

Modellbasierte Untersuchung zur Propagation von Gausstrahlen durch periodisch dicke Gitter

Bachelorarbeit

vorgelegt am: 2. September 2019

**zur Erlangung des Bachelor of Science in Elektrotechnik
und Informationstechnik**

**an der Fakultät für Elektrotechnik und
Informationstechnik der Technischen Universität
Dortmund**

Name:	Tobias Lange
Matrikelnummer:	209060
Studiengang:	Elektrotechnik und Informationstechnik
Erstgutachter:	apl. Prof. Dr.-Ing. Dirk Schulz
Institut:	Lehrstuhl für Hochfrequenztechnik an der TU Dortmund
betreuendes Institut:	Fraunhofer Institut für Lasertechnik

**Inhalt und Ergebnisse dieser Arbeit sind ausschließlich zum internen
Gebrauch bestimmt. Alle Urheberrechte liegen bei Fraunhofer ILT.
Ohne ausdrückliche Genehmigung des Fraunhofer ILT ist es nicht
gestattet, diese Arbeit oder Teile an Dritte weiterzugeben.**

Eidesstattliche Versicherung (Affidavit)

Name, Vorname
(Last name, first name)

Matrikelnr.
(Enrollment number)

Ich versichere hiermit an Eides statt, dass ich die vorliegende Bachelorarbeit/Masterarbeit* mit dem folgenden Titel selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

I declare in lieu of oath that I have completed the present Bachelor's/Master's* thesis with the following title independently and without any unauthorized assistance. I have not used any other sources or aids than the ones listed and have documented quotations and paraphrases as such. The thesis in its current or similar version has not been submitted to an auditing institution.

Titel der Bachelor-/Masterarbeit*:
(Title of the Bachelor's/ Master's* thesis):

*Nichtzutreffendes bitte streichen
(Please choose the appropriate)

Ort, Datum
(Place, date)

Unterschrift
(Signature)

Belehrung:

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden. (§ 63 Abs. 5 Hochschulgesetz - HG -).

Die Abgabe einer falschen Versicherung an Eides statt wird mit Freiheitsstrafe bis zu 3 Jahren oder mit Geldstrafe bestraft.

Die Technische Universität Dortmund wird gfls. elektronische Vergleichswerkzeuge (wie z.B. die Software „turnitin“) zur Überprüfung von Ordnungswidrigkeiten in Prüfungsverfahren nutzen.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Official notification:

Any person who intentionally breaches any regulation of university examination regulations relating to deception in examination performance is acting improperly. This offense can be punished with a fine of up to €50,000.00. The competent administrative authority for the pursuit and prosecution of offenses of this type is the chancellor of TU Dortmund University. In the case of multiple or other serious attempts at deception, the examinee can also be unenrolled, section 63, subsection 5 of the North Rhine-Westphalia Higher Education Act (*Hochschulgesetz*).

The submission of a false affidavit will be punished with a prison sentence of up to three years or a fine.

As may be necessary, TU Dortmund will make use of electronic plagiarism-prevention tools (e.g. the "turnitin" service) in order to monitor violations during the examination procedures.

I have taken note of the above official notification:**

Ort, Datum
(Place, date)

Unterschrift
(Signature)

****Please be aware that solely the German version of the affidavit ("Eidesstattliche Versicherung") for the Bachelor's/ Master's thesis is the official and legally binding version.**

Inhaltsverzeichnis

1	Motivation und Problemstellung	3
1.1	Zielsetzung und Methodik	4
2	Stand der Technik	5
3	Mathematische Formulierung	7
3.1	Kurzfassung des Grundprinzips	7
3.2	1D-RCWA	9
3.3	Multilayer und 2D-Formulierung	18
3.3.1	Multilayer	18
3.3.2	2D-RCWA	21
3.3.3	homogene Schicht	25
3.4	Feldzerlegung	25
4	Verifikation und Modelluntersuchung	28
4.1	RCWA-Verifikation	28
4.2	Modellbetrachtung der Feldzerlegung	36
5	Ergebnisse und Ausblick	44
6	Literaturverzeichnis und weitere Quellen	46

1 Motivation und Problemstellung

In den letzten Jahren sind, besonders angetrieben durch die stetige Verkleinerung von Transistoren auf Computerchips, Herstellungsverfahren entwickelt worden, die es möglich machen, Oberflächen mit Strukturgrößen kleiner als 50nm herzustellen. Besonders durch die Planartechnik mit EUV-Belichtung, Molekularstrahlepitaxie und Ionenimplantation ist es deshalb möglich, Strukturen in der Größenordnung der Wellenlänge des sichtbaren Lichtes (400nm - 700nm) und darunter herzustellen. Ferner entstanden Ideen, diese Produktionsmöglichkeiten nicht nur in der Herstellung von Computerchips, sondern beispielsweise auch in der Herstellung von Optiken im Nanometerbereich zu nutzen. Besonderer Vorteil sind hierbei einerseits die Möglichkeiten im Bezug auf Bauform und Größe; man stelle sich bsp. eine Displayfokussierung mit einer konventionellen Linse gegenüber einem Mikrolinsenfilm direkt auf dem Display aufgebracht vor. Auf der anderen Seite treten bei Strukturierungen in der Größenordnung der Wellenlänge des Lichtes Beugungseffekte auf, die man sich durch geschickte Auslegung bsp. zur Strahlfokussierung oder Abbildung zu Nutzen machen kann und die herkömmliche Strahlengangsoptiken nicht erzeugen sowie herkömmliche Optiks Simulationen wie Raytracing nicht simulieren können. Diese Arbeit beschäftigt sich besonders mit der Auslegung und Simulation solcher Optiken. Erklärtes Ziel ist es, eine Berechnungsmethode zu formulieren, mit der das Verhalten des elektromagnetischen Feldes an beliebigen Strukturen in der Größenordnung der Wellenlänge berechnet werden kann. Dieses Modell gilt es zu einem Simulationsprogramm zu implementieren und anhand von Beispielmustern, anderer Simulationsprogramme und Veröffentlichungen zu verifizieren. Hierfür wurde die Rigorous Coupled Wave Analysis (RCWA) ausgewählt, ein Verfahren, dessen Grundkonzept in den achtziger Jahren von den Professoren Gaylord und Moharam [1] vorgestellt wurde.

Dieses Grundkonzept besteht darin, das elektromagnetische Feld im Fourier-Raum zu Beugungsordnungen zu zerlegen und in der zu simulierenden Struktur in Differentialgleichungen zu definieren. Aus diesen Differentialgleichungen ergeben sich dann mittels Eigenwertzerlegung allgemeine periodische Feldlösungen, die an Randflächen mit Hilfe von Stetigkeitsbedingungen explizit angegeben werden. Diese Berechnungsmethode wurde ausgewählt, da sie, anders als übliche in der Optikauslegung verwendete Verfahren wie die FDTD, somit von vornherein analytische Grundlösungen (Beugungsordnungen) für die Felddarstellung definiert. Mit diesem Grundsatz wurde eine erheblich schnellere Simulation mit wenig Speicherbedarf erwartet sowie durch die Analyse der Beugungsordnungen ein gutes Verständnis der Ausgangsfelder von diffraktiven Elementen, beispielsweise im Bezug auf eine Parametervariation der Wellenlänge. Ferner bieten die Beugungsordnungen als Lösung eine mögliche Integration von diffraktiven Elementen in Raytracing.

Hauptproblem der RCWA waren zur Zeit der ersten Veröffentlichungen besonders die daraus entstehenden großen Matrixsysteme, die sowohl in Eigenwerte zerlegt und auch gelöst werden mussten. Dies verursacht auch heute noch große Probleme im Bezug auf Konvergenz und numerische Abweichung. Die genaue Ursache dieser Probleme als auch die mathematische Herleitung der Lösung, die teilweise von anderen Autoren als auch von diesem bereitgestellt wurden, werden in der folgenden Arbeit genauer vorgestellt.

Mit der RCWA ist es möglich, den Einfall einer homogenen ebenen Welle mit

Einfalls- und Polarisationswinkel auf eine beliebig strukturierten 3D-periodischen Zelle zu simulieren. Als Ergebnis entstehen Feldlösungen für das Gebiet über und unter der Struktur. Nun bestehen die meisten Anwendungen allerdings eher aus Strukturen, die sich nicht durch eine einzige solcher periodischen Zellen beschreiben lassen. Die Idee des Autors hierfür war, eine Gesamtstruktur durch mehrerer solcher, entweder gleicher oder verschiedener Zellen, zu beschreiben und das Eingangsfeld beliebiger Form in homogene ebene Wellen zu zerlegen, die RCWA-Simulation für jede verschiedene Zelle durchzuführen und anschließend das Feld wieder zusammenzusetzen.

Grundlegend gliedert sich die Arbeit wie folgt. Im folgenden zweiten Kapitel wird untersucht, inwieweit andere Autoren schon Formulierungen zur RCWA bereitgestellt haben und inwieweit diese als Fundament dieser Arbeit gelten. Ferner soll geklärt werden, ob und wie andere Autoren bei der Auslegung von diffraktiven Optiken vorgehen und besonders welche Verfahren genutzt werden. In Kapitel drei wird nach der kurzen Vorstellung des Grundkonzeptes der RCWA durch die gesamte Formulierung der RCWA geleitet. Das folgende Kapitel vier stellt den Kern dieser Arbeit dar. Im ersten Teil dieses Kapitels soll mit Hilfe verifizierter Simulationsergebnisse eindeutig geklärt werden, inwieweit die RCWA vorher definierten Ansprüchen an Konvergenz und Abweichung genügt, auch im Vergleich mit anderen Rechenmethoden. Im zweiten Teil des vierten Kapitels wird anhand einer realen und publizierten Linsenauslegung untersucht, inwieweit die eben kurz erwähnte Feldsynthese ebenfalls den Ansprüchen an Abweichungen entspricht und wie sie bei der Auslegung von Optiken genutzt werden kann. Im fünften und letzten Kapitel sollen die Kernaussagen und Ergebnisse auf diese Fragestellungen zusammengefasst werden und ein kurzer Überblick über mögliche Erweiterungen dieser Implementierung gegeben werden.

1.1 Zielsetzung und Methodik

Nach diesem kurzen Überblick über die folgende Arbeit werden an dieser Stelle die Ziele und Anforderungen dieser Arbeit konkretisiert. Die gesamten später vorgestellten Herleitungen, von deren Implementierung eine Nutzbarkeit beispielsweise in der Auslegung von diffraktiven Optiken erhofft wird, müssen gewissen Anforderungen genügen. Diese Anforderungen werden an verschiedene Ausgaben des Programmes wie Feldgrafiken in einer Ebene oder die Darstellung von Beugungsordnungen in Abhängigkeit eines Parameters wie der Wellenlänge gestellt. Dies umfasst erstens, dass im Vergleich zu Veröffentlichungen zur RCWA die Simulation zum einen nicht langsamer sowie auch gegen denselben Wert konvergiert. Zweitens wird der Anspruch definiert, dass in grafischen Darstellungen diesselben Artefakte wie bspw. das rasche Abklingen eines Feldes an einer Metalloberfläche oder diesselben Kurvensteigungen und Wendepunkte in parameterabhängigen Ergebnisdiagrammen zu erkennen sind. Drittens wird erwartet, dass besonders im Vergleich mit anderen Berechnungsmethoden eventuell vorhandene Abweichungen begründet werden und gegebenenfalls Lösungskonzepte vorgestellt werden können.

Um diese Ziele zu erreichen, wird wie in Abbildung(1) vorgegangen.

Die Einarbeitung in Wellenoptik sowie die Implementierung mit gleichzeitiger Fehlerkorrektur der eigenen Implementierung geschah im Vorfeld dieser Arbeit. Zum einen ist hierfür zuerst die Verifikation genau dieser Implementierung zur Feldlösung einer periodischen Zelle vonnöten. Diese Verifikation wird

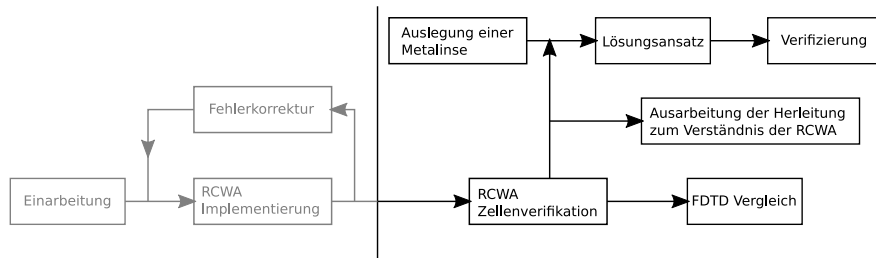


Abbildung 1: Grundlegende Herangehensweise zur Lösung der Problemstellung in dieser Arbeit; die grau hinterlegten Module wurden im Vorfeld dieser Bachelorarbeit behandelt

durch den Vergleich mit einer anderen Berechnungsweise, hier der FDTD, zu einem Vergleich ausgeweitet. Nebenbei werden auch grundlegende Herleitungen der RCWA durch Recherche aufgestellt, um Begründungen für Abweichungen und Feldartefakte geben zu können. Parallel dazu besteht die Fragestellung, ob eine Metalinse mit Hilfe der RCWA berechnet werden kann. Hierzu wird ein Lösungsansatz entwickelt, den es ebenfalls auf seine Richtigkeit und Abweichungen zu überprüfen gilt. Voraussetzung hierfür ist natürlich, dass die verwendeten Simulationsergebnisse einer Zelle verifiziert sind.

2 Stand der Technik

In diesem Kapitel wird ein Überblick gegeben, inwiefern sich schon vorherige Autoren mit der RCWA beschäftigt haben und welche Implementierungen und Lösungen schon verfügbar sind. Dies ist auch nötig, um Zugang zu geeigneten veröffentlichten Referenzergebnissen zu erlangen sowie eigene Formulierungen und Implementierungen auf ihre Richtigkeit zu überprüfen. Ferner muss ermittelt werden, welche Simulationen und Methoden andere Autoren bei der Auslegung ihrer Nanooptiken angewandt haben.

Wie im Eingangskapitel schon erwähnt, wurde das erste Grundkonzept von den Autoren Moharam und Gaylord 1982 [1] veröffentlicht. Eine vollständige Formulierung der RCWA für ein einzelnes Gitter mit einer periodischen Achse (innerhalb der Periode kann entlang der Achse für jeden Punkt ein Brechungsindex definiert werden) wurde 1994 [3] von beiden Autoren veröffentlicht. Im Gegensatz zu vorherigen einfachen Planarformulierungen ermöglicht diese eine komplette 3D-Betrachtung des Feldes und der Struktur mit azimuthalem und polarem Eingangswinkel sowie einem Polarisationswinkel für die lineare Polarisation der einfallenden Welle. Dieses Paper wurde als Startpunkt für die eigene RCWA-Implementierung gewählt. Um mit einem tiefgreifenden Verständnis der Feldtheorie und der RCWA zu beginnen, wurde die Formulierung dieses Papers zuerst in der vereinfachten TE und TM Polarisation implementiert. Anschließend wurden verschiedene Testgitter dieses Papers und anderer Veröffentlichungen nachsimuliert und die Simulationsergebnisse verglichen, um auch ein Verständnis für Konvergenzverhalten und eventuelle Fehlererkennung zu erlangen. Des Weiteren wurde von den gleichen Autoren im selben

Jahr ein Mehrschichtenansatz [4] publiziert. Mit diesem ist es möglich, bei planarem Einfall eine zusätzlich in ihrer Tiefe beliebig strukturierte Schicht durch Übereinanderlegen mehrerer Schichten darzustellen, wobei jede Einzelne wie in [3] beschrieben wird. Das Hauptproblem hierbei besteht in der dabei nötigen Invertierung der Hauptkoeffizientenmatrix, in welcher Matrizen mit Elementen $\exp(k_0 q d)$ mit k_0 als Wellenzahl, q als Eigenwert und d als Schichtdicke zu invertieren sind. Dies ist zwar analytisch möglich, führt jedoch numerisch sehr schnell zu Überläufen. Die Autoren stellten für den Planarfall eine Lösungsidee vor, in welcher diese Invertierung durch geschickte Substitution und anschließendes rückwärtiges Einsetzen nach dem Lösen umgangen werden kann. Diese Idee wurde vom Autor dieser Arbeit auf die allgemeine dreidimensionale Feldbetrachtung sowie auf 2D-periodische Strukturen erweitert, wodurch beliebige Oberflächen abgebildet werden können. Die Formulierung für ein 2D-periodisches Gitter zusammen mit der dazu nötigen Umordnung des Lösungssystems (siehe hierzu Kapitel 3.3) wurde von dem Autor aus [3] hergeleitet.

Bei der Verifikation der RCWA stellten einige Autoren in planaren Fällen mit einer TM-Polarisation der Eingangswelle eine sehr langsame Konvergenz fest, die aus der Synthetisierung der Fourierreihen von Feld und Permittivität herrührt. Dieses Problem war einige Zeit ungelöst, bis Lifeng Li [6] dieses Problem im planaren Fall durch eine reziproke Formulierung der Permittivitätsmatrix und anschließende Invertierung behob und hierzu auch in [6] eine mathematische Beschreibung des Problems erstellte. Im Laufe der Entwicklung der eigenen Simulation wurden die Ergebnisse von weiteren Autoren verglichen, die jedoch hier nicht alle genannt werden.

Diese Simulation reicht nun aus, um sowohl einzelne periodische Zellen als auch ganze periodische Strukturen beliebiger Form simulieren zu können. Es entstand die im vorherigen Kapitel kurz beschriebene Idee der Feldzusammensetzung aus mehreren Zellen, die bei Entdeckung der Veröffentlichung [2] zum ersten Mal angewandt wurde, um eine solche Metalinse mit der RCWA zu simulieren. Dies erschien besonders hierfür geeignet, da bsp. bei ebenem Feldeingang, wofür die Linsen in [2] gebaut wurden, nur sechs Simulationen vonnöten waren und ansonsten lediglich eine 2D-fft in python numpy vollzogen werden musste. Die Autoren dieses Papers simulierten ihre Linsen mit einer FDTD, sodass die Ergebnisse zwischen RCWA-Feldsynthese und FDTD verglichen werden können. Auch erscheint die FDTD als ein sinnvolles Mittel, um Referenzergebnisse mit einer anderen Berechnungsmethode nur für die periodische RCWA-Zelle oder komplett periodische Strukturen zu generieren. Die FDTD als „Ällroundtool“ wird beispielsweise auch für die Simulation von Halbleiterstrukturen und thermischen Simulationen verwendet, für welche die RCWA ebenfalls geeignet ist. Als Software entschied sich der Autor zu Beginn für Optiwave FDTD [7], die gerade als Freeware sehr geeignet erschien. Nach erheblichen Konvergenzproblemen, auf die bei der Vorstellung der Ergebnisse in Kapitel 4 näher eingegangen wird, wurde stattdessen eine 30 Tage Testversion von Lumerical [12] verwendet.

Nach Beginn der Referenzsimulation zu den Metalinsen entdeckte der Autor noch [8], wo eine ähnliche Idee verfolgt wurde, jedoch nur mit einem durchgehenden periodischen Gitter. Hierbei besteht jedoch mit dem Hintergrund des Anwendungsfalles der Metalinse ein Problem: Durch die Felderstellung mit Hilfe der Beugungsordnungen wird keine Phaseninformation des Feldes mitgenommen. Diese ist jedoch gerade bei den Metalinsen mit Strukturgrößen und Perioden unterhalb der Wellenlänge die einzige Möglichkeit der Simulation über das

Ausgangsfeld. Würde man versuchen, die Metalinse mit Beugungsordnungen zu simulieren, erhielte man eine homogene ebene Welle in Normalenrichtung mit einer konstanten Amplitude von näherungsweise 1.

Nach dieser Recherche ist dem Autor keine Veröffentlichungen bezüglich eines direkten Vergleichs der Simulationsergebnisse von FDTD und RCWA in verschiedenen Aspekten bekannt. Ferner wurde auch die Idee der Feldsynthese in keiner Publikation entdeckt.

3 Mathematische Formulierung

3.1 Kurzfassung des Grundprinzips

In diesem Teilabschnitt wird kurz zum Einstieg die Grundidee der RCWA-Simulation vorgestellt, damit der Leser auch ohne die Durcharbeitung des kompletten dritten Kapitels den Vorstellungen von Ergebnissen und Aussagen in Kapitel vier und fünf folgen kann. Hierzu betrachte man die Abbildung(2).

