ; SEWELL, P. : Cold-cavity thresholds of microdisks with uniform and nonuniform gain: quasi-3-D modeling with accurate 2-D analysis. In: *Selected Topics in Quantum Electronics, IEEE Journal of* 11 (2005), Nr. 5, S. 1135 – 1142. – ISSN 1077–260X
- [117] VAHALA, K. J.: Optical Microcavities. In: *Nature* 424 (2003), Nr. 6950, S. 839–846. <http://dx.doi.org/10.1038/nature01939>. – DOI 10.1038/nature01939
- [118] WEILAND, T. : On the computation of resonant modes in cylindrically symmetric cavities. In: *Nuclear Instruments and Methods in Physics Research* 216 (1983), Nr. 3, 329 - 348. [http://dx.doi.org/DOI:10.1016/0167-5087\(83\)90500-8](http://dx.doi.org/DOI:10.1016/0167-5087(83)90500-8). – DOI DOI: 10.1016/0167–5087(83)90500–8. – ISSN 0167–5087
- [119] WEILAND, T. ; SCHUHMAN, R. ; GREGOR, R. ; PARAZZOLI, C. ; VETTER, A. ; SMITH, D. ; VIER, D. ; SCHULTZ, S. : Ab initio numerical simulation of left-handed metamaterials: Comparison of calculations and experiment. In: *Journal of Applied Physics* 90 (2001), Nr. 10, S. 5419
- [120] WEITSCH, Y. ; EIBERT, T. : Continuous Beam-Steering Leaky-Wave Antenna based on Substrate Integrated Waveguide. In: *Antennas and Propagation, 2007. EuCAP 2007. The Second European Conference on*, 2007, S. 1 –5
- [121] WIERSIG, J. ; HENTSCHEL, M. : Unidirectional light emission from high- Q modes in optical microcavities. In: *Phys. Rev. A* 73 (2006), Nr. 3, S. 031802
- [122] YABLONOVITCH, E. : Inhibited Spontaneous Emission in Solid-State Physics and Electronics. In: *Phys. Rev. Lett.* 58 (1987), May, Nr. 20, S. 2059–2062. <http://dx.doi.org/10.1103/PhysRevLett.58.2059>. – DOI 10.1103/PhysRevLett.58.2059

Verwendete Software

- [123] COMPUTER SIMULATION TECHNOLOGY AG (CST): *CST Studio Suite*. 2010. – <http://www.cst.com/>
- [124] INTEL: *Math Kernel Library*. 2009. – <http://www.intel.com/>
- [125] JOHNSON, S. G.: *Nanostructures and Computation Wiki at MIT!*. – <http://ab-initio.mit.edu/wiki/index.php/Harminv>
- [126] THE MATHWORKS: *Matlab 2009b*. 2009. – <http://www.mathworks.com/>