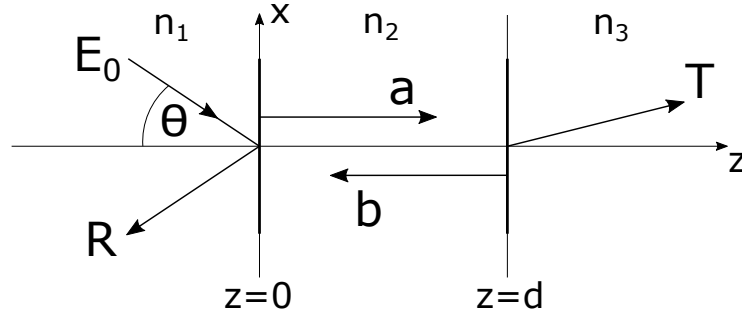


Abbildung 2: Grundprinzip der RCWA: Eine ebene homogene Welle E_0 trifft unter dem Winkel α auf die Grenzfläche zu einem Medium mit Brechungsindex n_2 . Die Subskripte der Brechungsindizes definieren in dieser Skizze die Gebiete

Eine ebene homogene Welle breitet sich in einem homogenen isotropen in negativer z -Richtung unendlich ausgedehnten Medium mit Brechungsindex n_1 aus. Sie trifft unter einem Winkel θ auf eine zur x -Achse parallele Grenzfläche, die ein weiteres homogenes isotropes Medium mit Brechungsindex n_2 berandet. Dieses zweite Medium sei bis zu $z = d$ ausgedehnt, wo sich eine zweite Grenzfläche zu einem in positiver z -Richtung ebenfalls unendlich ausgedehnten Medium mit Brechungsindex n_3 befindet. In dieser Anordnung sind nun die Feldlösungen R und T als Endergebnis gesucht, um das Feld in allen drei Gebieten vollständig angeben zu können. Zuerst betrachtet man die Eingangswelle, die sich gemäß gezeigter Einfallsrichtung in elektrotechnischer Notation folgendermaßen definieren lässt:

$$E_{ein} = E_0 \exp(-jk_{1z}z) \exp(-jk_x x) \quad (1)$$

$$\text{mit } k_{1z} = k_0 n_1 \cos(\theta) \quad \text{und} \quad k_x = k_0 n_1 \sin(\theta)$$

E_0 sei hierbei als skalarer Wert die Amplitude der einfallenden Welle. Die aufgeführten Wellenvektorkomponenten ergeben sich trigonometrisch mit dem Wellenvektor im Medium $k_0 n_1$ der Eingangswelle, wobei $k_0 = 2\pi/\lambda$ und λ der Vakuumwellenlänge der Eingangswelle entspricht. Da die gezeigten Grenzübergänge

allesamt parallel zur x-Achse liegen, ist die angegebene Wellenvektorkomponente k_x aufgrund der allgemein gültigen Stetigkeitsbedingung an Grenzflächen in allen drei Gebieten konstant. Mit der ersten ebenen Grenzfläche an den homogenen Medien lässt sich direkt die gesamte Feldlösung in Gebiet 1 angeben, da wie in Abbildung(2) gezeigt ein Teil der Welle unter dem Ausgangswinkel θ in das Gebiet 1 reflektiert wird und so die Gesamtlösung des Feldes aus einfallender und reflektierter Welle besteht:

$$E_1 = E_0 \exp(-jk_{1z}z) \exp(-jk_x x) + R \exp(jk_{1z}z) \exp(-jk_x x) \quad (2)$$

Die Vorzeichen der Propagationsterme ergeben sich aus Abbildung(2). R beschreibt hierbei die Amplitude der reflektierten Welle. Auf gleiche Weise muss sich in Gebiet 3 eine ebene homogene Welle als Transmission von der Grenzfläche ausbreiten. Auch hierfür lässt sich die vollständige Feldlösung in Gebiet 3 angeben:

$$E_3 = T \exp(-jk_{3z}z) \exp(-jk_x x) \quad (3)$$

Die transmittierte Amplitude T und die Wellenvektorkomponente k_{3z} seien vorerst noch unbekannt. Nun wendet man sich dem begrenzten Gebiet 2 zu. Aufgrund der ebenen Grenzflächen und der homogenen Brechungsindexverteilung in Gebiet 2 kann auch hier als allgemeine Form der Feldlösung eine ebene Welle definiert werden:

$$a = C \exp(-jk_0 \gamma z) \exp(-jk_x x) \quad (4)$$

Hierbei sei C die noch unbekannte Amplitude dieser ebenen Welle. Die z Brechzahlkomponente γ gilt es ebenfalls zu bestimmen. Hierzu zieht man die skalare Wellengleichung im homogenen isotropen Medium heran. Sie wird auf das elektrischen Feld angewandt und in rein ortsabhängiger Form verwendet:

$$\Delta a + k_0^2 n_2^2 a = 0 \quad (5)$$

Das Einsetzen von Gleichung(4) in die Wellengleichung, wobei die Ortsableitung nach y null ist, ergibt als Lösung für γ :

$$\gamma = \pm \sqrt{n_2^2 - n_1^2 \sin(\alpha)^2} \quad (6)$$

Dies zeigt, dass in Gebiet 2 zwei Fundamentallösungen $D \exp(\pm jk_0 \gamma z)$ in Form von ebenen Wellen existieren, wobei die erste in positive und die zweite in negativer z-Richtung propagiert. An dieser Stelle wird eine Konvention getroffen: Die Gesamtlösung in Gebiet zwei ergibt sich aus der Linearkombination beider Fundamentallösungen, die jeweils mit einem noch unbekannten skalaren komplexen Koeffizienten D multipliziert werden. Dieser Koeffizient entspricht der Amplitude und kann in eulerscher Darstellung als Phasorbetrag mit einem konstanten Phasenwinkel aufgefasst werden. Da dieser Koeffizient noch nicht bestimmt ist, kann die z-Variable in jedem der beiden Exponentialterme der Fundamentallösungen durch ein beliebiges $z - p$ substituiert werden, wobei p ein Wert zwischen 0 und d ist. Dies entspricht einem anderen Phasenwinkel bei $z = 0$, der allerdings durch den noch unbekannten Phasenwinkel von D wieder verändert wird. In den späteren Formulierungen des Gitters ist es sinnvoll, hier für die zweite Fundamentallösung z durch $z - d$ zu substituieren, da dann zur

Feldlösung nur eine einzige Exponentialmatrix mit Elementen $\exp(-k_0 q d)$ gebildet werden muss. Somit wird diese Formulierung hier schon eingeführt. Die gesamte Lösung in Gebiet 2 lautet nun:

$$E_2 = A \exp(-jk_0 \gamma z) \exp(-jk_x x) + B \exp(jk_0 \gamma (z - d)) \exp(-jk_x x) \quad (7)$$

mit A und B als Amplituden der beiden ebenen Wellen. Diese Lösungsform entspricht einer stehenden Welle, da direkt erkennbar ist, dass im Lösungsgebiet zwei entgegelaufende Wellen mit gleichen Propagationskonstanten überlagert werden.

Zur vollständigen Kenntnis der Feldlösung in allen drei Gebieten müssen nun noch die unbekannten Amplituden R, T, A und B berechnet werden. Erneut kommt hier die Stetigkeitsbedingung an Grenzflächen zum Einsatz. Die Feldlösungen zweier Gebiete sowie ihre Ableitungen nach z müssen demnach an der nicht metallischen Grenzfläche dazwischen exakt gleich sein. Somit werden die oben angegebenen Lösungen in Gebiet 1 und 2 für $z = 0$ und in Gebiet 2 und 3 für $z = d$ ausformuliert und an jeder Grenzfläche jeweils Feld und Feldableitung nach z gleichgesetzt. So entstehen vier Gleichungen, womit durch Umstellen die vier unbekannten Amplituden gelöst werden können.

Die in den folgenden Teilkapiteln vorgestellten Formulierungen beruhen auf genau dieser Grundidee und erweitern sie, um ebenfalls Feldlösungen in allen drei Gebieten ermitteln zu können, wobei Gebiet 2 jedoch nicht mehr homogen strukturiert ist.

3.2 1D-RCWA

Im nun folgenden Unterkapitel wird Schritt für Schritt durch die Formulierung der RCWA geleitet. Diese Formulierung wurde bis auf einige Modifikationen in ähnlicher Form von den Autoren Moharam und Gaylord in [3] veröffentlicht und wird hier mit einigen Modifikationen und andersartigen Formulierungen aufgeführt.

Als Ausgangspunkt dient das in Abbildung(3) dargestellte Gitterproblem.

Eine ebene, homogene, elektromagnetische Welle mit unendlich ausgedehnter ebener Phasenfront und Vakuumwellenlänge λ trifft aus einem unendlichen homogenen Gebiet mit Brechungsindex n_1 unter einem azimuthalen ϕ und einem polaren Winkel θ auf ein isotropes periodisches Gitter. Dieses Gitter sei ebenfalls unendlich ausgedehnt und durch eine einzelne Zelle beschreibbar, die durch Translation in x- und y-Richtung unendlich oft aneinander gereiht wird. Diese Zelle habe eine Dicke d und eine Periode Λ_x bzw. Λ_y in der jeweilige Koordinatenrichtung. Für jede Gitterachse stehen $2N_x - 1$ bzw. $2N_y - 1$ Punkte zur Verfügung, welche die xy-Ebene der Zelle diskretisieren. Hierbei werden N_x und N_y als die Gesamtzahl an Beugungsordnungen jeder Achse bezeichnet. Die Struktur der Zelle wird durch den Brechungsindex bzw. die Permittivität beschrieben, die für jeden diskretisierten Punkt (Mikrozelle) anzugeben ist und innerhalb dieser Mikrozelle konstant ist. Ferner ändert sich der Brechungsindex entlang der z-Achse über die Dicke d hinweg nicht. Die Permeabilität kann hier ebenfalls auf gleiche Weise dargestellt werden, jedoch können die magnetischen Momente in dem hier betrachteten visuellen optischen Wellenlängenbereich der Feldfrequenz nicht mehr folgen und sie kann somit als Konstante 1 ohne nötige Strukturdarstellung verwendet werden. Das ebenfalls unendlich ausgedehnte

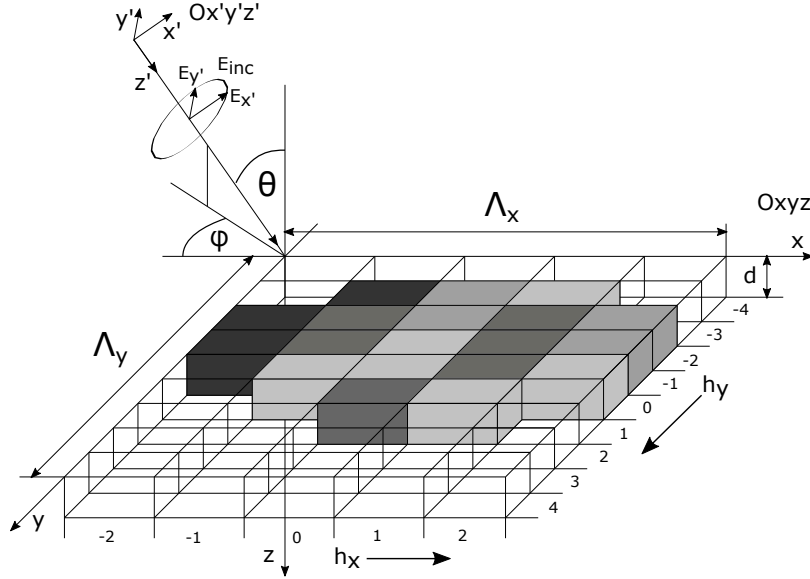


Abbildung 3: Allgemeiner Aufbau des Gitterproblems, wobei eine ebene Welle auf eine Einheitszelle des periodischen Gitters einfällt: Die Grautöne jeder Mikrozele repräsentieren unterschiedliche exemplarische Brechungsindizes, die innerhalb einer Mikrozele an jedem Punkt konstant sind

und homogene Medium unter der Struktur wird durch den Brechungsindex n_3 repräsentiert.

Für die einfallende ebene Welle sei das Koordinatensystem $Ox'y'z'$ definiert. Die Welle breite sich nur in z' Richtung aus, sodass als TEM-Welle ihr Wellenvektor in z' und alle elektrischen und magnetischen Feldkomponenten in der $x'y'$ -Ebene liegen. Die Komponenten des elektrischen Feldvektors bezüglich des Polarisationszustandes lassen sich folgendermaßen darstellen:

$$\mathbf{E}_{inc} = \begin{bmatrix} E_{x'} \\ E_{y'} \end{bmatrix} = \begin{bmatrix} \cos(\psi) x' \\ \sin(\psi) y' e^{j\delta} \end{bmatrix} \exp(-jk_0 n_1 z') \quad \text{mit} \quad k_0 = \frac{2\pi}{\lambda} \quad (8)$$

Hierbei beschreibt der Winkel ψ das Betragsverhältnis der x' zur y' -Komponente des elektrischen Feldes. Der Winkel δ stellt den konstanten Phasennachlauf der y' zur x' -Komponente dar. Somit lässt sich ein elliptischer Polarisationszustand definieren. Einer rein lineare Polarisation entspricht $\delta = 0, \pi, 2\pi, \dots$ mit dem Polarisationswinkel ψ .

Nun gilt es, die $Ox'y'z'$ Darstellung der Welle in das globale Koordinatensystem $Oxyz$ des Gitters zu transformieren. Mit Hilfe der allgemeinen Kugelkoordinatentransformation werden die Koordinaten z' durch \hat{e}_r , x' durch \hat{e}_θ und y' durch \hat{e}_ϕ substituiert und man gelangt zur globalen Darstellung der ebenen Welle:

$$\mathbf{E}_{inc} = \mathbf{e} \exp[-jk_0 n_1 (\sin \theta \cos \phi x + \sin \theta \sin \phi y + \cos \theta z)] \quad (9)$$

mit dem Richtungsvektor der Amplitude aus $(x, y, z) = (z', x', y') * S^T$ mit S

als Rotationsmatrix der Kugelkoordinatentransformation:

$$\mathbf{e} = \begin{bmatrix} \cos \theta \cos \phi \cos \psi - \sin \phi \sin \psi & e^{j\delta} \\ \cos \theta \sin \phi \cos \psi + \cos \phi \sin \psi & e^{j\delta} \\ -\sin \theta \cos \psi & \end{bmatrix}$$

Zu Beginn wird zur Übersichtlichkeit die Gesamtlösung eines eindimensional strukturierten Gitters vorgestellt. Hierbei sei der Brechungsindex in y-Richtung konstant und die in Abbildung(3) gezeigten Ordnungen h_y sowie die Periode Λ_y entfallen. Die Erweiterung auf ein zweidimensionales Gitter wird in Kapitel 3.3 vorgestellt.

Im nun folgenden Abschnitt soll die grundlegende Felddarstellung in allen drei Gebieten aufgezeigt werden. Die nötige Vorgehensweise orientiert sich an der Herleitung in [10] Abschnitt 1.2 und soll hier kurz begründet wiedergegeben werden.

Um die Felddösungen in allen drei Gebieten herzuleiten, betrachtet man an dieser Stelle noch einmal Abbildung(2) als Grundidee. Diese Abbildung wird dahingehend erweitert, dass Gebiet 2 nun nicht mehr homogen ist. Stattdessen sei die Permittivität ϵ für jedes x in z-Richtung über die komplette Dicke d von Gebiet 2 konstant, sodass sie im gesamten Gebiet 2 durch die Funktion $\epsilon(x)$ eindeutig definiert ist. Ferner sei die Permittivität in x-Richtung periodisch mit Λ_x , sodass gilt: $\epsilon(x) = \epsilon(x + \Lambda_x)$

Ferner sei an dieser Stelle angenommen, die Struktur in Gebiet 2 sei rein reflektiv. Man definiert nun die eingehende, ebene Welle:

$$E_i(x, z) = E_0 \exp(-jk_{1z}z) \exp(-jk_{1x}x) \quad (10)$$

Die Wellenvektorkomponenten ergeben sich hierbei auf gleiche Weise wie in Kapitel 3.1. Die gesamte Felddlösung in Gebiet 1 entspricht der Superposition $E_1 = E_i + E_R$ mit E_R als der reflektiven Teillösung. Man erweitere E_R nun folgendermaßen:

$$E_R(x, z) = E_p(x, z) \exp(-jk_{1x}x) \quad \text{mit} \quad E_p(x, z) = E_R(x, z) \exp(jk_{1x}x) \quad (11)$$

An dieser Stelle werde aus weiter unten aufgeführten Gründen die Periodizität $E_p(x + \Lambda_x, z) = E_p(x, z)$ gefordert. Diese Periodizität gilt es nun aufzuzeigen; ein konkreter Beweis auch im Bezug auf die Eindeutigkeit der Lösungen findet sich in [13]. Man formuliert mit Hilfe dieser Periodizität:

$$\begin{aligned} E_R(x + \Lambda_x, z) &= E_p(x + \Lambda_x, z) \exp(-jk_{1x}(x + \Lambda_x)) = \\ &= E_p(x, z) \exp(-jk_{1x}x) \exp(-jk_{1x}\Lambda_x) = E_R(x, z) \exp(-jk_{1x}\Lambda_x) \end{aligned} \quad (12)$$

Die Periodizität lässt sich zeigen, wenn diese Darstellung von E_R unter Ausnutzung der Periodizität einer eindeutigen Felddlösung in Gebiet 1 entspricht. Aufgrund der Phaseeigenschaft einer einzelnen ebenen, homogenen und unendlich ausgedehnten Welle ist ja $E_i(x + \Lambda_x, z) = E_i(x, z) \exp(-jk_{1x}\Lambda_x)$. Das Gebiet 2 wurde vorher als ideal leitendes Metall definiert. In Anwesenheit eines Feldes über einer solchen Grenzfläche entstehen somit auf der Oberfläche des Metalls Flächenladungen und das Feld verschwindet. Somit gilt bei $z=0$ $E_1 = 0$ und es lässt sich gleichsetzen:

$$E_R(x + \Lambda_x, 0) = -E_i(x + \Lambda_x, 0) \quad (13)$$

$$\Rightarrow E_R(x, 0) \exp(-jk_{1x}\Lambda_x) = -E_i(x, 0) \exp(-jk_{1x}\Lambda_x)$$

Daraus folgt nun auch $E_R(x, 0) = -E_i(x, 0)$. Somit ist E_R eine eindeutige Lösung und dementsprechend gilt die Periodizität. Im folgenden wird die Zusammensetzung gezeigt, um die Darstellung des Feldes überzuleiten.

Nach der Fouriertheorie lässt sich jede stückweise stetige Funktion durch eine Fourierreihe annähern. Aufgrund der eben präsentierten Periodizität kann E_p durch eine periodische Fouriertransformation im Intervall $0 \leq x < \Lambda_x$ dargestellt werden:

$$\begin{aligned} E_R(x, z) &= E_p(x, z) e^{-jk_{1x}x} = \left[\sum_{n=-\infty}^{\infty} S_{p,n}(z) e^{2\pi j \frac{n}{\Lambda_x} x} \right] e^{-jk_{1x}x} \quad (14) \\ &= \sum_{n=-\infty}^{\infty} S_{p,n}(z) e^{-jk_{xi}x} \quad \text{mit} \quad k_{xi} = k_{1x} - n \frac{2\pi}{\Lambda_x} \end{aligned}$$

Mit Hilfe dieses Konzeptes lassen sich wieder im allgemeinen Fall bezogen auf Abbildung(3) das elektrische Feld im eindimensionalen Gitterfall über und unter der Struktur darstellen als:

$$\mathbf{E}_1 = \mathbf{E}_{\text{inc}} + \sum_i \mathbf{R}_i \exp[-j(k_{xi}x + k_y y - k_{1,zi}z)] \quad (15)$$

$$\mathbf{E}_3 = \sum_i \mathbf{T}_i \exp(-j[k_{xi}x + k_y y + k_{3,zi}(z - d)]) \quad (16)$$

wobei \mathbf{R}_i bzw. \mathbf{T}_i die komplexen Amplitudenvektoren dieser ebenen Wellen sind. Die Wellenvektoren sind hierbei:

$$\begin{aligned} k_{xi} &= k_0 n_1 \sin \theta \cos \phi - i \frac{2\pi}{\Lambda_x} \quad (17) \\ k_y &= k_0 n_1 \sin \theta \sin \phi \\ k_{med z} &= \begin{cases} +\sqrt{(k_0 n_{med})^2 - k_{xi}^2 - k_y^2} & \text{für } k_0 n_{med} > \sqrt{k_{xi}^2 + k_y^2} \\ -j\sqrt{k_{xi}^2 + k_y^2 - (k_0 n_{med})^2} & \text{für } k_0 n_{med} < \sqrt{k_{xi}^2 + k_y^2} \end{cases} \end{aligned}$$

mit med = 1 bzw 3. und i als Iterator von $-n_x$ bis n_x .

Die z-Komponente als letzte unbekannte Größe ergibt sich aus dem Satz des Pythagoras mit den anderen Komponenten und wurde aus Gründen der Lesbarkeit in eine bedingte imaginäre und reale Formulierung aufgespalten. In Analogie zur Totalreflexion an einer Grenzschicht bedeutet der Übergang von der realen zur imaginären z-Komponente, dass die zugehörige ebene Welle hier in z-Richtung nicht mehr ausbreitungsfähig ist, da sie keinen realen Propagationsterm mehr enthält. Sie entspricht somit nur einem in z-Richtung abklingenden, nicht propagierenden (evaneszenten) Feld.

Da an Grenzflächen aufgrund der Stetigkeitsbedingung tangentielle Feld- und somit auch Wellenvektorkomponenten in beiden Gebieten konstant sein müssen, sind die x und y Wellenvektorkomponenten in allen 3 Gebieten konstant. Somit können auf gleiche Weise elektrisches und magnetisches Feld im Gitterbereich angegeben werden:

$$\mathbf{E}_g = \sum_i [S_{xi}(z) \hat{\mathbf{e}}_x + S_{yi}(z) \hat{\mathbf{e}}_y + S_{zi}(z) \hat{\mathbf{e}}_z] \exp[-j(k_{xi}x + k_y y)] \quad (18)$$

$$\mathbf{H}_g = -j \left(\frac{\epsilon_0}{\mu_0} \right)^{1/2} \sum_i [U_{xi}(z)\hat{\mathbf{e}}_x + U_{yi}(z)\hat{\mathbf{e}}_y + U_{zi}(z)\hat{\mathbf{e}}_z] \exp[-j(k_{xi}x + k_{yi}y)]$$

Hierbei sind die Amplituden S und U skalar und die Vektoren $\hat{\mathbf{e}}$ die Einheitsvektoren der Koordinatenachsen. Der Vorfaktor des H-Feldes entspricht der Verknüpfung von E und H-Feld einer TEM-Welle. Die Abwesenheit eines Propagationstermes in z -Richtung entspricht der Idee, dass das gesamte Gitter in z -Richtung homogen ist (weshalb die Permittivität auch nur in Abhängigkeit von x dargestellt werden kann) und die Propagation des Feldes in z -Richtung wie in Kapitel 3.1 über eine Eigenwertlösung formuliert wird.

Nun nutze man die Maxwellgleichungen in differenzieller Schreibweise im Gitterbereich. Hierbei wird von einer Abwesenheit von Stromdichten ausgegangen sowie dass alle Materialeigenschaften durch die komplexe Permittivität bzw. Permeabilität repräsentiert werden, die somit auch frequenzabhängig (Materialdispersion) sein können:

$$\vec{H} = \left(\frac{j}{\omega\mu} \right) \vec{\nabla} \times \vec{E} \quad \vec{E} = \left(\frac{-j}{\omega\epsilon_0 n^2} \right) \vec{\nabla} \times \vec{H} \quad (19)$$

Man setzt nun die allgemeinen Felddarstellungen aus Gleichung(18) in die Maxwellgleichungen(19) ein und teilt durch die Exponentialterme:

$$\frac{\Delta S_y}{\Delta z} = \frac{k_y}{k_0} \epsilon(x)^{-1} k_{xi} U_y + (k_0 - \frac{k_y}{k_0} \epsilon(x)^{-1} k_y) U_x \quad (20)$$

$$\frac{\Delta S_x}{\Delta z} = (\frac{k_{xi}}{k_0} \epsilon(x)^{-1} k_{xi} - k_0) U_y - \frac{k_{xi}}{k_0} \epsilon(x)^{-1} k_y U_x$$

$$\frac{\Delta U_y}{\Delta z} = \frac{k_y}{k_0} k_{xi} S_y + (k_0 \epsilon(x) - \frac{k_y}{k_0} k_y) S_x$$

$$\frac{\Delta U_x}{\Delta z} = (\frac{k_{xi}}{k_0} k_{xi} - k_0 \epsilon(x)) S_y - \frac{k_{xi}}{k_0} k_y S_x$$

Hierbei ist zu erkennen, dass die Permittivität als Funktion $\epsilon(x)$ auf gleiche Weise periodisch zu einer Fourierreihe zu entwickeln ist, um sie mit den Feldkoeffizienten S und U multiplizieren zu können.

An dieser Stelle sei auf die in Kapitel 2 erwähnte Arbeit von Lifeng Li [6] hingewiesen. In der dort aufgeführten Theorie betrachte man den allgemeinen Term der elektrischen Flussdichte $D_x = \epsilon_0 \epsilon(x) E_x$. In der Herleitung der Wellenvektoren wurde gezeigt, wie und warum das elektrische Feld in eine Fourierreihe zerlegt wird. Ferner ist daran direkt ersichtlich, dass eine Fourierzerlegung von $\epsilon(x)$ aufgrund der äquidistanten Terme zu den gleichen Basisvektoren im Fourierraum führt. Nun wird durch die Diskretisierung der Mikrozellen in Abbildung(3) für $\epsilon(x)$ eine Rechteckstruktur durch eine Fourierreihe angenähert. Diese Reihenentwicklung enthält Überschwinger in den Koeffizienten hoher Ordnungen ("Gibbssches Ringing"). Durch die Synthese der Permittivitätsreihe setzen sich diese Überschwinger auch auf das Feld fort. Li hatte hierfür in [6] eine genaue numerische Definition bezogen auf die Konvergenz bei einer endlichen Anzahl an Reihenordnungen vorgestellt, die hier kurz wiedergegeben werden soll.

Multipliziert man zwei Fourierreihen, die keine gemeinsamen Sprungstellen besitzen, kann die Multiplikation gemäß der Faltung geschrieben werden als:

$$\text{Laurent Umordnung: } D_n = \epsilon_0 \sum_{m=-\infty}^{\infty} \epsilon_{n-m} S_m \quad (21)$$

Die Faltung mit dem Verschiebungsindex $n - m$ kann im mehrdimensionalen Fall als $\epsilon_{n-m} \Rightarrow [[\epsilon]] = E$ geschrieben werden, wobei die doppelte Klammer als allgemeine Konvention die Darstellung einer eindimensionalen Reihe von Koeffizienten als Toeplitzmatrix mit Indexvorschrift $(i, j) = i - j$ für jedes Element beschreibt. So ergibt sich der Vektor aus der Matrixmultiplikation $\mathbf{D} = \mathbf{E}\mathbf{S}$ mit dem Vektor \mathbf{S} .

Haben beide Reihen jedoch gemeinsame Sprungstellen an den gleichen Orten x , so gilt:

$$\text{Li: } D_n = \epsilon_0 \sum_{m=-\infty}^{\infty} \left[\left[\begin{array}{c} 1 \\ \epsilon \end{array} \right] \right]_{n,m}^{-1} S_m \quad (22)$$

Nun lassen sich die Gleichungen (20) als Matrixsystem schreiben, wobei die Terme der Permittivität gemäß der eben vorgestellten Regeln formuliert werden:

$$\begin{bmatrix} \partial \mathbf{S}_y / \partial(k_0 z) \\ \partial \mathbf{S}_x / \partial(k_0 z) \\ \partial \mathbf{U}_y / \partial(k_0 z) \\ \partial \mathbf{U}_x / \partial(k_0 z) \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{K}_y \mathbf{E}^{-1} \mathbf{K}_x & \mathbf{I} - \mathbf{K}_y \mathbf{E}^{-1} \mathbf{K}_y \\ \mathbf{0} & \mathbf{0} & \mathbf{K}_x \mathbf{E}^{-1} \mathbf{K}_x - \mathbf{I} & -\mathbf{K}_x \mathbf{E}^{-1} \mathbf{K}_y \\ \mathbf{K}_y \mathbf{K}_x & \mathbf{E}_{\text{rez}}^{-1} - \mathbf{K}_y^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{K}_x^2 - \mathbf{E} & -\mathbf{K}_x \mathbf{K}_y & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{S}_y \\ \mathbf{S}_x \\ \mathbf{U}_y \\ \mathbf{U}_x \end{bmatrix} \quad (23)$$

wobei die Matrizen \mathbf{K}_x und \mathbf{K}_y Diagonalmatrizen mit Elementen k_{xi}/k_0 bzw. k_{yi}/k_0 sind. \mathbf{E}^{-1} beschreibt hier die Matrixinversion von \mathbf{E} und \mathbf{E}_{rez} die Toeplitzmatrix mit $[[\frac{1}{\epsilon}]]$

Durch erneutes Einsetzen dieser Matrixgleichungen ineinander lassen sich zwei separate Differenzialgleichungssysteme zweiter Ordnungen jeweils für das elektrische und magnetische Feld formulieren:

$$\begin{bmatrix} \partial^2 \mathbf{S}_y / \partial(k_0 z)^2 \\ \partial^2 \mathbf{S}_x / \partial(k_0 z)^2 \end{bmatrix} = \begin{bmatrix} \mathbf{K}_x^2 + \mathbf{D}\mathbf{E} & \mathbf{K}_y [\mathbf{E}^{-1} \mathbf{K}_x \mathbf{E}_{\text{rez}}^{-1} - \mathbf{K}_x] \\ \mathbf{K}_x [\mathbf{E}^{-1} \mathbf{K}_y \mathbf{E} - \mathbf{K}_y] & \mathbf{K}_y^2 + \mathbf{B}\mathbf{E}_{\text{rez}}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{S}_y \\ \mathbf{S}_x \end{bmatrix} \quad (24)$$

$$\begin{bmatrix} \partial^2 \mathbf{U}_y / \partial(k_0 z)^2 \\ \partial^2 \mathbf{U}_x / \partial(k_0 z)^2 \end{bmatrix} = \begin{bmatrix} \mathbf{K}_y^2 + \mathbf{E}_{\text{rez}}^{-1} \mathbf{B} & [\mathbf{K}_x - \mathbf{E}_{\text{rez}}^{-1} \mathbf{K}_x \mathbf{E}^{-1}] \mathbf{K}_y \\ [\mathbf{K}_y - \mathbf{E} \mathbf{K}_y \mathbf{E}^{-1}] \mathbf{K}_x & \mathbf{K}_x^2 + \mathbf{E}\mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{U}_y \\ \mathbf{U}_x \end{bmatrix} \quad (25)$$

mit $\mathbf{D} = \mathbf{K}_y \mathbf{E}^{-1} \mathbf{K}_y - \mathbf{I}$ und $\mathbf{B} = \mathbf{K}_x \mathbf{E}^{-1} \mathbf{K}_x - \mathbf{I}$

Falls die Diagonalmatrix \mathbf{K}_y wie hier im eindimensionalen Fall als ein Skalar multipliziert mit der Identität aufgefasst werden kann, darf sie nach allgemein gültigen Matrizenregeln als neutrales Element durch Multiplikationsterme geschoben werden. So kann zur numerischen Effizienz beim eindimensionalen Gitter zusammengefasst werden:

$$[\partial^2 \mathbf{U}_x / \partial(k_0 z)^2] = [\mathbf{K}_y^2 + \mathbf{A}] [\mathbf{U}_x] \quad (26)$$

$$[\partial^2 \mathbf{S}_x / \partial(k_0 z)^2] = [\mathbf{K}_y^2 + \mathbf{B}\mathbf{E}_{\text{rez}}^{-1}] [\mathbf{S}_x] \quad (27)$$

mit $\mathbf{A} = \mathbf{K}_x^2 - \mathbf{E}$ da sich so die Anzahl der Matrixelemente des Eigenwertproblems um den Faktor 4 reduziert.

Die Superposition der Eigenfunktionen dieser beiden Gleichungen lässt sich in gleicher Weise wie in Kapitel 3.1 als Gesamtlösung angeben:

$$U_{xi}(z) = \sum_{m=1}^n w_{1,i,m} \{ -c_{1,m}^+ \exp(-k_0 q_{1,m} z) + c_{1,m}^- \exp[k_0 q_{1,m} (z - d)] \} \quad (28)$$

$$S_{xi}(z) = \sum_{m=1}^n w_{2,i,m} \{ c_{2,m}^+ \exp(-k_0 q_{2,m} z) + c_{2,m}^- \exp[k_0 q_{2,m} (z - d)] \}$$

wobei q_1 bzw. q_2 die positiven Wurzeln der Eigenwerte aus dem jeweiligen Eigenwertproblem(26) bzw. (27) beschreiben und zu den Diagonalmatrizen \mathbf{Q}_1 bzw. \mathbf{Q}_2 zusammengefasst werden. Analog entsprechen $w_{1,m}$ bzw. $w_{2,m}$ den Eigenvektoren des jeweiligen Problems, die als Spaltenvektoren der Eigenvektormatrizen \mathbf{W}_1 bzw. \mathbf{W}_2 in den folgenden Matrixgleichungen verwendet werden. Die Vorzeichen wurden hierbei bewusst gewählt, um leicht Terme im späteren Gleichsetzen der Felder an den Grenzflächen zusammenfassen zu können. Zur Übersichtlichkeit werden die Gleichungen(28) folgend in Matrixschreibweise formuliert, sodass $\mathbf{U}_x = \mathbf{W}_1 [-\mathbf{c}_1^+ \mathbf{M}_{1z} + \mathbf{c}_1^- \mathbf{M}_{1d}]$ und $\mathbf{S}_x = \mathbf{W}_2 [\mathbf{c}_2^+ \mathbf{M}_{2z} + \mathbf{c}_2^- \mathbf{M}_{2d}]$. Hierbei werden die Exponentialterme in Gleichung(28) zu \mathbf{M}_{1z} bzw. \mathbf{M}_{1d} bzw. für das zweite Set Eigenwerte analog zusammengefasst. Setzt man diese beiden Lösungen in Gleichung(23) ein, so erhält man die Formulierungen der y-Komponenten der Felder im Gitterbereich:

$$\mathbf{S}_y = \mathbf{A}^{-1} \mathbf{W}_1 \mathbf{Q}_1 [\mathbf{c}_1^+ \mathbf{M}_{1z} + \mathbf{c}_1^- \mathbf{M}_{1d}] + \mathbf{A}^{-1} \mathbf{K}_x \mathbf{K}_y \mathbf{W}_2 [\mathbf{c}_2^+ \mathbf{M}_{2z} + \mathbf{c}_2^- \mathbf{M}_{2d}] \quad (29)$$

$$\mathbf{U}_y = \mathbf{B}^{-1} \mathbf{K}_x \mathbf{E}^{-1} \mathbf{K}_y \mathbf{W}_1 [-\mathbf{c}_1^+ \mathbf{M}_{1z} + \mathbf{c}_1^- \mathbf{M}_{1d}] + \mathbf{B}^{-1} \mathbf{W}_2 \mathbf{Q}_2 [-\mathbf{c}_2^+ \mathbf{M}_{2z} + \mathbf{c}_2^- \mathbf{M}_{2d}]$$

Die Terme vor den Matrizen \mathbf{M} werden in exakt der gezeigten Reihenfolge zu \mathbf{V}_{11} , \mathbf{V}_{12} , \mathbf{V}_{21} und \mathbf{V}_{22} zusammengefasst.

Nun hat man allgemeine Lösungen für die ebenen Wellen im Gitterbereich gefunden. Lediglich die komplexen Amplituden c_{12}^\pm dieser Wellen gilt es noch zu ermitteln. Hierfür wird wie in Kapitel 3.1 die Stetigkeitsbedingung der Felder an den Grenzflächen genutzt und die Felder dort gleichgesetzt.

Zuvor wird noch eine geeignete Darstellung der Amplitudenkoeffizienten \mathbf{R} und \mathbf{T} in Gebiet 1 und 3 definiert. Hierzu liegen die jeweiligen Beugungsordnungen i projiziert auf die xy-Ebene in einem azimuthalen Winkel zur x-Achse:

$$\varphi_i = \tan^{-1}(k_y/k_{xi}) \quad (30)$$

Des Weiteren stellt man sich in Analogie zum Ox'y'z' Koordinatensystem der eingehenden ebenen Welle, welches zu Beginn des Kapitels vorgestellt wurde, für die ebene Welle einer jeden Beugungsordnung ein eigenes Koordinatensystem vor. Aufgrund der TEM-Wellennatur im isotropen und homogenen äußeren Halbraum liegen die Feldvektoren von E und H-Feld in einer Ebene senkrecht zur Ausbreitungsrichtung. Diese Ebene lässt sich durch zwei Richtungsvektoren beschreiben, sodass für die Feldamplituden in diesem System nur zwei statt den drei kartesischen Koordinaten angegeben werden müssen:

$$\hat{\mathbf{e}}_s = \frac{\hat{\mathbf{e}}_z \times \mathbf{k}}{|\hat{\mathbf{e}}_z \times \mathbf{k}|} \quad \text{und} \quad \hat{\mathbf{e}}_p = \frac{\hat{\mathbf{e}}_s \times \mathbf{k}}{|\hat{\mathbf{e}}_s \times \mathbf{k}|} \quad (31)$$

wobei im oberen reflektiven Halbraum der \mathbf{k} -Vektor eine negative z -Komponente besitzt. Darin lassen sich die Amplituden der jeweiligen Ordnung wie folgt darstellen:

$$R_{s,i} = \mathbf{R}_i \cdot \hat{\mathbf{e}}_s = \cos \varphi_i R_{yi} - \sin \varphi_i R_{xi} \quad (32)$$

$$R_{p,i} = j n_1 \mathbf{R}_i \cdot \hat{\mathbf{e}}_p = \frac{-j}{k_0} [\cos \varphi_i (k_{1,zi} R_{xi} + k_{xi} R_{zi}) + \sin \varphi_i (k_y R_{zi} + k_{1,zi} R_{yi})]$$

Nun können die Felder der Gebiete 1 und 2 an der oberen Grenzfläche gleichgesetzt werden und dabei die Terme nach s und p -Polarisation zusammengefasst werden:

$$\begin{aligned} \sin \psi \delta_{i0} e^{j\delta} + R_{s,i} &= \cos \varphi_i S_{yi}(0) - \sin \varphi_i S_{xi}(0) \\ -j \sin \psi n_1 \cos \theta \delta_{i0} e^{j\delta} + j(k_{1,zi}/k_0) R_{s,i} &= \cos \varphi_i U_{xi}(0) + \sin \varphi_i U_{yi}(0) \\ j n_I \cos \psi \delta_{i0} + R_{p,i} &= \cos \varphi_i U_{yi}(0) - \sin \varphi_i U_{xi}(0) \\ \cos \psi \cos \theta \delta_{i0} + j[(k_{I,zi}/(k_0 n_1^2)) R_{p,i} &= \cos \varphi_i S_{xi}(0) + \sin \varphi_i S_{yi}(0)] \end{aligned} \quad (33)$$

δ_{i0} beschreibt hierbei das Kroneckerdelta; einen eindimensionalen Vektor, dessen mittleres Element (nullte Ordnung) genau eins und alle weiteren Elemente null sind. So lässt sich die einfallende ebene Welle ebenfalls als Vektor im Matrixsystem darstellen. Welcher resultierende Vektor in diesem System gleichgesetzt wird, lässt sich an der rechten Seite der Gleichungen erkennen. Diese lassen sich zu einem Matrixsystem zusammenfassen:

$$\begin{bmatrix} \sin \psi \delta_{i0} e^{j\delta} \\ -j \sin \psi n_1 \cos \theta \delta_{i0} e^{j\delta} \\ j \cos \psi n_1 \delta_{i0} \\ \cos \psi \cos \theta \delta_{i0} \end{bmatrix} + \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ j \mathbf{Y}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & j \mathbf{Z}_1 \end{bmatrix} \begin{bmatrix} \mathbf{R}_s \\ \mathbf{R}_p \end{bmatrix} = \begin{bmatrix} \mathbf{V}_{ss} & \mathbf{V}_{ss} \mathbf{X}_1 & \mathbf{V}_{sp} & \mathbf{V}_{sp} \mathbf{X}_2 \\ -\mathbf{W}_{ss} & \mathbf{W}_{ss} \mathbf{X}_1 & -\mathbf{W}_{sp} & \mathbf{W}_{sp} \mathbf{X}_2 \\ -\mathbf{W}_{ps} & \mathbf{W}_{ps} \mathbf{X}_1 & -\mathbf{W}_{pp} & \mathbf{W}_{pp} \mathbf{X}_2 \\ \mathbf{V}_{ps} & \mathbf{V}_{ps} \mathbf{X}_1 & \mathbf{V}_{pp} & \mathbf{V}_{pp} \mathbf{X}_2 \end{bmatrix} \begin{bmatrix} \mathbf{c}_1^+ \\ \mathbf{c}_1^- \\ \mathbf{c}_2^+ \\ \mathbf{c}_2^- \end{bmatrix} \quad (34)$$

wobei \mathbf{Y}_1 und \mathbf{Z}_1 Diagonalmatrizen mit den Elementen k_{1z}/k_0 bzw. $k_{1z}/(k_0 n_1^2)$ sind. Die Winkelterme und Eigenvektormatrizen werden zusammengefasst zu:

$$\begin{aligned} \mathbf{V}_{ss} &= \mathbf{F}_c \mathbf{V}_{11} & \mathbf{W}_{ss} &= \mathbf{F}_c \mathbf{W}_1 + \mathbf{F}_s \mathbf{V}_{21} \\ \mathbf{V}_{sp} &= \mathbf{F}_c \mathbf{V}_{12} - \mathbf{F}_c \mathbf{W}_2 & \mathbf{W}_{sp} &= \mathbf{F}_s \mathbf{V}_{22} \\ \mathbf{W}_{pp} &= \mathbf{F}_c \mathbf{V}_{22} & \mathbf{V}_{pp} &= \mathbf{F}_c \mathbf{W}_2 + \mathbf{F}_s \mathbf{V}_{12} \\ \mathbf{W}_{ps} &= \mathbf{F}_c \mathbf{V}_{21} - \mathbf{F}_s \mathbf{W}_1 & \mathbf{V}_{ps} &= \mathbf{F}_s \mathbf{V}_{11} \end{aligned} \quad (35)$$

wobei \mathbf{F}_c eine Diagonalmatrix mit Elementen $\cos \phi_i$ bzw. \mathbf{F}_s mit $\sin \phi_i$ und die Diagonalmatrizen $\mathbf{X}_1 = \exp(-k_0 q_{1,m} d)$ und $\mathbf{X}_2 = \exp(-k_0 q_{2,m} d)$ darstellen. Analog werden an der zweiten Grenzfläche die Amplituden in sp -Form angegeben:

$$\begin{aligned} T_{s,i} &= \cos \varphi_i T_{yi} - \sin \varphi_i T_{xi} \\ T_{p,i} &= (j/k_0) [\cos \varphi_i (k_{3,zi} T_{xi} - k_{xi} T_{zi}) - \sin \varphi_i (k_y T_{zi} - k_{3,zi} T_{yi})] \end{aligned} \quad (36)$$

und die Felder gleichgesetzt:

$$\cos \varphi_i S_{yi}(d) - \sin \varphi_i S_{xi}(d) = T_{si} \quad (37)$$

$$\begin{aligned}
\cos \varphi_i U_{xi}(d) + \sin \varphi_i U_{yi}(d) &= -j(k_{3,zi}/k_0)T_{si} \\
\cos \varphi_i U_{yi}(d) - \sin \varphi_i U_{xi}(d) &= T_{pi} \\
\cos \varphi_i S_{xi}(d) + \sin \varphi_i S_{yi}(d) &= -j(k_{3,zi}/k_0 n_1^2)T_{pi}
\end{aligned}$$

sodass das untere Grenzflächensystem lautet:

$$\begin{bmatrix} \mathbf{V}_{ss}\mathbf{X}_1 & \mathbf{V}_{ss} & \mathbf{V}_{sp}\mathbf{X}_2 & \mathbf{V}_{sp} \\ -\mathbf{W}_{ss}\mathbf{X}_1 & \mathbf{W}_{ss} & -\mathbf{W}_{sp}\mathbf{X}_2 & \mathbf{W}_{sp} \\ -\mathbf{W}_{ps}\mathbf{X}_1 & \mathbf{W}_{ps} & -\mathbf{W}_{pp}\mathbf{X}_2 & \mathbf{W}_{pp} \\ \mathbf{V}_{ps}\mathbf{X}_1 & \mathbf{V}_{ps} & \mathbf{V}_{pp}\mathbf{X}_2 & \mathbf{V}_{pp} \end{bmatrix} \begin{bmatrix} \mathbf{c}_1^+ \\ \mathbf{c}_1^- \\ \mathbf{c}_2^+ \\ \mathbf{c}_2^- \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -j\mathbf{Y}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & -j\mathbf{Z}_3 \end{bmatrix} \begin{bmatrix} \mathbf{T}_s \\ \mathbf{T}_p \end{bmatrix} \quad (38)$$

wobei \mathbf{Y}_3 und \mathbf{Z}_3 Diagonalmatrizen mit den Elementen k_{3z}/k_0 bzw. $k_{3z}/(k_0 n_3^2)$ sind.

Beide Grenzflächengleichungen lassen sich zu einem Matrixsystem umordnen, das eindeutig nach den c-Amplituden gelöst werden kann, indem die R und T Vektoren ineinander eingesetzt werden. Anschließend erfolgt rückwärtiges Einsetzen der c-Koeffizienten in die Gleichungen(34) bzw. (38), um die Amplituden R und T zu ermitteln.