Eigene Veröffentlichungen im Rahmen dieser Arbeit

- [127] BANDLOW, B. ; CLASSEN, C. ; SCHUHMANN, R. : Time and Frequency Domain Simulation of Photonic Nanocavities. In: *International URSI Symposium on Electromagnetic Theory EMTS 2010 - Symposium Digest, 16. - 19. August*. Berlin, Germany, 2010, S. 275–278
- [128] BANDLOW, B. ; LUBKOWSKI, G. ; SCHUHMANN, R. ; WEILAND, T. : An Analysis of Single-Cell Modeling of Periodic Metamaterial Structures. In: *Proceedings of the 16th Conference on the Computation of Electromagnetic Fields COMPUMAG 2007, 24. - 28. Juni*. Aachen, Germany, June 2007, S. 501–502
- [129] BANDLOW, B. ; SCHUHMANN, R. : Analysis of Higher Order Modes in the Design of Metamaterials by the Use of Periodic Unit Cells. In: BILOTTI, F. (Hrsg.) ; VEGNI, L. (Hrsg.) ; University "Roma Tre", Rome, Italy (Veranst.): *Proceedings of First International Congress on Advanced Electromagnetic Materials in Microwaves and Optics (METAMATERIALS 2007), 22. - 24. Oktober*. Rome, Italy, October 2007, S. 736–739
- [130] BANDLOW, B. ; SCHUHMANN, R. : Electromagnetic Simulation of Unit Cells in Metamaterial Arrays. In: *386. WE-Heraeus-Seminar Computational Nano-Photonics, 25. - 28. Februar*. Bad Honnef, Germany, 2007
- [131] BANDLOW, B. ; SCHUHMANN, R. : Computation of Eigenmodes in Dispersive Materials. In: *The 7th International Conference on Scientific Computing in Electrical Engineering, SCEE 2008, 28. September - 3. Oktober*. Espoo, Finland, 2008, S. 9–10
- [132] BANDLOW, B. ; SCHUHMANN, R. : Eigenmodeberechnung von Periodischen Strukturen aus Dispersiven Materialien. In: *Kleinheubacher Tagung 2008 des U.R.S.I. Landesausschuss in der Bundesrepublik Deutschland e.V., 22. - 25. September*. Miltenberg, Germany, 2008
- [133] BANDLOW, B. ; SCHUHMANN, R. : Scattering Matrix Based Analysis of Metamaterial Unit Cells. In: *The 9th International Workshop on Finite Elements for Microwave Engineering, 8. - 9. Mai*. Bonn, Germany, 2008
- [134] BANDLOW, B. ; SCHUHMANN, R. : Simulation of Metamaterial Unit Cells in the THz Region. In: *403. WE-Heraeus-Seminar on Periodic Structures, 27. Februar*. Bad Honnef, Germany, 2008

- [135] BANDLOW, B. ; SCHUHMANN, R. : Some Constraints of the Concept of Effective Material Parameters in Metamaterials. In: *German Microwave Conference, 10. - 12. März* Bd. ITG-Fachbericht 206. Hamburg-Harburg, Germany, 2008. – ISBN 978-3-8007-3086-5
- [136] BANDLOW, B. ; SCHUHMANN, R. : 3-D Eigenmode Calculation of Metallic Nanostructures. In: *Advances in Radio Science* 7 (2009), 23-27. <http://www.adv-radio-sci.net/7/23/2009/>. – ISSN 1684-9965