Nun betrachte man die allgemeine Definition der Flächenleistungsdichte (Poyntingvektor S) für TEM-Wellen im Vakuum mit $|S| = \frac{|E|^2}{Z_0}$, wobei Z_0 der Wellenwiderstand des Vakuums ist. Die auf das Gitter einfallende Welle hat eine Leistungsdichte von 1, die über ein infinitesimal kleines Flächenstück normal zum Wellenvektor definiert ist. Betrachtet wird allerdings ein Flächenstück in der xy-Ebene, weshalb die Leistungsdichte auf die z-Komponente der Ausbreitungsrichtung (somit der Realteil einer Wellenvektorkomponente) mit Hilfe des Skalarproduktes projiziert werden muss. Dieser ergibt sich für die einfallende Welle zu $k_0 n_1 \cos(\theta)$. Man bildet nun den Quotienten beider Leistungsdichten zu:

$$\begin{aligned}
DE_{ri} &= \left(|R_{s,i}|^2 + \frac{1}{n_1^2} |R_{p,i}|^2 \right) \operatorname{Re} \left(\frac{k_{1,zi}}{k_0 n_1 \cos \theta} \right) \\
DE_{ti} &= \left(|T_{s,i}|^2 + \frac{1}{n_3^2} |T_{p,i}|^2 \right) \operatorname{Re} \left(\frac{k_{3,zi}}{k_0 n_1 \cos \theta} \right)
\end{aligned} \quad (39)$$

Dieser Quotient wird allgemein als Beugungseffizienz bezeichnet und beschreibt den Anteil der Leistungsdichte von Beugungsordnung i bezogen auf die gesamte Leistungsdichte der einfallenden Welle.

Für verlustfreie Materialien muss die Energieerhaltung über alle Beugungsordnungen gelten:

$$\sum_i (DE_{ri} + DE_{ti}) = 1 \quad (40)$$

Auch bei verlustbehafteten Materialien (bsp. Gold bei 400nm Wellenlänge) ist die Gesamteffizienz eine gute Orientierung für die Konvergenz abhängig von der Anzahl der verwendeten Beugungsordnungen sowie für die allgemeine Richtigkeit und numerische Abweichung des Codes. Hierauf wird in Kapitel 4 näher eingegangen.

In Kapitel 4 werden ferner einige Felddarstellungen durch die Struktur gezeigt. Hierfür muss zusätzlich von die z-Komponente der Felder in der Gitterregion ermittelt werden, was mit Hilfe von Maxwell (19) geschieht:

$$\mathbf{U}_z(z) = j\mathbf{K}_x \mathbf{S}_y(z) - j\mathbf{K}_y \mathbf{S}_x(z) \quad (41)$$

$$\mathbf{S}_z(z) = \mathbf{E}^{-1}j(\mathbf{K}_x \mathbf{U}_y(z) - \mathbf{K}_y \mathbf{U}_x(z))$$

Hierzu wird eine beliebig orientiertes ebenes Samplinggitter durch die Struktur gelegt und die x,y und z-Komponenten der Felder explizit berechnet. Hierbei muss in einem Mehrschichtensystem, dessen Formulierung im folgenden Teilkapitel vorgestellt wird, genau beachtet werden, in welchem Schicht man sich im aktuellen Gitterpunkt befindet.

3.3 Multilayer und 2D-Formulierung

Im vorhergehenden Abschnitt wurde die Berechnung der Feldlösung an einer Struktur vorgestellt, die als Brechungsindexvariation entlang einer Achse repräsentiert werden kann. Diese Formulierung wird nun auf die verbleibenden beiden Achsen erweitert. Das Konzept für die Strukturierung in der z-Achse ist als "multilayer approach" bekannt. Hierbei wird die Struktur in z-Richtung diskretisiert, indem sie entlang der z-Richtung aus Schichten aufgebaut wird. Durch Hinzunahme einer zweiten periodischen Achse in y-Richtung können auch Strukturveränderungen in y-Richtung betrachtet werden; somit wird die im vorigen Kapitel vorgestellte 1D-Lösung zu einer 2D-Formulierung erweitert. Diese beiden Methoden werden im folgenden Kapitel vorgestellt.

Mit Hilfe dieser beiden Erweiterungen kann eine beliebig strukturierte dreidimensionale Zelle mit Ausdehnung in allen drei Achsen simuliert werden.

3.3.1 Multilayer

Die Grundidee dieser Erweiterung wurde von den Autoren Moharam und Gaylord in [4] vorgestellt. Diese Veröffentlichung bezieht sich auf einen planaren Gitterfall. Die nun folgende Beschreibung orientiert sich an der Formulierung in [4], unterscheidet sich jedoch an einigen Punkten und erweitert sie zu einem allgemeinen Fall mit drei Eingangswinkeln, wie er schon im vorherigen Kapitel beschrieben wurde.

Zu Beginn werde die Abbildung(4) betrachtet.

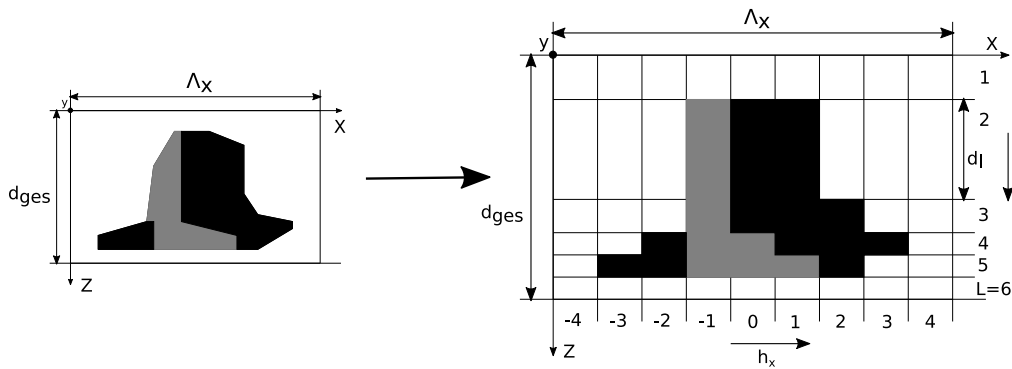


Abbildung 4: Grundkonzept des Multilayeransatzes: eine Struktur wird mit Hilfe von 9 Fourierordnungen ($N_x = 5$) und 6 Layern diskretisiert; die beiden Graufärbungen repräsentieren hierbei jeweils einen Brechungsindex

Die Dicke d_l kann hierbei für jedes einzelne Layer beliebig je nach Struktur festgelegt werden. Innerhalb wird der Brechungsindex auf exakt selbe Weise wie

im Kapitel zuvor angegeben, wobei der Brechungsindex innerhalb einer Schicht konstant in z-Richtung ist. Diese Darstellung ist wie in der Abbildung(4) ersichtlich nur eine Näherung, weshalb auch genau definiert werden muss, welcher Brechungsindex innerhalb eines Blockelementes (h_x, l) herrscht, falls beispielsweise eine Grenzfläche zwischen zwei Materialien quer durch einen Block hindurchgeht. Daraus ist direkt ersichtlich, dass für jede einzelne Schicht allein die gleiche Formulierung wie im Kapitel zuvor vollzogen werden kann sich so für jede Schicht jeweils eine eigene rechte Seite von Gleichung(34) und eine eigene linke Seite von Gleichung(38) ergibt. Diese Seiten der Gleichung sind dementsprechend allerdings individuell, weshalb sich im folgenden alle Elemente \mathbf{V} , \mathbf{W} und \mathbf{X} auf ihre jeweilige Schicht beziehen und deshalb den Index l erhalten. Schreibt man nun die Gleichungen(34) und (38) in dieser Indexschreibweise an jeder Grenzfläche aus, so ist direkt ersichtlich, dass durch Matrixinversion die c_l Koeffizienten ineinander eingesetzt werden können. Beginnt man nun an der letzten Grenzfläche zwischen Schicht L und den unteren Halbraum und iteriert das Einsetzen bis zur obersten Grenzfläche, so gelangt man zu folgender Zusammenfassung:

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} + \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ j\mathbf{Y}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & j\mathbf{Z}_1 \end{bmatrix} \begin{bmatrix} \mathbf{R}_s \\ \mathbf{R}_p \end{bmatrix} = \left[\prod_{l=1}^L \begin{bmatrix} \mathbf{V}_{ssl} & \mathbf{V}_{ssl}\mathbf{X}_{1l} & \mathbf{V}_{spl} & \mathbf{V}_{spl}\mathbf{X}_{2l} \\ -\mathbf{W}_{ssl} & \mathbf{W}_{ssl}\mathbf{X}_{1l} & -\mathbf{W}_{spl} & \mathbf{W}_{spl}\mathbf{X}_{2l} \\ -\mathbf{W}_{psl} & \mathbf{W}_{psl}\mathbf{X}_{1l} & -\mathbf{W}_{ppl} & \mathbf{W}_{ppl}\mathbf{X}_{2l} \\ \mathbf{V}_{psl} & \mathbf{V}_{psl}\mathbf{X}_{1l} & \mathbf{V}_{ppl} & \mathbf{V}_{ppl}\mathbf{X}_{2l} \end{bmatrix} \right]^{-1} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -j\mathbf{Y}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & -j\mathbf{Z}_3 \end{bmatrix} \begin{bmatrix} \mathbf{T}_s \\ \mathbf{T}_p \end{bmatrix} \quad (42)$$

Die Eingangswinkelmatrix aus Gleichung(34) wird hier mit den Elementen α bezeichnet. Dieses Matrixsystem lässt sich analytisch nach den Koeffizienten \mathbf{R} und \mathbf{T} lösen. Real entsteht hierbei jedoch ein grundsätzliches numerisches Problem:

In Gleichung(42) werden die gezeigten $L \times 4 \times 4$ Blockmatrizen invertiert und miteinander multipliziert. Sie enthalten jedoch Terme der Form $\exp(-k_0 q d)$, deren Invertierung und kumulative Multiplikation numerisch Terme der Größenordnung e^{-150} , im schlimmsten Fall sogar overflows, erzeugen. Dieses Problem war schon den Autoren Gaylord und Moharam bekannt, die in [4] eine Idee zur Lösung dieser Problems präsentierten, indem sie in ihrem planaren Multilayerfall durch eine geschickte Substitution und späterem rückwärtigen Einsetzen diese Invertierung umgingen. Sie präsentierten hierfür zwei Ansätze: Erstens den partial approach, der sich numerisch besonders für entweder hochreflektive oder hochtransmittive Gitter eignet. Zweitens den full approach, mit dem sowohl das reflektierte als auch das transmittierte Feld berechnet werden kann. Beide Ansätze wurden von diesem Autor unabhängig von und zeitlich vor dieser Arbeit auf den dreidimensionalen Einfall sowie für zweidimensionale Gitter erweitert. Im Folgenden wird nur der full approach vorgestellt:

Beginnend mit der Gleichung an der letzten Grenzfläche werden die Exponen-

zialmatrizen \mathbf{X} von der zu invertierenden Hauptkoeffizientenmatrix isoliert:

$$\begin{aligned} \mathbf{H}\mathbf{K}\mathbf{R}_l & \begin{bmatrix} \mathbf{X}_{1l} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{X}_{2l} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{V}_{ssl} & \mathbf{V}_{ssl} & \mathbf{V}_{spl} & \mathbf{V}_{spl} \\ -\mathbf{W}_{ssl} & \mathbf{W}_{ssl} & -\mathbf{W}_{spl} & \mathbf{W}_{spl} \\ -\mathbf{W}_{psl} & \mathbf{W}_{psl} & -\mathbf{W}_{ppl} & \mathbf{W}_{ppl} \\ \mathbf{V}_{psl} & \mathbf{V}_{psl} & \mathbf{V}_{ppl} & \mathbf{V}_{ppl} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{f}_{1,L+1} & \mathbf{f}_{2,L+1} \\ \mathbf{g}_{1,L+1} & \mathbf{g}_{2,L+1} \\ \mathbf{h}_{1,L+1} & \mathbf{h}_{2,L+1} \\ \mathbf{i}_{1,L+1} & \mathbf{i}_{2,L+1} \end{bmatrix} \begin{bmatrix} \mathbf{T}_s \\ \mathbf{T}_p \end{bmatrix} \\ & = \mathbf{H}\mathbf{K}\mathbf{R}_l \begin{bmatrix} \mathbf{X}_{1l} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{X}_{2l} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{a}_{1L} & \mathbf{a}_{2L} \\ \mathbf{b}_{1L} & \mathbf{b}_{2L} \\ \mathbf{d}_{1L} & \mathbf{d}_{2L} \\ \mathbf{e}_{1L} & \mathbf{e}_{2L} \end{bmatrix} \begin{bmatrix} \mathbf{T}_s \\ \mathbf{T}_p \end{bmatrix} = \mathbf{H}\mathbf{K}\mathbf{R}_l \begin{bmatrix} \mathbf{X}_{1L}^{-1}(\mathbf{a}_{1L}\mathbf{T}_s + \mathbf{a}_{2L}\mathbf{T}_p) \\ \mathbf{b}_{1L}\mathbf{T}_s + \mathbf{b}_{2L}\mathbf{T}_p \\ \mathbf{X}_{2L}^{-1}(\mathbf{d}_{1L}\mathbf{T}_s + \mathbf{d}_{2L}\mathbf{T}_p) \\ \mathbf{e}_{1L}\mathbf{T}_s + \mathbf{e}_{2L}\mathbf{T}_p \end{bmatrix} \end{aligned} \quad (43)$$

Hierbei bezeichnet $\mathbf{H}\mathbf{K}\mathbf{R}_L$ der 4×4 Blockmatrix der rechten Seite von Gleichung(34). Ferner ist in diesem ersten Schritt $l = L$. Die Matrix mit Elementen $\mathbf{f}_{L+1}, \mathbf{g}_{L+1}$ entspricht der Matrix mit den Elementen \mathbf{Y}_3 und \mathbf{Z}_3 auf der rechten Seite in Gleichung(38).

Nun führt man neue Transmissionskoeffizienten $\mathbf{T}_{sl}, \mathbf{T}_{pl}$ für jede Schicht l ein. Sie ergeben sich durch die Substitutionen

$$\begin{bmatrix} \mathbf{x}_L & \mathbf{y}_L \end{bmatrix} = \begin{bmatrix} \mathbf{b}_{1L} & \mathbf{b}_{2L} \end{bmatrix} \begin{bmatrix} \mathbf{a}_{1L} & \mathbf{a}_{2L} \\ \mathbf{d}_{1L} & \mathbf{d}_{2L} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{X}_{1L} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_{2L} \end{bmatrix} \quad (44)$$

Analog wird $\begin{bmatrix} \mathbf{s}_L & \mathbf{r}_L \end{bmatrix}$ mit $\begin{bmatrix} \mathbf{e}_{1L} & \mathbf{e}_{2L} \end{bmatrix}$ statt den Termen \mathbf{b}_L in Gleichung(44) berechnet. So wurde eine Berechnung der Transmissionskoeffizienten ohne die Invertierung der Matrizen \mathbf{X} gefunden. So ergibt sich die Gleichung(43) des letzten Layers weiter zu

$$= \mathbf{H}\mathbf{K}\mathbf{R}_l \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{x}_L & \mathbf{y}_L \\ \mathbf{0} & \mathbf{I} \\ \mathbf{s}_L & \mathbf{r}_L \end{bmatrix} \begin{bmatrix} \mathbf{T}_{sL} \\ \mathbf{T}_{pL} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_{1L} & \mathbf{f}_{2L} \\ \mathbf{g}_{1L} & \mathbf{g}_{2L} \\ \mathbf{h}_{1L} & \mathbf{h}_{2L} \\ \mathbf{i}_{1L} & \mathbf{i}_{2L} \end{bmatrix} \begin{bmatrix} \mathbf{T}_{sL} \\ \mathbf{T}_{pL} \end{bmatrix} \quad (45)$$

Die Koeffizienten \mathbf{T}_L mit Index sind hierbei reine Substitutionen und keine physikalischen Amplituden. Diese Prozedur wird nun beginnend mit Gleichung(43) für das darüberliegende Layer $l = L - 1$ wiederholt, sodass nun in (43) der Index l nun durch $L-1$ und $L+1$ durch L ersetzt ist. Diese Prozedur wird nun für alle weiteren Layer iterativ bis zum ersten Layer wiederholt. Bei $l=1$ ergibt sich somit das lösbare System

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} + \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ j\mathbf{Y}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{0} & j\mathbf{Z}_1 \end{bmatrix} \begin{bmatrix} \mathbf{R}_s \\ \mathbf{R}_p \end{bmatrix} = \begin{bmatrix} \mathbf{f}_{11} & \mathbf{f}_{21} \\ \mathbf{g}_{11} & \mathbf{g}_{21} \\ \mathbf{h}_{11} & \mathbf{h}_{21} \\ \mathbf{i}_{11} & \mathbf{i}_{21} \end{bmatrix} \begin{bmatrix} \mathbf{T}_{s1} \\ \mathbf{T}_{p1} \end{bmatrix} \quad (46)$$

welches direkt nach den Koeffizienten \mathbf{R} und \mathbf{T} gelöst werden kann. Für die Berechnung der Transmissionkoeffizienten in Gebiet drei muss nun die gesamte Substitution der T-Koeffizienten der einzelnen Layer rückgängig gemacht werden. Hierzu wird iterativ eingesetzt

$$\begin{bmatrix} \mathbf{T}_{s,l+1} \\ \mathbf{T}_{p,l+1} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_{1l} & \mathbf{a}_{2l} \\ \mathbf{d}_{1l} & \mathbf{d}_{2l} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{X}_{1l} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_{2l} \end{bmatrix} \begin{bmatrix} \mathbf{T}_{s,l} \\ \mathbf{T}_{p,l} \end{bmatrix} \quad (47)$$

3.3.2 2D-RCWA

Im Folgenden wird nun die 2D-Formulierung der RCWA vorgestellt. Auch die Lösung dieses Problems ist in einigen Veröffentlichungen wie bspw. [10] zu finden, jedoch zum Teil mit anderen Konventionen und Formulierungen als den nun folgenden. Hierzu wird eine zweite periodische Achse in y-Richtung eingeführt. Es ist auch möglich, dass die zweite Zellenachse nicht im rechten Winkel zur ersten x-Achse steht oder das gesamte Zellkoordinatensystem zu einem globalen System gedreht ist. Dieser Fall wird in einigen Veröffentlichungen behandelt, jedoch hier nicht weiter betrachtet.

Die zweite Zellenachse entspricht der y-Achse mit einer eigenen Periode Λ_y . Da in der zweiten Achsenrichtung ebenfalls Beugungsordnungen diskretisiert werden, wird die Gesamtzahl der y-Ordnungen mit N_y bezeichnet. N_x und N_y können beide je nach Komplexität der Struktur unterschiedlich groß gewählt werden.

Somit ergibt sich für die y-Komponente des Wellenvektors eine neue Formulierung analog zur x-Komponente in Gleichung(17):

$$k_{yi} = k_0 n_1 \sin \theta \sin \phi - i_y \frac{2\pi}{\Lambda_y} \quad (48)$$

Die anderen Wellenkomponenten bleiben in gleicher Berechnungsweise erhalten. Grundsätzlich ergeben sich durch die zwei Ordnungsvariablen zweidimensionale Feldlösungen. Hierbei wird der erste Index für die x-Ordnung, der zweite für die y-Ordnung festgelegt:

$$\mathbf{S}_x = (S_{x,-i_x,-i_y}) \quad (49)$$

wobei die Laufvariablen $-n_x \dots i_x \dots n_x$ mit $n_x = N_x // 2$ und $-n_y \dots i_y \dots n_y$ mit $n_y = N_y // 2$ definiert sind. Der Operator $//$ entspricht hierbei einer Ganzzahldivision, da allgemein die Gesamtzahl der Beugungsordnungen wie im 1D-Fall beschrieben ungerade ist.

Zur Lösbarkeit des Systems als zweidimensionales Matrixsystem ist nun eine Umordnung der Indizes notwendig, in welcher diese Feldlösungen durch einen einzelnen Index indiziert und als Vektor geschrieben werden können. Dieses Umordnungsprinzip bzw. die in dieser Arbeit gewählte Konvention wird beispielsweise an den beiden Diagonalmatrizen \mathbf{K}_x und \mathbf{K}_y vorgestellt.

$$k_{x,i_x,i_y} \Rightarrow \begin{bmatrix} \begin{bmatrix} k_{x,-n_x} \\ k_{x,-n_x} \\ \vdots \\ k_{x,-n_x} \end{bmatrix} \\ \begin{bmatrix} k_{x, -(n_x-1)} \\ k_{x, -(n_x-1)} \\ \vdots \\ k_{x, -(n_x-1)} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} k_{x,n_x} \\ k_{x,n_x} \\ \vdots \\ k_{x,n_x} \end{bmatrix} \end{bmatrix} \quad k_{y,n_x,n_y} \Rightarrow \begin{bmatrix} \begin{bmatrix} k_{y,-n_y} \\ k_{y, -(n_y-1)} \\ \vdots \\ k_{y,n_y} \end{bmatrix} \\ \begin{bmatrix} k_{y,-n_y} \\ k_{y, -(n_y-1)} \\ \vdots \\ k_{y,n_y} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} k_{y,-n_y} \\ k_{y, -(n_y-1)} \\ \vdots \\ k_{y,n_y} \end{bmatrix} \end{bmatrix} \quad (50)$$

Die gleiche Umsortierung ist ebenfalls für die Erstellung der Permittivitätsmatrizen \mathbf{E} , \mathbf{E}_{rez}^{-1} und \mathbf{E}^{-1} nötig. Hierzu betrachte man die Abbildung(5).

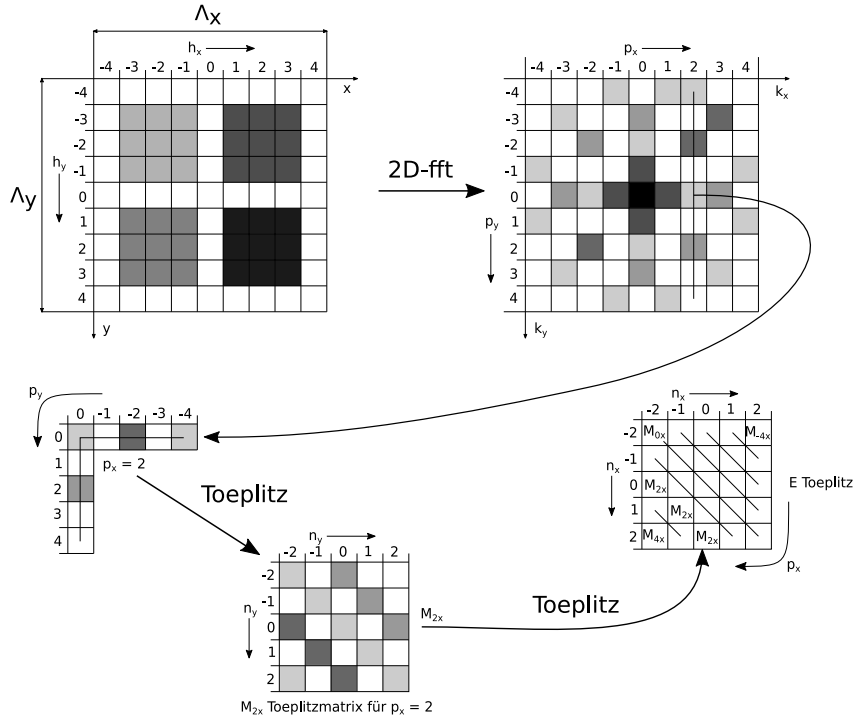


Abbildung 5: Erstellung der Permittivitätsmatrizen entsprechend der vorgestellten Umordnungskonvention; in diesem Schaubild ist exemplarisch $N_x = N_y = 5$

Die Erstellung der Matrix \mathbf{E} erfolgt hierbei entlang der Pfeilrichtung. Man ha-

be eine periodische Zelle mit Gitterperioden Λ_x und Λ_y , wobei in den Achsen Indizes h_x bzw. h_y mit hier gleichem Indexbereich von $-(2N_x - 1)/2$ bis $(2N_x - 1)/2$ die Struktur diskretisieren. Im ersten Schritt definieren die Graufärbungen jeweils den Brechungsindex ihres Quadrates. Diese Struktur zweidimensional fouriertransformiert, wobei der resultierende Fourierkoeffizient der Ordnung (0,0) exakt in der Mitte positioniert ist. Diese Fourierkoeffizienten werden durch die Indizes p_x bzw. p_y mit gleichem Wertebereich wie h_x iteriert. Nun wird in dieser Konvention jede einzelne Spalte zu einer eigenen Toeplitzmatrix entsprechend der bekannten Toeplitzindizierung aus dem eindimensionalen RCWA-Fall. Nun stelle man sich eine weitere Matrix der Dimension $N_x N_y \times N_x N_y$ vor, wobei die Toeplitzmatrizen p_x wiederum gemäß der Toeplitzindizierung in diese positioniert werden. Nach dieser Indexvorschrift entstehen die drei Matrizen \mathbf{E} , \mathbf{E}_{rez}^{-1} und \mathbf{E}^{-1} .

Nun beginnt man bei den Matrizen der Urformulierung (Gleichungen(24) und (25)), die auf gleiche Weise mit den umsortierten Matrizen bestückt werden. Hierbei muss jedoch die reziproke Formulierung aufgrund der Periodizität beider Achsen auch auf die y-Fourierreihensynthese ausgeweitet werden, sodass alle vorhandenen Matrizen \mathbf{E} , jedoch nicht \mathbf{E}^{-1} , durch die reziproken \mathbf{E}_{rez}^{-1} ersetzt werden müssen. Somit ist auch die Vereinfachung zu Gleichung(26) und (27) nicht möglich, wodurch die dort genannte Optimierung der Rechenzeit nicht möglich ist.

Man kann sich nun für die Eigenwertzerlegung eines dieser beiden Gleichungssysteme entscheiden. Beispielsweise ergibt die Lösung nach den S-Feldelementen aus Gleichung(26):

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_y \\ \mathbf{S}_x \end{bmatrix} = \begin{bmatrix} \mathbf{W}_y \\ \mathbf{W}_x \end{bmatrix} \times [\mathbf{c}^+ \exp(-k_0 \mathbf{q} z) + \mathbf{c}^- \exp[k_0 \mathbf{q}(z - d)]] \quad (51)$$

Hierbei entspricht \times einer Matrixmultiplikation. Die Elemente \mathbf{W} sind hierbei die Eigenvektormatrix und \mathbf{q} der Vektor bestehend aus den positiven Wurzeln der Eigenwerte. Der Exponentialterm beschreibt hierbei einen Vektor, der aus der elementweisen Exponentenbildung des Termes $\pm k_0 \mathbf{q} z$ entsteht. Ferner werden der resultierende Exponentenvektor $\exp(\pm k_0 \mathbf{q} z)$ und der Vektor \mathbf{c}^\pm ebenfalls elementweise miteinander multipliziert.

Mit Hilfe der einfachen Ableitungen aus Gleichung(23) erhält man die komplementären Elemente des H-Feldes:

$$\mathbf{U}_y = \mathbf{K}_y \mathbf{K}_x \mathbf{W}_y \mathbf{Q}^{-1} \mathbf{c}_{res} + (\mathbf{E}_{rez}^{-1} - \mathbf{K}_y^2) \mathbf{W}_x \mathbf{Q}^{-1} \mathbf{c}_{res} \quad (52)$$

mit $\mathbf{c}_{res} = [\mathbf{c}^- \exp(k_0 \mathbf{q}(z - d)) - \mathbf{c}^+ \exp(-k_0 \mathbf{q} z)]$

$$\mathbf{U}_x = (\mathbf{K}_x^2 - \mathbf{E}_{rez}^{-1}) \mathbf{W}_y \mathbf{Q}^{-1} \mathbf{c}_{res} - \mathbf{K}_x \mathbf{K}_y \mathbf{W}_x \mathbf{Q}^{-1} \mathbf{c}_{res} \quad (53)$$

\mathbf{Q} ist hierbei die Diagonalmatrix mit den Elementen des Vektors \mathbf{q} . Analog zu Gleichung(33) werden nun die Grenzflächengleichungen erstellt, wobei die Blockmatrix auf der rechten Seite der Gleichung(34) zu

$$\begin{bmatrix} \mathbf{V}_{ss} & \mathbf{V}_{ss} \mathbf{X} \\ -\mathbf{V}_{sp} & \mathbf{V}_{sp} \mathbf{X} \\ -\mathbf{V}_{ps} & \mathbf{V}_{ps} \mathbf{X} \\ \mathbf{V}_{pp} & \mathbf{V}_{pp} \mathbf{X} \end{bmatrix} \quad (54)$$

wird mit den Elementen

$$\begin{aligned}\mathbf{W}_{HY1} &= \mathbf{K}_y \mathbf{K}_x \mathbf{W}_y \mathbf{Q}^{-1} \\ \mathbf{W}_{HY2} &= (\mathbf{E}_{rez}^{-1} - \mathbf{K}_y^2) \mathbf{W}_x \mathbf{Q}^{-1} \\ \mathbf{W}_{HX1} &= (\mathbf{K}_x^2 - \mathbf{E}_{rez}^{-1}) \mathbf{W}_y \mathbf{Q}^{-1} \\ \mathbf{W}_{HX2} &= -\mathbf{K}_x \mathbf{K}_y \mathbf{W}_x \mathbf{Q}^{-1}\end{aligned}$$

$$\begin{aligned}\mathbf{V}_{ss} &= \mathbf{F}_c \mathbf{W}_{EY} - \mathbf{F}_s \mathbf{W}_{EX} \\ \mathbf{V}_{sp} &= -\mathbf{F}_c (\mathbf{W}_{HX1} + \mathbf{W}_{HX2}) - \mathbf{F}_s (\mathbf{W}_{HY1} + \mathbf{W}_{HY2}) \\ \mathbf{V}_{ps} &= -\mathbf{F}_c (\mathbf{W}_{HY1} + \mathbf{W}_{HY2}) + \mathbf{F}_s (\mathbf{W}_{HX1} + \mathbf{W}_{HX2}) \\ \mathbf{V}_{pp} &= \mathbf{F}_c \mathbf{W}_{EX} + \mathbf{F}_s \mathbf{W}_{EY}\end{aligned}\tag{55}$$

wobei \mathbf{F}_c eine Diagonalmatrix mit Elementen $\cos \phi_i$ bzw \mathbf{F}_s mit $\sin \phi_i$ und die Diagonalmatrizen $\mathbf{X} = \exp(-k_0 \mathbf{q} d)$ darstellen. Der Index i beschreibt hierbei den globalen umgeordneten Index von $-n_x n_x \dots n_x n_y$.

Nun wird zur Multilayererstellung erneut das Konzept des full-approaches angewandt, sodass die letzte Grenzflächengleichung wie in Gleichung(43) umgestellt wird zu:

$$\begin{aligned}\mathbf{H} \mathbf{K} \mathbf{R}_l \begin{bmatrix} \mathbf{X}_l & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{V}_{ssl} & \mathbf{V}_{ssl} \\ -\mathbf{V}_{spl} & \mathbf{V}_{spl} \\ -\mathbf{V}_{psl} & \mathbf{V}_{psl} \\ \mathbf{V}_{ppl} & \mathbf{V}_{ppl} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{f}_{1,L+1} & \mathbf{f}_{1,L+2} \\ \mathbf{g}_{1,L+1} & \mathbf{g}_{1,L+2} \\ \mathbf{h}_{1,L+1} & \mathbf{h}_{1,L+2} \\ \mathbf{i}_{1,L+1} & \mathbf{i}_{1,L+2} \end{bmatrix} \begin{bmatrix} \mathbf{T}_s \\ \mathbf{T}_p \end{bmatrix} \\ = \mathbf{H} \mathbf{K} \mathbf{R}_l \begin{bmatrix} \mathbf{X}_l & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{a}_{1,L} & \mathbf{a}_{2,L} \\ \mathbf{b}_{1,L} & \mathbf{b}_{2,L} \\ \mathbf{d}_{1,L} & \mathbf{d}_{2,L} \\ \mathbf{e}_{1,L} & \mathbf{e}_{2,L} \end{bmatrix} \begin{bmatrix} \mathbf{T}_s \\ \mathbf{T}_p \end{bmatrix}\end{aligned}\tag{56}$$

mit exakt derselben Definition der Matrizen und der Indizes wie im 1D-Multilayerfall. Nun wird die Substitution zur Umgehung der \mathbf{X}_L -Invertierung eingefügt:

$$\begin{bmatrix} \mathbf{x} & \mathbf{y} \\ \mathbf{s} & \mathbf{r} \end{bmatrix} = \begin{bmatrix} \mathbf{d}_1 & \mathbf{d}_2 \\ \mathbf{e}_1 & \mathbf{e}_2 \end{bmatrix} \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 \\ \mathbf{b}_1 & \mathbf{b}_2 \end{bmatrix}^{-1} \mathbf{X}_l\tag{57}$$

Somit ergibt sich Gleichung(56) zu:

$$\mathbf{H} \mathbf{K} \mathbf{R}_l \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{x}_L & \mathbf{y}_L \\ \mathbf{s}_L & \mathbf{r}_L \end{bmatrix} \begin{bmatrix} \mathbf{T}_{s,L} \\ \mathbf{T}_{p,L} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_{1,L} & \mathbf{f}_{2,L} \\ \mathbf{g}_{1,L} & \mathbf{g}_{2,L} \\ \mathbf{h}_{1,L} & \mathbf{h}_{2,L} \\ \mathbf{i}_{1,L} & \mathbf{i}_{2,L} \end{bmatrix} \begin{bmatrix} \mathbf{T}_{s,L} \\ \mathbf{T}_{p,L} \end{bmatrix}\tag{58}$$

Für die Rücksubstitution nach dem Lösen der Gleichungen an der ersten Grenzfläche muss

$$\begin{bmatrix} \mathbf{T}_{s,l+1} \\ \mathbf{T}_{p,l+1} \end{bmatrix} = \begin{bmatrix} \mathbf{a}_{1,l} & \mathbf{a}_{2,l} \\ \mathbf{b}_{1,l} & \mathbf{b}_{2,l} \end{bmatrix}^{-1} \mathbf{X}_l \begin{bmatrix} \mathbf{T}_{s,l} \\ \mathbf{T}_{p,l} \end{bmatrix}\tag{59}$$

iterativ eingesetzt werden.

3.3.3 homogene Schicht

In diesem Abschnitt wird kurz die Formulierung für ein Layer mit homogener Brechungsindexverteilung vorgestellt. Hierbei enthält in Abbildung(3) jede Mikrozeile denselben Brechungsindex. In diesem Fall vereinfachen sich die Permittivitätsmatrizen zu Diagonalmatrizen mit $\mathbf{E} = \mathbf{E}_{rez}^{-1} = \epsilon_2 \mathbf{I}$ und $\mathbf{E}^{-1} = \frac{1}{\epsilon_2} \mathbf{I}$. Somit werden die Systemmatrizen in Gleichungen(26) und (27) ebenfalls zu Diagonalmatrizen und die Eigenwerte $\mathbf{Q}_1 = \mathbf{Q}_2 = \sqrt{\mathbf{K}_y^2 + \mathbf{K}_x^2 - \mathbf{E}}$ und Eigenvektoren $\mathbf{W}_1 = \mathbf{W}_2 = \mathbf{I}$ können direkt abgelesen werden. Zur Vermeidung von Divisionen durch Null werden ersetzt

$$\mathbf{B}^{-1} = \text{Diag}\left(\frac{1}{\left(\frac{k_{xi}}{k_0}\right)^2 \frac{1}{\epsilon_2} - 1}\right) \quad \text{und} \quad \mathbf{A}^{-1} = \mathbf{E}^{-1} \mathbf{B}^{-1} \quad (60)$$

Im 2D-Fall ergeben sich dieselben Elemente, nur muss hier die Kopplung des Systems beachtet werden, sodass nun

$$\mathbf{W} = \mathbf{I} \quad \text{und} \quad \mathbf{Q} = \begin{bmatrix} \sqrt{\mathbf{K}_y^2 + \mathbf{K}_x^2 - \mathbf{E}} & \mathbf{0} \\ \mathbf{0} & \sqrt{\mathbf{K}_y^2 + \mathbf{K}_x^2 - \mathbf{E}} \end{bmatrix} \quad (61)$$

3.4 Feldzerlegung

Mit den Formulierungen der beiden vorherigen Kapitel ist es nun möglich, bei einer auf eine Zelle einfallenden ebenen homogenen Welle die gesamte Feldlösung im Zellenbereich sowie in den Gebieten darüber und darunter zu berechnen. In [8] wird eine Methodik vorgestellt, mit welcher die Feldlösung ober- und unterhalb von Strukturen, die aus mehreren periodischen Zellen bestehen, bei beliebigem Eingangsfeld berechnet werden können. Abbildung(6) verdeutlicht die zugehörige Anordnung.

Im einfachsten Fall trifft eine ebene homogene Welle auf die Struktur. In der Abbildung(6) handelt es sich jedoch um einen Gaußstrahl, dessen Hauptachse im Punkt Zentrum die Strukturoberfläche schneidet. Ferner ist ein strahleigenes Koordinatensystem $Ox'y'z'$ gezeigt, dessen Ursprung sich genau in der Strahlmitte des Gaußstrahls befindet. In [8] wird für diesen Strahl eine analytische Darstellung im strahleigenen Koordinatensystem $Ox'y'z'$ präsentiert.

In der Strahlmitte bei $z' = 0$ wird der Strahlradius $w_{0x'}$ bzw. $w_{0y'}$ definiert, wobei im Falle eines nicht zur z' -Achse rotationssymmetrischen Intensitätsprofils beide Strahlradien angegeben werden müssen. Der Strahlradius wird als der Abstand des Punktes bzw. der Kreislinie von der z' -Achse definiert, an welchem die Intensität nur noch $\frac{1}{e}$ der Intensität des zugehörigen Hauptachsenpunktes beträgt.

Alternativ kann beispielsweise eine ebene homogene ebene Welle als Eingangsfeld definiert werden. Dies ist mit $A \exp -jk_0 n_1 z'$ möglich. Hierbei entspricht A beispielsweise einer begrenzten Kreis- oder Rechteckform in der $x'y'$ Fläche des Strahls, wobei jeder Punkt eine Leistungsdichte enthält. In den meisten Anwendungen ist diese über die komplette Form konstant. Ein allgemein schiefer Einfall mit den Winkeln θ_e und ϕ_e entspricht einer Drehung des Koordinatensystems $Ox'y'z'$ zum globalen Koordinatensystem $Oxyz$ der Struktur. Somit wird im Koordinatensystem $Ox'y'z'$ die Oberfläche der Struktur als eine schiefe Ebene mit den Eingangswinkeln als Drehwinkeln definiert. Die Oberfläche ist

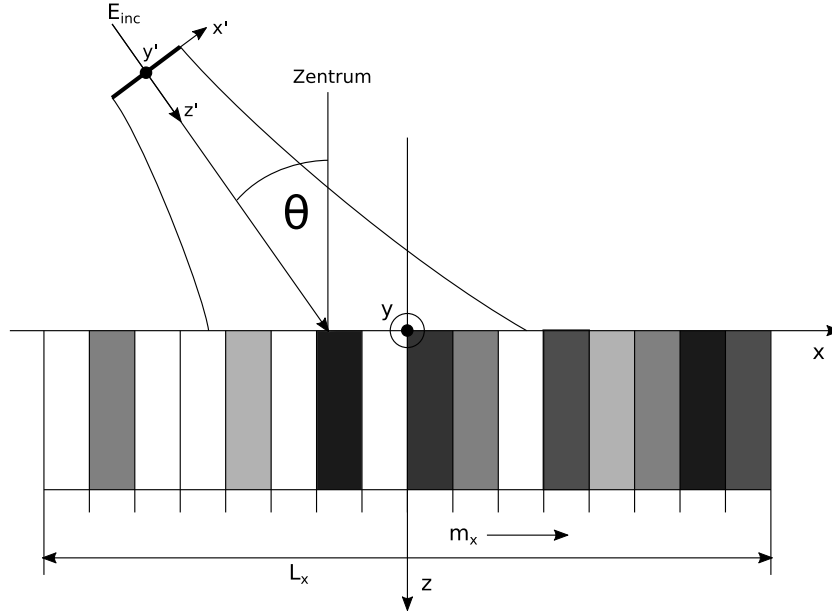


Abbildung 6: Ein Gaußstrahl trifft mit azimuthalem $\phi = 0^\circ$ und polarem $\theta = 0^\circ$ Winkel der Hauptachse auf eine Struktur. Jede einzelne Grauschattierungen für sich repräsentiert hier jeweils eine komplette periodische Zelle. L_x beschreibt den Durchmesser der fouriertransformierten Fläche und der Index m_x identifiziert die jeweilige periodische Zelle

mit einem Samplinggitter belegt, wobei jeder Gitterpunkt eine Zelle der Struktur repräsentiert. Nun berechnet man mit Hilfe der Strahldarstellung an jedem Gitterpunkt mit Index m_x den komplexen Feldvektor an dieser Stelle. Aus ihm lassen sich nach den Formulierungen in [8] die drei Eingangswinkel θ , ϕ und ψ ableiten. Eine exemplarische Darstellung dieser resultierenden Winkelbereiche ist in Kapitel 4.2 angegeben. Ferner werden das komplexe Winkelargument des Vektors als Eingangsphase sowie der Absolutbetrag des Vektors für die Gitterzelle abgespeichert. Hierauf wird im folgenden Ergebniskapitel nochmals eingegangen, da diese Phase für das Design der Oberfläche in präsentierter Anwendung verwendet werden kann.

Die Struktur selbst kann nun aus gleichen Zellen oder mit Hilfe einer Vorschrift bestückt werden; ein Beispiel wird hierfür in Kapitel 4 vorgestellt. Für jede verwendete Zelle werden in der RCWA-Implementation für alle nötigen Eingangswinkelkombinationen berechnet. Die azimuthalen und polaren Winkelschritte sind beliebig definierbar; zur Effizienz werden vorher einmal minimal und maximal nötige Eingangswinkel berechnet und Diskretisierungsschritte festgelegt; bspw. 1° . Für jede dieser Kombinationen wird im RCWA-Programm das E- und H-Feld an der unteren xy-Grenzfläche der Struktur berechnet und in das Feldzerlegungsprogramm eingelesen. Soll in einer Anwendung ein Feld über der Gesamtstruktur berechnet werden, muss die obere xy-Grenzfläche in der RCWA-Simulation bestimmt werden. Nun wird eine zweidimensionale Datenstruktur

initialisiert, die das Feld direkt unterhalb der Struktur enthalten soll, welches nach den Eingangswinkeln dieser Zelle ausgewählt, mit der Eingangsintensität der Zelle skaliert und, falls nötig, um die Eingangsphase verschoben ist. So erhält man das Gesamtfeld an der Unterseite der Struktur. Es kann nun mit den Standardalgorithmen 2D-fouriertransformiert werden.

Diese Fouriertransformation lässt sich durch Multiplikation mit dem Term $\exp(-jk_{3zi}z)$ an den Ort z propagieren. Danach kann durch Rücktransformation das Feld so in einer xy -Ebene am Ort z oder einer xz -Schnittebene angegeben werden.

In [8] wird das Feld auf exakt gleiche Weise zerlegt. Bei dieser Methode wird jedoch zum einen die Phaseninformation des Feldes nicht mitgenommen sowie bei der Feldtransformation können aufgrund der Beugungseffizienzterme keine Standard 2D-fft-Algorithmen genutzt werden.

In den folgenden Analysen wird die Intensität des Feldes für Beleuchtungsanwendungen dargestellt. Diese berechnet sich aus $|S| = E^2/Z_3$, wobei Z_3 dem Feldwellenwiderstand des homogenen Mediums unter der Struktur entspricht. In den meisten Anwendungen ist dies Luft mit dem Vakuumwellenwiderstand $Z_3 = Z_0 = 120\pi$.

4 Verifikation und Modelluntersuchung

Im nun folgenden Kapitel sollen die in Kapitel 2 geforderten quantitativen Aussagen anhand von Modellbeispielen und Referenzen formuliert werden.

Im nun folgenden Kapitel werden die gesamten obigen Formulierungen, die in zwei Softwarepaketen in Python anwenderorientiert implementiert wurden, getestet. Hierbei gilt es, Simulationsergebnisse und Konvergenzverhalten verschiedener Beispielstrukturen und -anordnungen auf verschiedene Weise zu verifizieren. Ferner sollen etwaige relative Abweichungen beim Vergleich von bspw. Beugungsordnungen oder Feldverteilungen in einer Ebene mit anderen Simulationsergebnissen eine Schwelle von 5 Prozent nicht überschreiten sowie die Ursachen, Konvergenzverhalten und Umgang bzw. Lösung dieser Probleme ergründet werden. Ferner soll auch an eher realen Beispielen ein Überblick gegeben werden, welche Simulationsmöglichkeiten diese beiden Module bieten.