- [137] BANDLOW, B. ; SCHUHMANN, R. : Computation Of Dispersion Relations Of Fishnet Metamaterial Unit Cells. In: *2009 IEEE International Symposium on Antennas and Propagation and USNC/URSI National Radio Science Meeting, 1. - 5. Juni*. Charleston, SC, USA, 2009
- [138] BANDLOW, B. ; SCHUHMANN, R. : A Mode Selecting Eigensolver for 2D FIT Models of Waveguides. In: *Applied Computational Electromagnetics Society Journal* 24 (2009), December, Nr. 6, 592-599. <http://aces.ee.olemiss.edu/search.php?vol=24&no=6&type=2>. – ISSN 1054-4887
- [139] BANDLOW, B. ; SCHUHMANN, R. : A Mode Selecting Eigensolver for 2D FIT Models of Waveguides. In: *25th Annual Review of Progress in Applied Computational Electromagnetics (ACES), 8. - 12. März*. Monterey, California, USA, 2009, S. 222-226
- [140] BANDLOW, B. ; SCHUHMANN, R. : Computation of Eigenmodes in Dispersive Materials. In: ROOS, J. (Hrsg.) ; COSTA, L. R. (Hrsg.): *Scientific Computing in Electrical Engineering SCEE 2008* Bd. 14. 1st. Springer, 2010. – ISBN 978-3-642-12293-4, S. 61-68
- [141] BANDLOW, B. ; SCHUHMANN, R. : Mode Selecting Eigensolver for 3D Computational Models. In: *The 8th International Conference on Scientific Computing in Electrical Engineering, SCEE 2010, 19. - 24. September*. Toulouse, France, 2010, S. 101-102
- [142] BANDLOW, B. ; SCHUHMANN, R. ; LUBKOWSKI, G. ; WEILAND, T. : Analysis of Single-Cell Modeling of Periodic Metamaterial Structures. In: *IEEE Transactions on Magnetics* 44 (2008), June, Nr. 6, S. 1662-1665. <http://dx.doi.org/10.1109/TMAG.2007.916037>. – DOI 10.1109/TMAG.2007.916037. – ISSN 0018-9464
- [143] BANDLOW, B. ; SIEVERS, D. ; SCHUHMANN, R. : An Improved Jacobi-Davidson Method for the Computation of Selected Eigenmodes in Waveguides. In: *Proceedings of the 17th Conference on the Computation of Electromagnetic Fields (COMPUMAG 2009), 22. - 26. November*. Florianopolis, Brazil, 2009, S. 548-549
- [144] BANDLOW, B. ; SIEVERS, D. ; SCHUHMANN, R. : An Improved Jacobi-Davidson Method for the Computation of Selected Eigenmodes in Waveguide Cross Sections. In: *Magnetics, IEEE Transactions on* 46 (2010), Nr. 8, S. 3461-3464. <http://dx.doi.org/10.1109/TMAG.2010.2046315>. – DOI 10.1109/TMAG.2010.2046315. – ISSN 0018-9464
- [145] CLASSEN, C. ; BANDLOW, B. ; SCHUHMANN, R. : Computational Analysis of Whispering Gallery Modes in Flat Dielectric Disks. In: *International Conference on Electromagnetics in Advanced Applications, 20. - 24. September*. Sydney, Australia, 2010
- [146] CLASSEN, C. ; BANDLOW, B. ; SCHUHMANN, R. ; TSUKERMAN, I. : FIT & FLAME for Sharp Edges in Electrostatics. In: *International URSI Symposium on Electromagnetic Theory EMTS 2010 - Symposium Digest, 16. - 19. August*. Berlin, Germany, 2010, S. 56-59