Der erste Teil des folgenden Kapitels widmet sich nur der Verifikation und Vergleich einer einzelnen periodische RCWA-Zelle. Im zweiten Teil werden reale Modellbeispiele für die Feldsynthese in Kapitel 3.3 vorgestellt und ebenfalls untersucht, ob sowohl die Grundidee mit ihrer mathematischen Formulierung als auch die dazugehörige Implementierung ebenfalls den in Kapitel 2 aufgeführten Ansprüchen genügen.

4.1 RCWA-Verifikation

In Korrelation zu Kapitel 3 wird noch einmal das Grundproblem dieses Abschnittes definiert. Es existiere eine periodische Zelle mit dreidimensionaler Ausdehnung sowie möglicher Periodizität in zwei Achsen; man betrachte hierzu nochmals die Abbildung(3). Die Zelle kann bezüglich aller drei Achsen bezogen auf ihren Brechungsindex/Permittivität beliebig strukturiert sein. Auf diese Zelle falle eine ebene, homogene und polarisierte Welle mit unendlicher Phasenfläche und beliebigen Eingangswinkeln auf die Zelle ein.

Der Lösung dieses Problems entspricht beispielsweise die Angabe des gesamten elektrischen Feldes in einer xy-Ebene innerhalb der Struktur oder der Angabe aller ausgehenden Beugungsordnungen; sowohl im Halbraum über als auch unter der Struktur unter Nennung der zugehörigen Winkel und Leistung dieser Ordnung.

Im folgenden werden verschiedene Lösungen dieses Problems vorgestellt und anhand dieser die geforderten Aussagen getroffen. Hierbei steigern sich Komplexität der Struktur sowie Umfang der Simulation von Fall zu Fall und orientieren sich am Aufbau von Kapitel 3. Somit wird zu Beginn die 1D-Multilayer Formulierung anhand von den Ergebnissen anderer Autoren, die ebenfalls mit RCWA-Implementierungen berechnet wurden, verglichen. Gleiches wird anschließend mit der 2D-Formulierung vollzogen. Zuletzt wird eine rein akademische Struktur sowohl in der RCWA als auch in einer frei verfügbaren FDTD-Software nachgebaut. Der Hauptaugenmerk der Verifikation liegt genau in diesem Vergleichsmodell.

Viele der zur RCWA publizierten Rechenergebnisse beschreiben Strukturen einfacher Geometrie wie Zylinder und Quader, die durch ein bis zwei Layer vollständig definierbar sind. Der Umgang mit der RCWA hat jedoch gezeigt, dass manche Implementierungs- oder analytische Formulierungsfehler als auch ein langsame-

res Konvergenzverhalten, das auf einer ineffizienten Implementierung beruht, oft erst bei Strukturen mit einer Layeranzahl größer als 50 sowie mit verschiedenen Materialien besonders mit komplexer und wellenlängenabhängiger Permittivität erkennbar sind. Dementsprechend sind diese Ergebnisse und das oftmals sehr gute Konvergenzverhalten leicht nachzubilden, dennoch dementsprechend wenig aussagekräftig. Für komplexere Strukturen konnten nahezu keine publizierten Ergebnisse gefunden werden, weshalb ein Vergleich mit einer selbst aufgesetzten Simulation in einem anderen Rechentool als geeigneter eingeschätzt wird, um die Grundfragen dieser Arbeit zu klären. Somit kann noch eine zusätzliche Aussage zu einem Vergleich zu anderen Rechenansätzen gewonnen werden. Denn auch wenn die RCWA-Ergebnisse in den betrachteten Fällen den Konvergenz- und Abweichungsansprüchen bezüglich anderer Veröffentlichungen genügt, kann man so eine Einschätzung gewinnen, wie sich die RCWA im Vergleich zu anderen Berechnungsweisen verhält und wo eventuelle Probleme auftreten könnten.

Die erste Struktur ist ein eindimensional strukturiertes Gitter mit einer Periode von 600nm. Sie besteht aus zwei Schichten mit jeweils einer Dicke von 300nm, wobei die obere Schicht einen Materialblock mit einer Höhe und Breite von 300nm und die untere komplett mit Material gefüllt ist. Die Halbräume über und unter der Struktur sowie die Umgebung des Blockes in Schicht 1 sind Luft; ferner wurde als Material Gallium-Arsenid (GaAs) mit einem hier konstanten Brechungsindex von $n = 3.65 - 0.0034j$ für alle Wellenlängen gewählt. Das Eingangsfeld sei eine normal einfallende Welle mit TM-Polarisation ($\psi = 0^\circ$) und einer Wellenlänge von $1\mu m$. Als Referenz dient die Abbildung 5.3 b in [10]. Simulation und Referenz sind in Abbildung(7) angegeben.

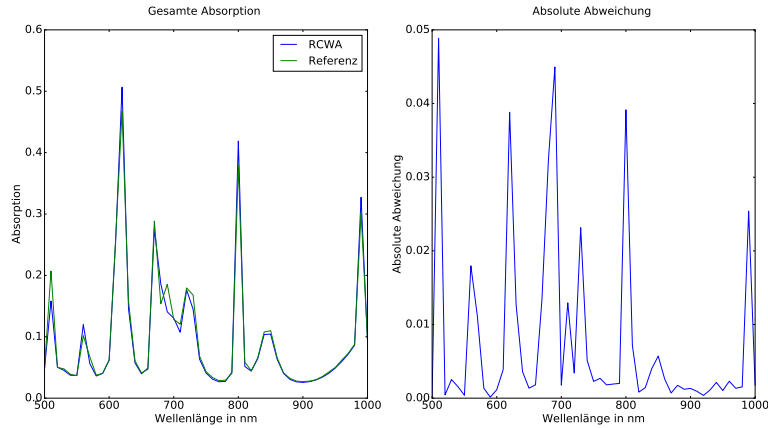


Abbildung 7: 1D-Simulation - links: Referenzsimulation in [10] Abbildung 5.3 b und eigene RCWA, rechts: Absolute Abweichung $|A_{RCWA} - A_{Referenz}|$, $N_x = 141$

Die Absorption ist hierbei durch die Differenz von 1 zur Summe aller Beugungsordnungen (Gleichung(40)) gegeben und ist somit bei diesem verlustbehafteten Medium größer null. Diese Simulation ist jedoch eine reine Modellbetrachtung; der Brechungsindex verlustbehafteter Materialien ist in der Regel stark frequenzabhängig. Mit eingelesenen Indexdaten von [11], woher die meisten Ma-

terialdaten in den hier präsentierten Simulationen stammen (die Datenbank setzt sich aus labortechnischen Messergebnissen sowie Simulationen von bspw. dem Drude-Lorentz Schwingungsmodell zusammen), ergibt sich der tatsächliche Absorptionsverlauf in Abbildung(8).

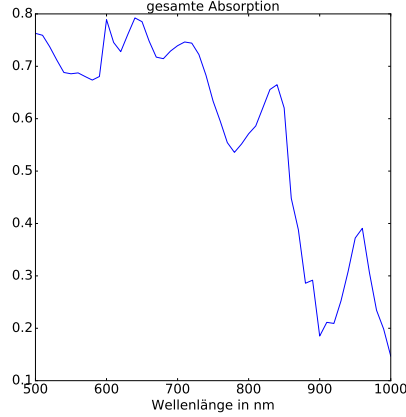


Abbildung 8: Simulation der Struktur aus Abbildung(7) mit wellenlängenabhängigem Brechungsindex, $N_x = 141$

In Abbildung(7) ist zu erkennen, dass beide Kurven grundlegend übereinstimmen. Die Kurvensteigungen und Wendepunkte liegen entsprechen einander, sodass grundlegende Implementierungs- oder analytische Fehler ausgeschlossen werden können. Die einzigen vorhandenen Abweichungen liegen wie gezeigt genau in den Absorptionsspitzen der Kurven. Da hier die Kurve relativ steile Flanken enthält, kann diese Abweichung mit leicht unterschiedlichen Diskretisierung der Wellenlängen als Variable zusammenhängen, wodurch an solch steilen Flanken leicht Abweichungsspitzen entstehen. Ferner zeigte sich in weiteren Vergleichssimulationen, dass an den Punkten von Absorptionsspitzen sehr leicht Abweichungen zwischen verschiedenen RCWA-Implementierungen entstehen. Dies kann durch bspw. der Verwendung eines anderen Algorithmus zur Eigenwertzerlegung verursacht werden.

Im nächsten Beispiel wird eine Struktur mit 2D-Periodizität in Fig. 3 aus [9] auf das Konvergenzverhalten hin untersucht. Hierbei handelt es sich um zylinderförmige Luftlöcher in einer Quarzschicht mit wellenlängenunabhängigem Brechungsindex $n = 1.4622$. Der untere Halbraum sei ebenfalls mit Quarz gefüllt, der obere ist Luft. Eine ebene Welle trifft mit polarem Winkel $\theta = 60^\circ$, $\phi = 0^\circ$ und Polarisationswinkel $\psi = 0^\circ$ bei einer Wellenlänge von 500nm auf die 100nm dicke Schicht. Die Gitterperiode sei in beiden Achsen 300nm weit. Die Beugungseffizienz der ausgehenden reflektiven p-polarisierten (TM) Beugungsordnung (0,0) ist in Abbildung (9) dargestellt.

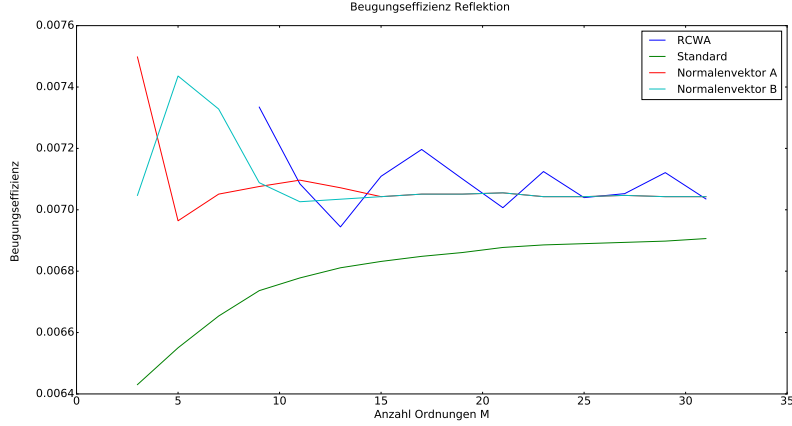


Abbildung 9: 2D-Simulation Hauptordnung $(0,0)_{pp}$ reflektiv im oberen Halbraum - blaue Kurve links: eigene Simulation, rote Kurve rechts: Referenzsimulation in Referenz [9] Fig. 3b

Hierbei beschreibt $M = 2N_x + 1$, wobei $N_x = N_y$, die Gesamtanzahl an Beugungsordnungen in einer Achse der Zelle. Der ideale Grenzwert von 0.00704, auf welchen die RCWA sowie zwei der Referenzen zu konvergieren, lässt sich identifizieren. Dennoch sind im Verlauf der eigenen Simulationskurve größere Schwingungen um den idealen Konvergenzwert zu sehen als in der gezeigten roten und blauen Kurve aus [9]. Die Abweichung der eigenen Kurve von den beiden Referenzkurven ist gravierender als im Beispiel zuvor. Dieses Verhalten ist auf ein Grundproblem der RCWA zurückzuführen. Die hier simulierte Struktur besteht aus zylinderförmigen Luftlöchern in einer Quartzsschicht. Demnach wird in der eigenen Implementierung, wie in Kapitel 3 gezeigt, ein Kreis durch eine Rechteckstruktur mit endlicher Anzahl an Diskretisierungspunkten angenähert. Dadurch entstehen die abgebildeten Sprünge um den Idealwert herum. Mit steigender Gesamtzahl an verwendeten Beugungsordnungen liegt die rechteckige Randkurve mal etwas über der tatsächlichen Kreiskurve, mal etwas darunter. Dieses Verhalten kann durch eine Normalenvektortransformation der Kreiskurve behoben werden, wodurch die verbesserte Konvergenz in der roten Kurve der Referenzsimulation [9] erreicht wird. Hierbei wird für die Permittivität an jedem diskretisierten Punkt der Zelle eine Normalenvektortransformation entlang einer Kurve durchgeführt, sodass im transformierten System die Permittivität einer homogenen Schicht entspricht. Diese Methode wird jedoch zusammen mit der anisotropen Formulierung der RCWA in dieser Arbeit aus Umfangsgründen nicht behandelt. Die grüne Kurve in der Teilabbildung des Referenzplots beschreibt die Standardformulierung der RCWA, die bspw. keine reziproke Epsilonformulierung enthält und welche in den meisten Veröffentlichungen verwendet wird. Hier ist das erheblich langsamere Konvergenzverhalten gegen einen abweichenden Wert sehr gut zu sehen. Somit können auch hier Fehler in Implementierung oder Formulierung ausgeschlossen werden, dennoch sollte ein eventueller Anwender dieses Grundproblem der Diskretisierung von Kurven beim Aufsetzen einer Simulation und Verwendung der Ergebnisse beachten.

Im nun folgenden Abschnitt werden die Ergebnisse der FDTD-Referenzsimulation vorgestellt. Zu Beginn wurde hierfür die Testversion Optiwave [7] eingesetzt, da hiervon eine 32bit Version frei verfügbar ist. Zuerst wurden einige einfache Geometrien wie Quader und Zylinder, die mittig in einer periodischen Zelle liegen, simuliert und die Felddarstellungen in verschiedenen Ebenen durch die Zelle hindurch betrachtet. Bei diesen Strukturen stimmten RCWA und Optiwave gut überein; trotz symmetrisch abgestimmter händischer Diskretisierung zeigte die FDTD jedoch Abweichungen bezogen auf die erwartete Symmetrie der Felddarstellungen. Ein kurzes Ergebnis dieser Vergleichssimulationen ist in Abbildung(10) zu sehen. Hierin ist ein TiO_2 -Zylinder mit 400nm Höhe und Durchmesser von 140nm mittig in einer luftgefüllten periodischen Zelle der Dicke 600nm mit einer Gitterperiode von 250nm in x- und y-Richtung. Auf die Zelle fällt eine x-polarisierte ebene Welle mit Vakuumwellenlänge 532nm unter normalem Einfallswinkel ($\theta = 0^\circ$ und $\phi = 0^\circ$) ein.

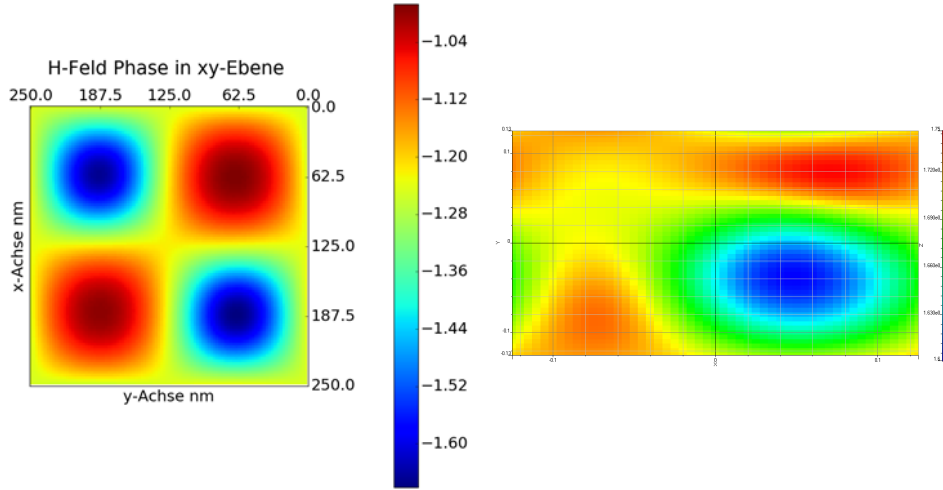


Abbildung 10: Symmetrieproblem bei der FDTD Referenzsimulation in Optiwave: links Phase in rad der x-Komponente des H-Feldes in der xy-Ebene am Boden der Zelle - rechts gleiche Darstellungsweise in Optiwave

In Abbildung(10) sind im Feldverlauf aufgrund der symmetrischen Struktur und des normalen Feldeinganges rein symmetrische Verläufe zu erwarten. Wie zu sehen ist erfüllt die RCWA diesen Anspruch, die FDTD zeigt jedoch auch nach händisch eingestellter Diskretisierung mit symmetrisch liegendem Nullpunkt die gezeigte Abweichung.

Anschließend wurde, um noch mehr Aussagen zum Vergleich der Rechenmethoden treffen zu können, ein erheblich kompliziertes Modellbeispiel gewählt. Abbildung(11) zeigt den Aufbau des Beispiels. Es handelt sich hierbei um eine luftgefüllte Zelle, in der zentral ein hier nur prinzipiell aufgebauter FINFET-Transistor positioniert ist. Um die Interpretation und den Vergleich der Ergebnisse zu erleichtern, wurde die Struktur komplett symmetrisch aufgebaut. Hierbei hat die vordere längere x-Achse eine Periode von $1.5\mu m$ und die kürzere

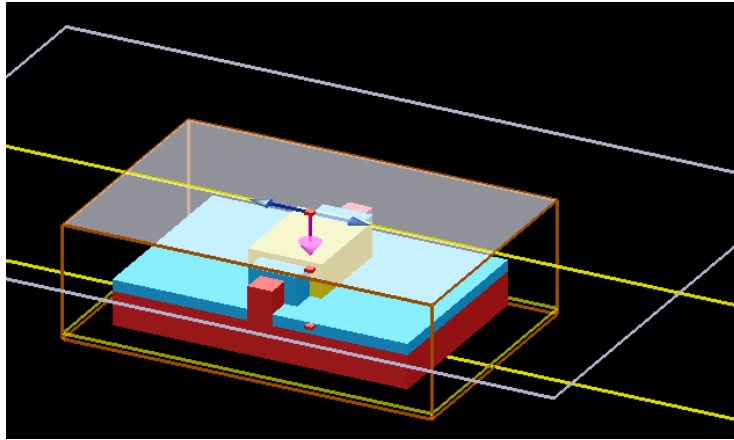


Abbildung 11: Darstellung des FINFET-Modellbeispiels: rot - undotiertes Silizium $n=4.2581-0.04514j$, blau - TiO_2 -Oxid $n=2.563$, gelb - Gold $n=0.6084-2.5886j$; der orangene Rahmen definiert die Simulationszelle und für seine Randflächen werden Randbedingungen definiert; die weiße Ebene entspricht der Phasenfläche des Eingangsfeldes

hintere y-Achse eine Periode von $1\mu m$. Als Eingangsfeld dient eine ebene Welle mit Wellenlänge $532nm$ in TE-Polarisation. Somit liegt der E-Feldvektor des Eingangsfeldes nur in x-Richtung. Finne und Schichtoxid weisen eine Dicke von $100nm$ auf, die des Gateoxids und der Goldschicht dagegen $50nm$. Die gesamte Struktur ist nach allen Seiten von Luft umschlossen, sodass eventuelle Abweichungen oder Probleme der FDTD mit komplexen Brechungsindizes an Randflächen als Abweichung ausgeschlossen werden können.

In der FDTD selbst wird die Eingangsebene als die Deckelfläche der orangefarbenen Zelle gewählt. Für die seitlichen Flächen sind allesamt periodische Randbedingungen (PBC) festgelegt. Die Boden- und Deckelfläche der Zelle sind als perfekt absorbierende anisotrope Schichten mit 10 Absorptionsschichten definiert.

Dieses Modell wurde in Optiwave aufgesetzt. Diese Software zeigte jedoch bei beliebiger Variation von Zeit- und Ortsdiskretisierung in verschiedensten Größenordnungen keinerlei Konvergenzverhalten. Die Ursache des Problems blieb ungeklärt und so wurde alternativ das Modell in einer Testversion von Lumerical [11] aufgesetzt. Hier war sofort eine Konvergenz erkennbar und aus Beobachtung des Variation der Simulation bei verschiedenen Diskretisierungen wurde die Lösung bei einem örtlichen Schrittweite von $0.25nm$ und Zeitschritten von jeweils $0.02346fs$ bei einer gesamten Simulationszeit von $1000fs$ als konvergentes Ergebnis ausgewählt. Diese FDTD-Lösung sowie die Felddarstellungen der RCWA-Simulation sind in Abbildung(12) dargestellt. Die RCWA ist hierbei mit $N_x = 41$ und $N_y = 21$ Ordnungen diskretisiert, was zu einer Simulationszeit von unterhalb einer Minute führt. Die FDTD benötigt hingegen mehr als 8min Ausführungszeit.

Gut zu erkennen sind hierbei die prinzipiell gleichen Felder beider Simulationsarten. Die FDTD errechnet höhere Feldspitzen an den Goldecken am oberen Bildrand, wodurch andere Feldartefakte, die in der RCWA vorkommen, weniger zu sehen sind. Grundsätzlich lässt sich die hohe absolute Abweichung beider Simulationen an den Goldecken mit der Art und Weise von Diskretisierung und Strukturdarstellung erklären. Da, wie in Kapitel 3 vorgestellt, die RCWA Ecken

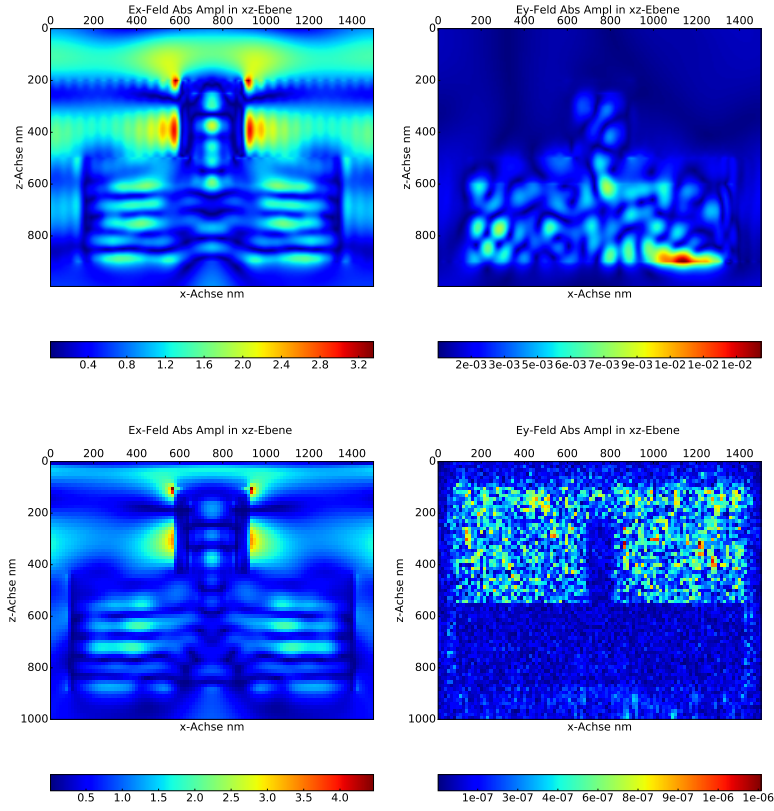


Abbildung 12: FDTD-Referenz: oben RCWA, unten FDTD Simulation; gezeigt ist der Absolutbetrag der Amplitude der elektrischen Feldkomponente x und y in einer xz-Ebene, die genau auf der Hälfte der y-Periodenlänge positioniert ist; aufgrund der komplexen Permittivität errechnete die RCWA einen Absorptionsanteil von 0.594735

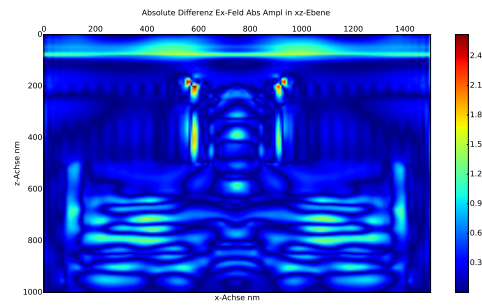


Abbildung 13: Differenz der Absolutbeträge von E_x beider Simulationen nach $||E_{x,RCWA}| - |E_{x,FDTD}||$

durch eine Fourierreihe mit endlicher Koeffizientenzahl annähert, herrscht ein Grundunterschied zwischen beiden Strukturabbildern in den jeweiligen Simulatio-

nen. In der FDTD können solche Ecken exakt nachgebaut werden, da für jeden Gitterpunkt der Brechungsindex definiert wird und keine Fourierreihenentwicklung stattfindet. An den Ecken der Goldschicht kommt dieser Effekt besonders zum Tragen, da an leitenden Metalloberflächen hohe Feldbeträge, die wiederum im Gold selbst sehr rasch abklingen, entstehen. Ferner ist die Struktur bezogen auf die Diskretisierung im FDTD-Programm etwas anders gewählt worden, wodurch die Struktur in der FDTD etwas ausgedehnter erscheint. Zuletzt wird als mögliche Abweichungsursache die leichten Unterschiede in der Angabe des Brechungsindex vermutet. In der RCWA können komplexe wie reelle Brechungsindizes explizit angegeben werden, in dem verfügbaren FDTD-Programm werden sie nicht einsehbar aus einem vordefinierten Lorentz-Drude-Schwingungsmodell berechnet. Auch sind in Abbildung(13) im oberen Bereich Lamellen zu erkennen, die in der RCWA-Simulation vorhanden sind und als numerische Artefakte resultierend aus der Eigenwertzerlegung großer Matrizen in diesen Schichten interpretiert werden können.

Als Grundaussage des modellbasierten Vergleichs zur FDTD ist gezeigt, dass grundsätzlich beide Lösungen gegeneinander konvergieren. Das abklingende Feldverhalten beispielsweise im Bezug auf Metallschichten wie auch andere Feldmerkmale werden in gleicher Weise reproduziert und die Felder sind grundsätzlich vergleichbar. Grundsätzlich lässt sich so im Bezug auf die RCWA sagen, dass die Fourierzerlegung der Strukturpermittivität einige Abweichungen gegenüber Verfahren, in welchen die Permittivität ohne Zerlegung in die Berechnung einfließt, verursacht. Einerseits sind exakte Eckennachbildungen aufgrund der endlichen Anzahl an Ordnungen nicht möglich, was explizit im letzten Beispiel auch gezeigt wurde. Die FDTD hingegen verwendet die Brechungsindizes, die für einzelne Gitterpunkte angegeben sind, direkt ohne Fouriertransformation in den Maxwellgleichungen. Auch spielt die Diskretisierung selbst eine Rolle beispielsweise im Bezug auf die Lage des mittleren Samplingpunktes zwecks Symmetrie als auch zwecks dem Umgang von Brechungsindexsprüngen zwischen zwei Meshgitterpunkten. Dies ist Hauptursache für die gezeigten größten Abweichungen, an denen eine quantitativ hohe Abweichung herrscht. Dieses Problem beschreibt jedoch Abweichungen zur Diskretisierung, die umgangen in echt quantitativer Vergleich deshalb schwer, da in echt tatsächlich verschiedene Strukturen abgebildet werden

Im restlichen Simulationfeld hingegen sind nur leichte Artefakte gerade im oberen Bildteil zu sehen, die sich als höherfrequente Eigenwertlösungen interpretieren lassen. Dies kann numerische Ursachen in der Eigenwertzerlegung haben, allerdings besteht auch die Möglichkeit, dass diese Artefakte bei weiterer Erhöhung der Ordnungsanzahl verschwinden.

4.2 Modellbetrachtung der Feldzerlegung

Im folgenden Teilkapitel soll nun untersucht werden, ob das in Kapitel 3.3 vorgestellte Konzept der Feldzerlegung richtig analytisch formuliert und implementiert ist sowie den in Kapitel 2 definierten Ansprüchen bezüglich Konvergenz und Abweichungen genügt.

Wie in Kapitel 3.3 schon erwähnt wurde dieses Konzept aus der Fragestellung, ob Metalinsen mit der RCWA berechnet werden können, entwickelt. Metalinsen bestehen allgemein aus gleichartigen, einfachen geometrischen Körpern wie bspw. Zylindern oder rechteckigen Blöcken, die in Planartechnologie aus bspw. TiO_2 oder Polysilizium auf einen Glaswafer aufgebracht werden. Eine einzelne dieser periodischen Zellen eignet sich sehr gut für eine Feldsimulation mit der RCWA, jedoch reicht hierbei die Betrachtung der Beugungseffizienzen als Ergebnis nicht aus. Da diese Elemente in ihren Perioden und Elementgrößen kleiner als die einfallende Wellenlänge sind, entstehen keine Beugungseffekte und man erhielte nur einen Transmissionskoeffizient in der nullten Ordnung. Die Funktion der Metalinse beruht jedoch auf der unterschiedlichen Phasenverschiebung des Eingangsfeldes durch jede einzelne Zelle. Somit wurde das Lösungskonzept formuliert, dass die RCWA ein komplettes Ausgangsfeld in einer Ebene unterhalb eines Elementes sowie die dazugehörige Phasenverschiebung berechnet. In Abbildung (14) sind genau diese Phasenverschiebung abhängig vom Durchmesser der hier verwendeten Zylinder dargestellt.

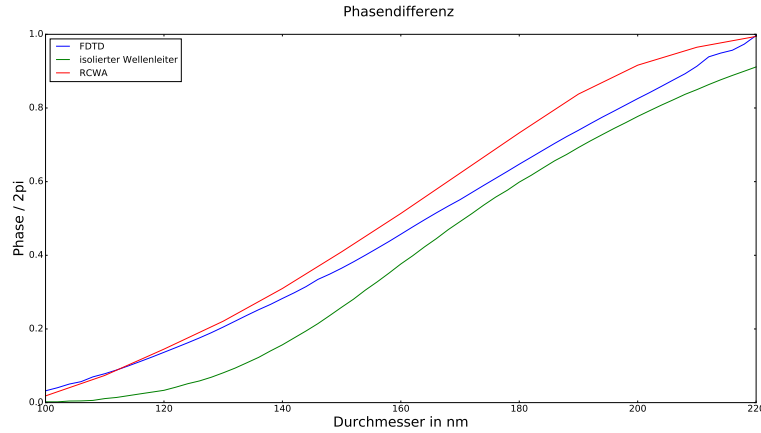


Abbildung 14: Abhängig vom Zylinderdurchmesser induzierte normierte Phasenverschiebung eines einzelnen Zellelementes; basierend auf verschiedenen Berechnungsmethoden; die Zylinder haben eine Höhe von 600nm und sind mittig in der zugehörigen Zelle mit einer Periode von 250nm in jeder Achse positioniert

Die beiden Referenzkurven (grün und blau) entstammen der Abbildung 1 e) in [2]. Die Berechnung eines isolierten Wellenleiters entspricht der Modellvorstellung, dass das Feld innerhalb einer periodischen Zelle maßgeblich durch den Zylinder bestimmt wird. Das Feld in einem solchen Zylinder wird als Feldlösung eines allgemeinen runden Wellenleiters betrachtet. Hierbei wird die Wellengleichung in Zylinderkoordinaten ausformuliert und mit Hilfe der paraxialen Näherung vereinfacht, sodass sich für die radiale Feldkomponente die Bessel-

sche Differenzialgleichung ergibt. Deren Lösung sind die Besselschen Funktionen in ganzzahligen Ordnungen. Durch Gleichsetzen der allgemeinen Feldlösungen an der Grenzfläche des Zylinders und Lösen der daraus entstehenden charakteristischen quadratischen Gleichung erhält man die allgemein als HE und EH (die Benennung orientiert sich am Vorzeichen der beiden Lösungen der Gleichungen) bekannten Feldlösungen ganzzahliger Ordnungen. In [2] wird davon ausgegangen, dass nur die HE_{11} Grundmode im Zylinder ausbreitungsfähig sei und propagiert diesen analytischen Term durch den Zylinder. In Abbildung(14) sind besonders die Abweichung des isolierten Wellenleitermodells gut zu erkennen. Ein Erklärungsansatz hierfür ist, dass in dem Modell des Wellenleiters das Mantelgebiet des Wellenleiters als homogen und rotationssymmetrisch definiert ist. Durch diese Kreisform lässt sich jedoch eine Fläche wie die Metalinse nicht akkurat beschreiben.

Die Phasenverschiebung aus der RCWA berechnet sich aus dem komplexen Winkel der transmittierten Amplituden T_s und T_p .

In [2] wurde mit der allgemeinen Definition der numerischen Apertur formuliert, dass an jedem Punkt (x,y) der Metalinse eine Phasenverschiebung von

$$\Phi_t(x, y) = 2\pi - \frac{2\pi}{\lambda} (\sqrt{x^2 + y^2 + f^2} - f) \quad (62)$$

in das Eingangsfeld induziert werden muss. Dieses wird im Folgenden als eine normal einfallende ebene homogene Welle (gleiche Eingangsphase an jedem Linsenpunkt) definiert. Durch die Phasenverschiebung wird eine Fokussierung im Abstand $z = f$ von der Linse realisiert. Die gesamte Linsenfläche wird nun in periodische Zellen aufgeteilt, wobei sich mit Hilfe von Gleichung(62) für jede Zelle der für diese Zelle nötige Zylinderdurchmesser angeben lässt. In der RCWA werden nun für jeden Zellentyp (in den folgenden Ergebnissen wurden 11 Zylindertypen mit einer Durchmesserdiffferenz von 10nm von 100nm bis 200nm verwendet) das Ausgangsfeld berechnet und eine virtuelle Luftschicht einige Nanometer unterhalb der Linse mit diesen Felddarstellungen belegt. Um diese Idee zu verdeutlichen ist das resultierende Feld aus der so erstellten Phasenkarte zur besseren Sichtbarkeit für eine sehr kleine Linse von $12\mu m$ bei einer Wellenlänge von 532nm und einer Brennweite von $8\mu m$ in Abbildung(15) gezeigt.

Die Zellenperiode in diesem Beispiel wie auch in den folgenden Analysen beträgt 250nm in x- und y-Richtung; die Vakuumwellenlänge des einfallenden Lichtes beträgt 532nm.

Im Folgenden wird dieses Modell nun verifiziert. Hierfür werden Messergebnisse sowie FDTD-Simulationsergebnisse aus [2] verwendet; zusätzlich muss die Feldverteilung im Brennpunkt gegen die Airyfunktion(63) konvergieren, die sich aus der Besselfunktion für Beugungsscheibchen ergibt und die in den folgenden Abbildungen als Idealverlauf dazu gezeigt wird.

$$I(\theta) = I_0 \left(\frac{2J_1(kasin(\theta))}{kasin(\theta)} \right)^2 \quad \text{mit} \quad x = kasin(\theta) \quad (63)$$

Hierbei entspricht J_1 der Besselfunktion erster Art, a dem Linsenradius, I_0 der maximalen Intensität in der Zentrumsachse und θ dem polaren Winkel am aktuellen Flächenpunkt.

In Abbildung(16) sind die Messergebnisse einer von den Autoren von [2] real hergestellten Metalinse mit einem Durchmesser von $300\mu m$ und einer

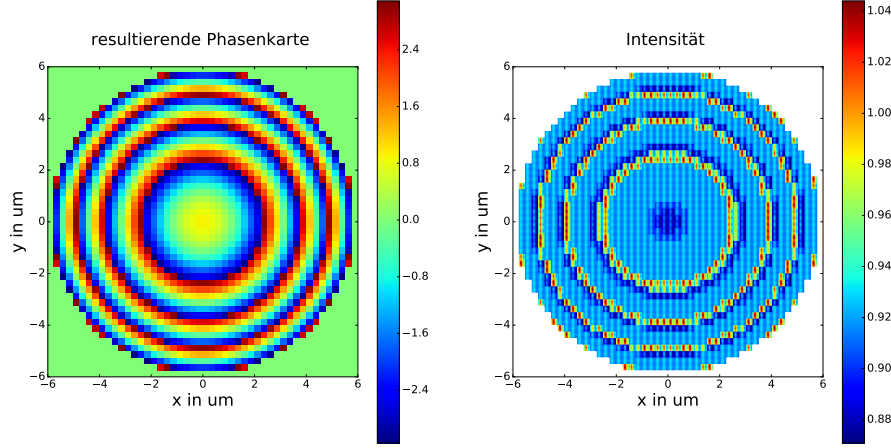


Abbildung 15: links: diskretisierte Phasenkarte nach Gleichung(62); rechts: resultierende Feldbelegung unterhalb der Linse als Betragsquadrat des Feldes

Brennweite $f = 200\mu\text{m}$ gezeigt. Dies entspricht einer relativ hohen numerischen Apertur von 0.6.

Zur Abschätzung der Korrektur sind jeweils die Airyfunktion und die Position der ersten Minima als Linien eingezeichnet. Der Intensitätsverlauf in der xz -Ebene ist in Abbildung(18) gezeigt. Diese Linse hat hierbei nur einen Durchmesser von $24\mu\text{m}$ und eine numerische Apertur von 0.85, was in einer Brennweite von $7.437\mu\text{m}$ resultiert. In jeder Zelle ist das Feld hierbei an 20 Samplingpunkten pro Achse eingelesen.

Der in Abbildung(16) gezeigte Vergleich ergibt, dass exakt gleiche Feldartefakte vorhanden sind und dass dieses Simulationskonzept somit für diese Auslegung von Metalinsen funktioniert. Die einzige erkennbare Abweichung ist die leichte Verschiebung des ersten Minimums des Intensitätsverlaufs im Vergleich zur Airyfunktion. Die Differenz beider Minima beträgt etwa 50nm . Einerseits ist in den präsentierten Messergebnissen von [2] ebenfalls eine Differenz der beiden Minimapunkte zu sehen. Andererseits werden in diesem Simulationskonzept periodische Felder der einzelnen Zellen zusammengesetzt und propagiert. Damit ist der Feldübergang an den Grenzen zwischen Zellen und somit das in Abbildung(16) gezeigte, daraus propagierte Feld physikalisch nicht korrekt. Aufgrund der hohen Zellenanzahl sowie der reinen Phasenverschiebung des Feldes nahezu ohne Beugung innerhalb einer Zelle konvergiert diese Methode jedoch wie aufgezeigt gegen das ideale Ergebnis.

Im Folgenden wird deshalb die Charakteristik dieses propagierten Feldes näher betrachtet. Hierzu eignet sich aufgrund der starken Fokussierung eine logarithmische Betrachtung in Abbildung(18).

Es wird erwartet, dass sich die eben geschilderten Artefakte aufgrund der Zellengrenzflächen bei der Propagation des Feldes in z -Richtung fortpflanzen. Genau dies ist in Abbildung(18) zu sehen. Ein Verbesserungsansatz ist hierbei die Erhöhung der fouriertransformierten Fläche, in der das Feld aus den RCWA-

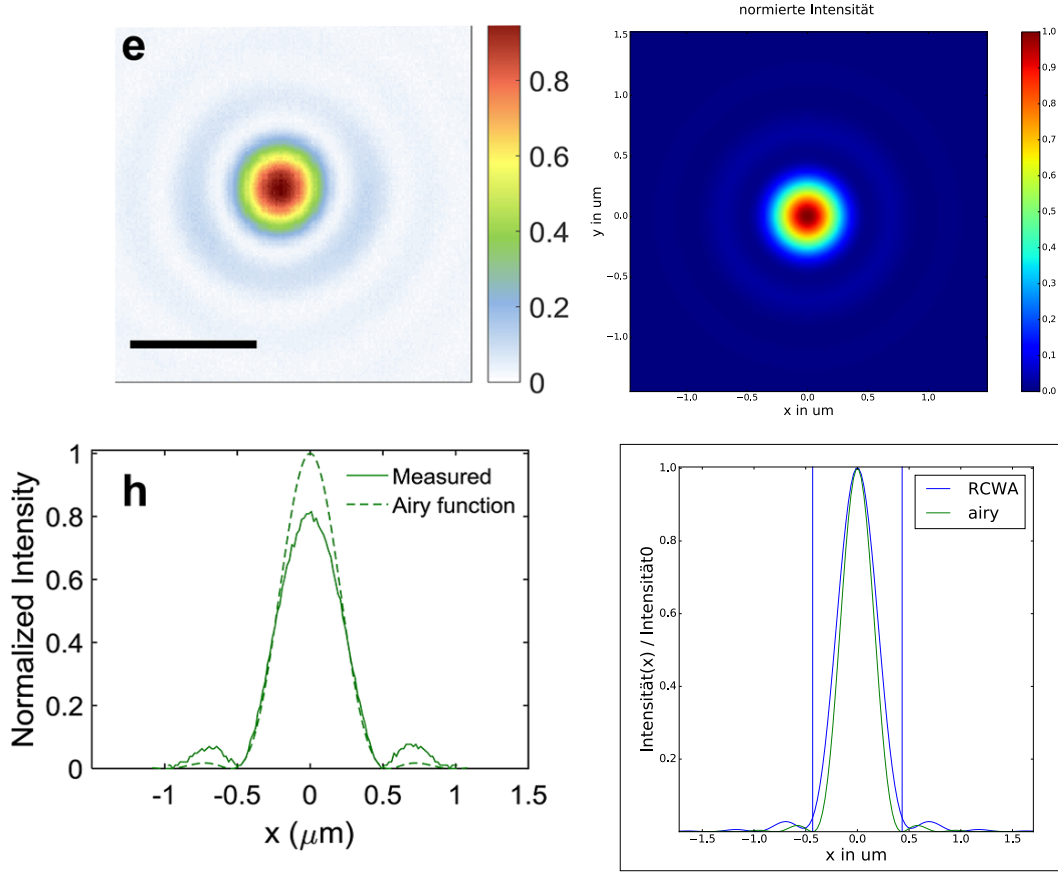


Abbildung 16: links: gemessene Intensitätsverteilung in der xy-Ebene bei $z = f$ aus den Messergebnissen von [2] Figure 3, rechts: Simulationsergebnisse der RCWA-Simulation bei gleichen Parametern

Zellsimulationen erstellt wird. Abbildung(18) zeigt den positiven Effekt bezüglich dieser Grenzartefakte bei bspw. Verdopplung dieser Fläche. Die Linse liegt hierbei immer noch im Zentrum der Fläche, jedoch macht ihr Durchmesser nur noch die Hälfte der Fourierfläche aus. Dieser positive Effekt kann zur Veranschaulichung des Prinzips in Abbildung(18) ein Schnittbild des Feldes in der xz-Ebene gezeigt. Hier sind ebenfalls gleiche Feldartefakte zu sehen und besonders die Intensitätsmaxima liegen an den selben Orten. Die Autoren von [2] simulierten diesen Verlauf in einer FDTD und führten an, dass sie mit dieser Simulation einer Linse von $24\mu\text{m}$ Durchmesser im Bezug auf Speicher- und Rechenleistung ihr Maximum erreicht haben. Somit wird an dieser Stelle direkt ein Vorteil der Feldsynthese deutlich, da die in Abbildung(16) simulierte Linse mit einem Durchmesser von $300\mu\text{m}$ eine Berechnungsdauer auf einem Desktop-PC von zwei Minuten aufweist.