- [147] LUBKOWSKI, G. ; BANDLOW, B. ; SCHUHMANN, R. ; WEILAND, T. : Effective Modeling of Double Negative Metamaterial Macrostructures. In: *IEEE Transactions on Microwave Theory and Techniques* 57 (2009), May, Nr. 5, S. 1136–1146. <http://dx.doi.org/10.1109/TMTT.2009.2017349>. – DOI 10.1109/TMTT.2009.2017349
- [148] SCHUHMANN, R. ; BANDLOW, B. ; LUBKOWSKI, G. ; WEILAND, T. : Mikro- und makroskopische Simulation von periodischen Metamaterialien. In: *Kleinheubacher Tagung 2007 des U.R.S.I. Landesausschuss in der Bundesrepublik Deutschland e.V., 24. - 27. September*. Miltenberg, Germany, 2007
- [149] SCHUHMANN, R. ; BANDLOW, B. ; LUBKOWSKI, G. ; WEILAND, T. : Micro- and Macroscopic Simulation of Periodic Metamaterials. In: *Advances in Radio Science* 6 (2008), 77-82. <http://www.adv-radio-sci.net/6/77/2008/>. – ISSN 1684–9965

Eigene Veröffentlichungen außerhalb dieser Arbeit

- [150] BANDLOW, B. ; FÖRSTNER, J. ; MEIER, T. ; SCHUHMANN, R. : Computational Optics at the University of Paderborn. In: *Heraeus Seminar, 02.06.2008*. Bad Honnef, Germany, 2008
- [151] LUBKOWSKI, G. ; DAMM, C. ; BANDLOW, B. ; SCHUHMANN, R. ; SCHÜSSLER, M. ; WEILAND, T. : Waveguide Miniaturization Using Spiral Resonators and Dipole Arrays. In: *36th European Microwave Conference (EuMC), 10. - 15. September*. Manchester, UK, 2006, S. 1312–1315
- [152] LUBKOWSKI, G. ; DAMM, C. ; BANDLOW, B. ; SCHUHMANN, R. ; SCHÜSSLER, M. ; WEILAND, T. : Broadband Transmission Below the Cutoff Frequency of a Waveguide Loaded with Resonant Scatterers Arrays. In: *IET Microwaves, Antennas & Propagation* 1 (2007), Feb, Nr. 1, S. 165–169. <http://dx.doi.org/10.1049/iet-map:20050322>. – DOI 10.1049/iet-map:20050322
- [153] LUBKOWSKI, G. ; DAMM, C. ; BANDLOW, B. ; SCHUHMANN, R. ; SCHÜSSLER, M. ; WEILAND, T. : Experimental Study of Subwavelength Waveguides Loaded by Electric and Magnetic Resonant Scatterer Arrays. In: *38th European Microwave Conference (EuMc), 27. - 31. Oktober*. Amsterdam, The Netherlands, 2008, S. 1050–1053
- [154] LUBKOWSKI, G. ; DAMM, C. ; BANDLOW, B. ; SCHUHMANN, R. ; SCHÜSSLER, M. ; WEILAND, T. : Metamaterial Loaded Waveguides For Filter Applications. In: *Metamaterials Week 2008, 5-8 February 2008, Casa Convalescència*. Barcelona, Spain, 2008
- [155] LUBKOWSKI, G. ; DAMM, C. ; BANDLOW, B. ; SCHUHMANN, R. ; SCHÜSSLER, M. ; WEILAND, T. : Metamaterial Loaded Waveguides for Miniaturized Filter Applications. In: *Frequenz - TICMO Special* 62 (2008), Nr. 3-4, 71-74. <http://frequenz.schiele-schoen.de/108/12920/fre20803071/>
- [156] LUBKOWSKI, G. ; HIRTENFELDER, F. ; BANDLOW, B. ; SCHUHMANN, R. ; WEILAND, T. : Macromodeling of Parabolic Double Negative Metamaterial Antennas. In: *Frequenz - TICMO Special* 62 (2008), Nr. 3-4, 67-70. <http://frequenz.schiele-schoen.de/108/12919/fre20803067/>

- [157] SCHUHMANN, R. ; BANDLOW, B. ; FISCHER, C. ; GLAHN, T. : A Study on Wave Propagation in Cylindrical Optical Structures Using Finite Difference Methods. In: *2010 IEEE International Symposium on Antennas and Propagation and CNC-USNC/URSI Radio Science meeting, 11. - 17. Juli*. Toronto, Ontario, Canada, 2010
- [158] STALLEIN, M. ; BANDLOW, B. ; SCHUHMANN, R. : Ansatz quasi-geführter Moden zur Berechnung des Übertragungsverhaltens gekrümmter dielektrischer Schichtwellenleiter. In: *Kleinheubacher Tagung 2009 des U.R.S.I. Landesausschuss in der Bundesrepublik Deutschland e.V., 28. September - 01. Oktober*. Miltenberg, Germany, 2009. – 29.09.2009
- [159] STALLEIN, M. ; BANDLOW, B. ; SCHUHMANN, R. : Using Quasi-Guided Modes for Modeling the Transfer Behavior of Bent Dielectric Slab Waveguides. In: *Advances in Radio Science* 8 (2010), 19-26. <http://dx.doi.org/10.5194/ars-8-19-2010>. – DOI 10.5194/ars-8-19-2010
- [160] STALLEIN, M. ; BANDLOW, B. ; SCHUHMANN, R. : Validation of Ray Optics Capability to Analyze the Transfer Behavior of Bent Optical Slab Waveguides. In: *International URSI Symposium on Electromagnetic Theory EMTS 2010 - Symposium Digest, 16. - 19. August*. Berlin, Germany, 2010, S. 295–298

Betreute studentische Arbeiten

- [161] JONES, C. A.: *Implementation of Perfectly Matched Layers for Evanescent Waves in FIT*, Universität Paderborn, EIM-E, TET, Studienarbeit, Dezember 2009
- [162] KIRSCH, S. : *Absorbierende Randbedingungen für die kreiszylindrische Formulierung der Methode der Finiten Integration*, Universität Paderborn, EIM-E, TET, Diplomarbeit, April 2010