Im Folgenden wird nun die Feldsynthese bei einem einfallenden Gausstrahl auf

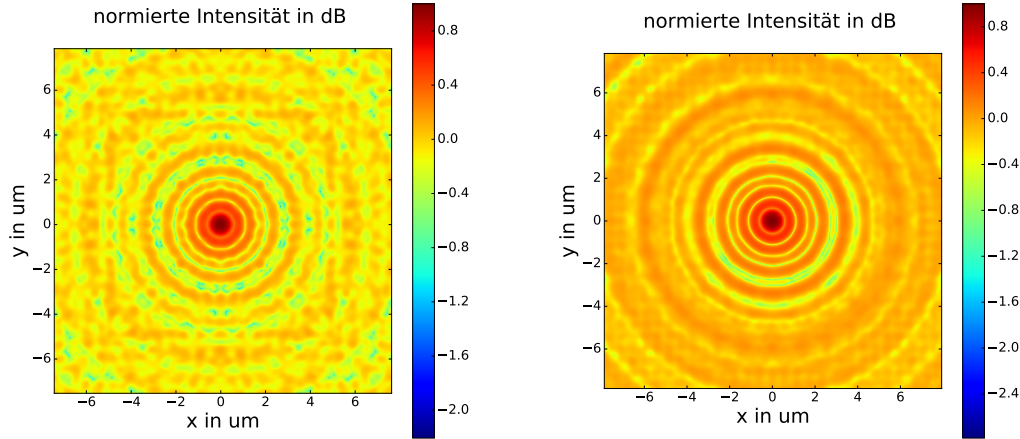


Abbildung 17: Beispielhaft Linse mit $d = 50\mu m$ Durchmesser und Brennweite $40\mu m$:
links $L_x = L_y = d$, rechts $L_x = L_y = 2d$

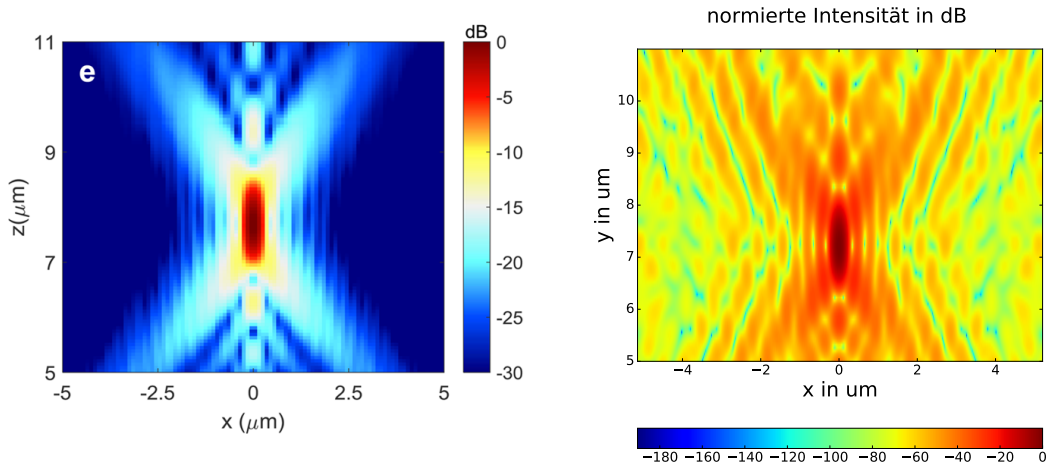


Abbildung 18: links: Intensitätsverteilung in der xz -Ebene der FDTD Simulation in [2], rechts: Ergebnisse der RCWA-Simulation

ein Material komplexer Permittivität vorgestellt. Hierbei sind im folgenden Beispiel alle Zellen gleich strukturiert, um die gezeigten Felddbilder mit Hilfe analytischer Überlegung zum Feldverhalten nachvollziehen zu können. Es handelt sich hierbei um eine kreisförmige Fläche mit Durchmesser $60\mu m$, die mittig in der Zerlegungsfläche mit Kantenlänge $120\mu m$ positioniert ist. Sie ist aus periodischen Zellen mit Kantenlängen $1\mu m$ in beiden Achsen besetzt, die jede ein Goldplättchen umgeben von Luft enthalten. Jedes Plättchen ist quadratisch mit einer Kantenlänge von $500nm$ und mittig in seiner Zelle positioniert. Jede Zelle

ist 500nm dick und die Vakuumwellenlänge des einfallenden Lichtes entspricht 600nm. Der Gaußstrahl habe einen Radius von $5\mu m$ in seiner Strahltaile, welche bezogen auf die Hauptachse des Strahls $100\mu m$ von der Struktur entfernt liegt.

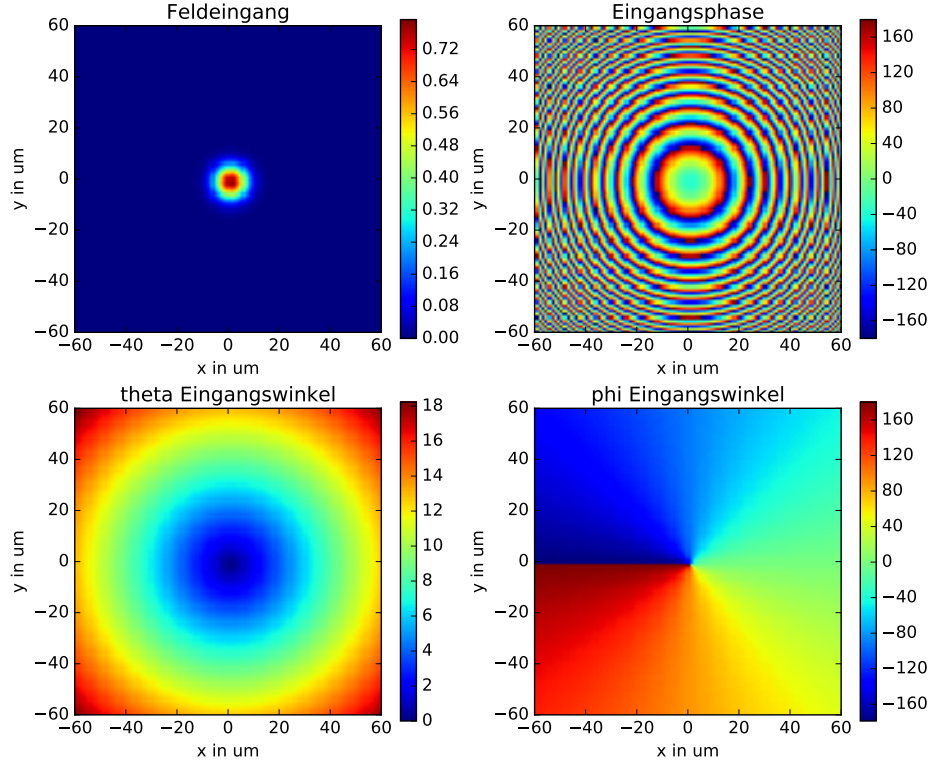


Abbildung 19: Karten des Eingangsfeldes, welche sich aus der Vektorrechnung ergeben und für die Bestückung des Ausgangsfeldes mit den zugehörigen Zellenfeldern nötig sind; hier exemplarisch für $\theta = 0^\circ$ und $\phi = 0^\circ$ sowie $\psi = 45^\circ$, Winkelangabe in Grad

In Abbildung(19) ist die nötige Zerlegung des eingehenden Gausschen Strahles gezeigt. Wie in Kapitel 3 beschrieben muss aus diesem Strahl für jede Zelle der gesamten Struktur ein azimuthaler, polarer und Polarisationswinkel sowie die Eingangsintensität identifiziert werden. Der Intensitätsplot zeigt so auch die Lage und Grundcharakteristik des Eingangsfeldes an. In Abbildung(20), hier für einen nicht normalen Einfall, sind die so resultierenden Winkelkarten gezeigt, welche direkt im Programm als Indizierung für die jeweilige Zellsimulation verwendet werden. Die RCWA-Zelle wird vorher bei einer Diskretisierung von einem Winkelgrad für alle θ Winkel und einen azimuthalen Winkel ϕ von 0° bis 45° simuliert. Hierbei wird die Symmetrie der RCWA-Zelle ausgenutzt und der verbleibende restliche azimuthale Winkelbereich bis 360° mit den passenden rotierten und gespiegelten Feldsimulationen aus dem Urbereich von 0° bis 45° bestückt.

Nun wird nach gleichem Prinzip ein schräger Einfall simuliert, wobei für die Hauptachse des Strahls ein azimutahler Winkel von $\phi = 0^\circ$ und ein polarer Winkel von $\theta = 30^\circ$ definiert werden. So kann das grundsätzliche Verhalten aufgrund der Symmetrie im Ergebnis besser abgeschätzt werden. Alle anderen Eingangsparameter des Strahls bleiben konstant.

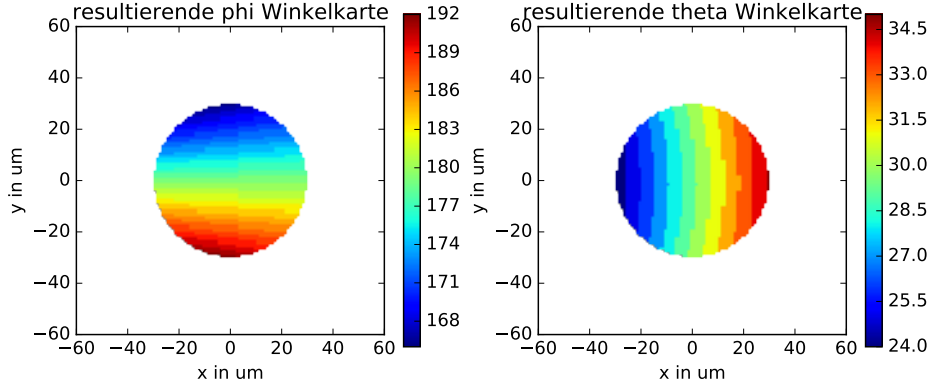


Abbildung 20: Karten des Eingangsfeldes, welche sich aus der Vektorrechnung ergeben und für die Bestückung des Ausgangsfeldes mit den zugehörigen Zellenfeldern nötig sind; hier exemplarisch für $\theta = 30^\circ$ und $\phi = 0^\circ$ sowie $\psi = 45^\circ$, alle Winkelangaben in Grad

Die Simulationsergebnisse dieser Einstrahlung auf ein Gitter mit Goldelementen ist in Abbildung(21) und (22) dargestellt. In allen drei Darstellungen des Feldes trifft der Strahl von der linken Seite parallel zur x-Achse unter einem polaren Winkel von 30° auf die Struktur. Zur Verdeutlichung des Prinzips ist noch einmal das Feld bzw die Synthetisierung der Zellen in der Abbildung(21) links gezeigt.

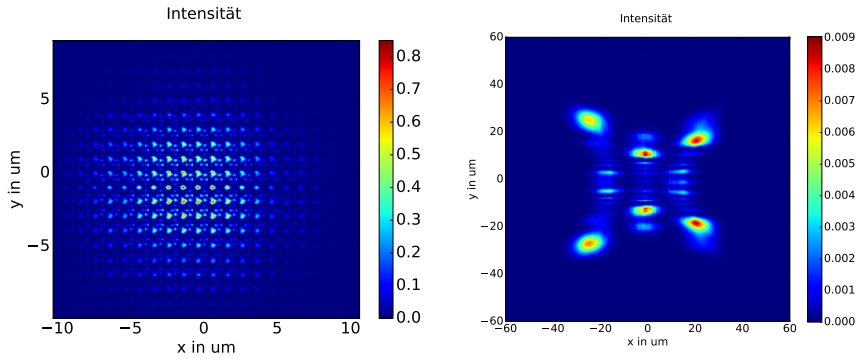


Abbildung 21: links: Intensitätsverteilung in der xy-Ebene bei $z = 0\mu m$ - rechts: xy-Ebene bei $z = 20\mu m$

Die Felddarstellung in der xy-Ebene bei $z = 20\mu m$ in Abbildung(21) rechts entspricht hierbei dem erwarteten Ergebnis. Der Strahl wird in x und y-Richtung gebeugt und erzeugt aufgrund des azimuthalen Winkels von 0° ein zur x-Achse

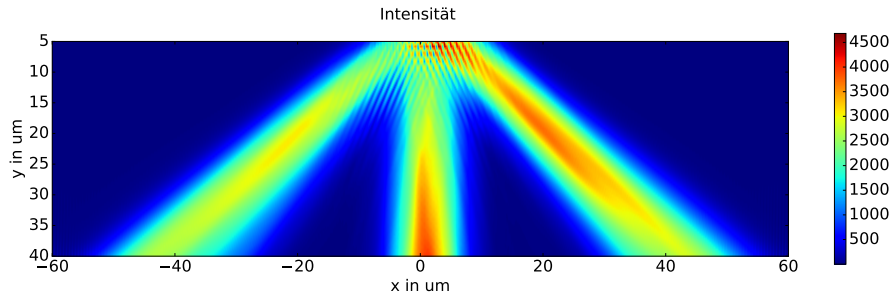


Abbildung 22: Intensitätsverteilung in der xz-Ebene, hierbei wurden die Absolutbeträge des Feldes entlang der y-Achse addiert und auf die betrachtete Ebene projiziert.

symmetrisches Ausgangsfeld. Analytisch gesehen entspricht eine ebene homogene Welle im ortstransformierten Wellenvektorraum (k-Raum) aufgrund ihrer unendlichen Ausleuchtung und Phasenfläche einem Diracpuls. Somit stellt jede Beugungsordnung des Gitters einen Diracpuls im k-Raum dar. Die Transformation eines Eingangsfeldes an einem Gitter oder einer anderen Struktur kann als Faltung betrachtet werden. Diese entspricht im k-Raum, in welchen man über eine örtliche Fouriertransformation gelangt, einer Multiplikation. Der Gaußstrahl wiederum behält seine grundlegende Strahlform als Gaußverteilung beim Übergang in den Fourierraum bei. Somit kann das Ausgangsfeld als die Multiplikation einer Gaußverteilung mit jeder einzelnen Beugungsordnung betrachtet werden. Somit sind im Ausgangsfeld in einer xy-Ebene einzelne Gaußverteilungen zu sehen, deren Lage bzw Richtung der Hauptachse durch die Beugungsordnungen bestimmt sind. Besonders ist dies in Abbildung(21) rechts zu identifizieren, wo die größten Intensitäten in der -1.,0. und +1. Ordnung in x-Richtung liegen sowie in der -1. und +1. Ordnung der y-Richtung.

5 Ergebnisse und Ausblick

In dieser Arbeit wurde das Grundkonzept und die nötigen Formulierungen der RCWA erläutert und mit Hilfe eigener Formulierungen und Korrekturen von bestehenden Veröffentlichungen erweitert. Mit Hilfe dieser darauf basierenden Software kann somit die Feldlösung einer in allen drei Achsen beliebig strukturierte periodische Zelle, auf die eine ebene homogene Welle einfällt, bestimmen werden. Diese Implementierung wurde auf verschiedenste Weise verifiziert und getestet, wo sich zum einen gezeigt hat, dass in der Formulierung keine analytischen Fehler identifiziert werden konnten sowie im Vergleich von Simulationsergebnissen in Veröffentlichungen zur RCWA keine langsamere Konvergenz zu beobachten ist. Besonderes Augenmerk lag auf dem Vergleich eines Modellbeispiels mit der FDTD. Das hier betrachtete Modell bestand aus verschiedensten ineinander übergehende Geometrien und Materialien mit komplexer Permittivität, wobei letztere üblicherweise ein schlechteres Konvergenzverhalten oder auch gar starke Abweichungen zwischen verschiedenen Simulationsarten verursachen können. Mit Vergleich dieses eher realen Modells wurde gezeigt, dass die RCWA als semianalytischer Lösungsansatz auch gegenüber einem komplett verschiedenen Lösungsansatz, in diesem Falle einem iterativen Ansatz mit Iterationen über Zeit und Raum, in einem realen Modell gegen dieselbe Lösung konvergiert. Dies geschieht ohne beispielsweise das Auftreten numerischer Probleme, die eine Divergenz verursachen. Die beobachteten Abweichungen lassen sich größtenteils auf leicht verschiedene Brechungsindizes der verwendeten Materialien sowie Unterschiede in der Diskretisierung und Darstellung der Struktur zurück zu führen. Diese Diskretisierung, beispielsweise bezogen auf symmetrische Strukturen oder den Umgang von Geometriekanten, die zwischen zwei Diskretisierungspunkten liegen, können in der RCWA manipuliert werden. Sie sind jedoch in kommerziellen Programmen nicht exakt gleich einstellbar, woraus diese Abweichungen resultieren. Dennoch verursacht eine kleinere zur Verfügung stehende Rechen- bzw. Speicherkapazität bei der RCWA direkt ein kleineres Auflösungsvermögen der Struktur. Die Grundursache liegt in der Kopplung von Fourierreihenkoeffizienten mit der Felddarstellung. Ist für eine gegebene komplizierte Struktur eine hohe Anzahl an Ordnungen nötig, um sie als Fourierreihe genau genug darstellen zu können, ist die Größe der Matrizen in der RCWA direkt daran gekoppelt. Ein iteratives Verfahren wie die FDTD kann dieses Grundsatzproblem in geschickter sequentieller Implementierung umgehen. An dieser Stelle wurde allerdings als Vorteil der RCWA auch aufgezeigt, dass die RCWA im Vergleich zur FDTD die Berechnung einer Zelle auf einem Desktop-PC teilweise um den Faktor 8 schneller durchführen kann.

Ferner wurde die RCWA-Zellensimulation auf ein reales Problem, in diesem Fall der Auslegung von Metalinsen angewandt. Aus diesem Problem heraus wurde ein Ansatz entwickelt und vorgestellt, der es ermöglicht, für beliebig örtliche begrenzte Strahlformen, die auf eine in der Größenordnung der Wellenlänge und darunter strukturierte Oberfläche, einfallen, eine Feldlösung im gesamten Gebiet unter oder über der Struktur anzugeben.

Im Falle der Metalinsen wurde dieser Ansatz direkt mit Forschungsergebnissen einer Veröffentlichung verglichen, in welcher die simulierte Linse tatsächlich hergestellt und im Labor vermessen wurde. Die Ergebnisse der Veröffentlichung und der hier vorgestellten Simulation stimmten exakt überein. Ferner wurde auch das dort publizierte FDTD-Simulationsergebnis einer Metalinse fehlerfrei nach-

gebildet. Hierbei waren gravierende Vorteile des in dieser Arbeit vorgestellten Ansatzes bemerkbar, da die Simulationen von Linsen mit einem Durchmesser von 3mm mit der RCWA-Implementierung in der Größenordnung von Minuten bei Berechnung auf einem Desktop-PC liegen, die Autoren des Papers jedoch darauf hinwiesen, dass sie mit $24\mu\text{m}$ die maximale Größe der FDTD für den von ihnen verwendeten Arbeitsgruppenrechner erreicht hatten. Somit erweist sich diese Implementierung als gerade für die Auslegung von Metalinsen sehr nützliches und schnelles Hilfsmittel.

Dennoch darf an dieser Stelle die nicht physikalische Charakteristik der Grenzflächenübergänge an den Zellrändern nicht vergessen werden. Es wurde gezeigt, dass bei Anwendungen wie Metalinsen dieses Konzept zum richtigen Ergebnis führt. Weiterhin konnten bspw. durch die Variation der Flächengröße, in welcher die Fouriertransformation vollzogen wird, Feldartefakte wie bspw. Ecken, die durch die falschen Grenzübergänge erzeugt werden, behoben werden. Dennoch ist ersichtlich, dass bspw. gleiches Vorgehen bei einer Struktur mit Zellen in der Größenordnung der Wellenlänge und Zellenanzahlen weit unter der Elementenzahl in der Metalinsenanwendung nicht zum erwarteten Ergebnis führen würde.

Der Ansatz angewandt auf Gaußstrahlen zeigte Ergebnisse, die analytischen Vorstellungen wie Symmetrieforderungen oder dem Grundverhalten von Gittern bezüglich ihrer Ordnungen entsprechen. Jedoch wurde hierzu noch kein Anwendungsfall entdeckt, mit dem ein Vergleich der Simulationsergebnisse vollzogen werden könnte. Dies ist jedoch für die Zukunft geplant.

Für die bestehende reine RCWA-Implementierung ist des Weiteren vorgesehen, die in Kapitel 4.1 kurz erwähnte Normalenvektortransformation hinzuzunehmen. Dieses Prinzip beruht darauf, jeden Punkt einer Zelle, deren Geometrie in z-Richtung nur durch einen Multilayeransatz angenähert werden müsste, mit Hilfe von Kurvenverläufen in ein lokales Koordinatensystem zu transformieren. In diesem neuen Koordinatensystem werden dann Lösungen nach gleichem Prinzip ermittelt. Somit wird nur eine einzige Schicht zur Simulation verwendet. Auch kann diese Transformation in der xy-Ebene angewandt werden, um beispielsweise das Grundproblem der Annäherung eines Kreises durch Rechtecke zu beheben. Von dieser Erweiterung wird eine erhebliche Verbesserung des Darstellungsproblems mit Fourierkoeffizienten in der RCWA, wie es am Vergleich mit der FDTD klar wurde, erhofft. Ferner ist eine Erweiterung zur anisotropen Formulierung für Strukturen, deren Permittivitäten nicht mehr nur durch skalare Werte, sondern durch Tensoren beschrieben werden können, angedacht.

6 Literaturverzeichnis und weitere Quellen

Literatur

- [1] T.K. Gaylord and M.G.Moharam: *Planar Dielectric Grating Diffraction Theories*; Applied Physics B, Atlanta; 1982
- [2] M. Khorasaninejad, A. Y. Zhu, C. Roques-Carmes, W. T. Chen, J. Oh, I. Mishra, R. C. Devlin and F. Capasso: *Polarization-Insensitive Metalenses at Visible Wavelengths* ; Nanoletters, Massachusetts; 2016
- [3] M. G. Moharam, Eric B. Grann, and Drew A. Pommet, T. K. Gaylord: *Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings*; J. Opt. Soc. Am., Florida; 1994
- [4] M. G. Moharam, Eric B. Grann, and Drew A. Pommet, T. K. Gaylord: *Stable implementation of the rigorous coupled-wave analysis for surface-relief gratings: enhanced transmittance matrix approach*; J. Opt. Soc. Am., Florida; 1994
- [5] Thomas Schuster, Johannes Ruoff, Norbert Kerwien, Stephan Rafler and Wolfgang Osten: *Normal vector method for convergence improvement using the RCWA for crossed gratings* ; J. Opt. Soc. Am., Stuttgart; 2007
- [6] Lifeng Li: *Use of Fourier series in the analysis of discontinuous periodic structures* ; J. Opt. Soc. Am., Arizona; 1996
- [7] Optiwave Photonic Software: <https://optiwave.com/resources/academia/free-fdtd-download/>
- [8] Shun-Der Wu, Thomas K. Gaylord, Elias N. Glytsis and Yu-Ming Wu: *Three-dimensional converging-diverging Gaussian beam diffraction by a volume grating* ; J. Opt. Soc. Am., Atlanta; 2005
- [9] Roman Antos: *Fourier factorization with complex polarization bases in modeling optics of discontinuous bi-periodic structures* ; J. Opt. Soc. Am., Praha; 2009
- [10] Maximilian Auer: *Numerical treatment of localized fields in rigorous diffraction theory and its application to light absorption in structured layers* ; Dissertation, Universität Heidelberg; 2016
- [11] M. N. Polyanskiy: *Refractive index database* ; <https://refractiveindex.info>.
- [12] Testversion Lumerical FDTD: <https://lumerical.com>
- [13] M. Cadilhac “Some Mathematical Aspects of the Grating Theory”; Electromagnetic Theory of Gratings, Edited by R. Petit, Vol. 22. of Topics in Current Physics, Springer Berlin Heidelberg, 1980, pp. 53–62.